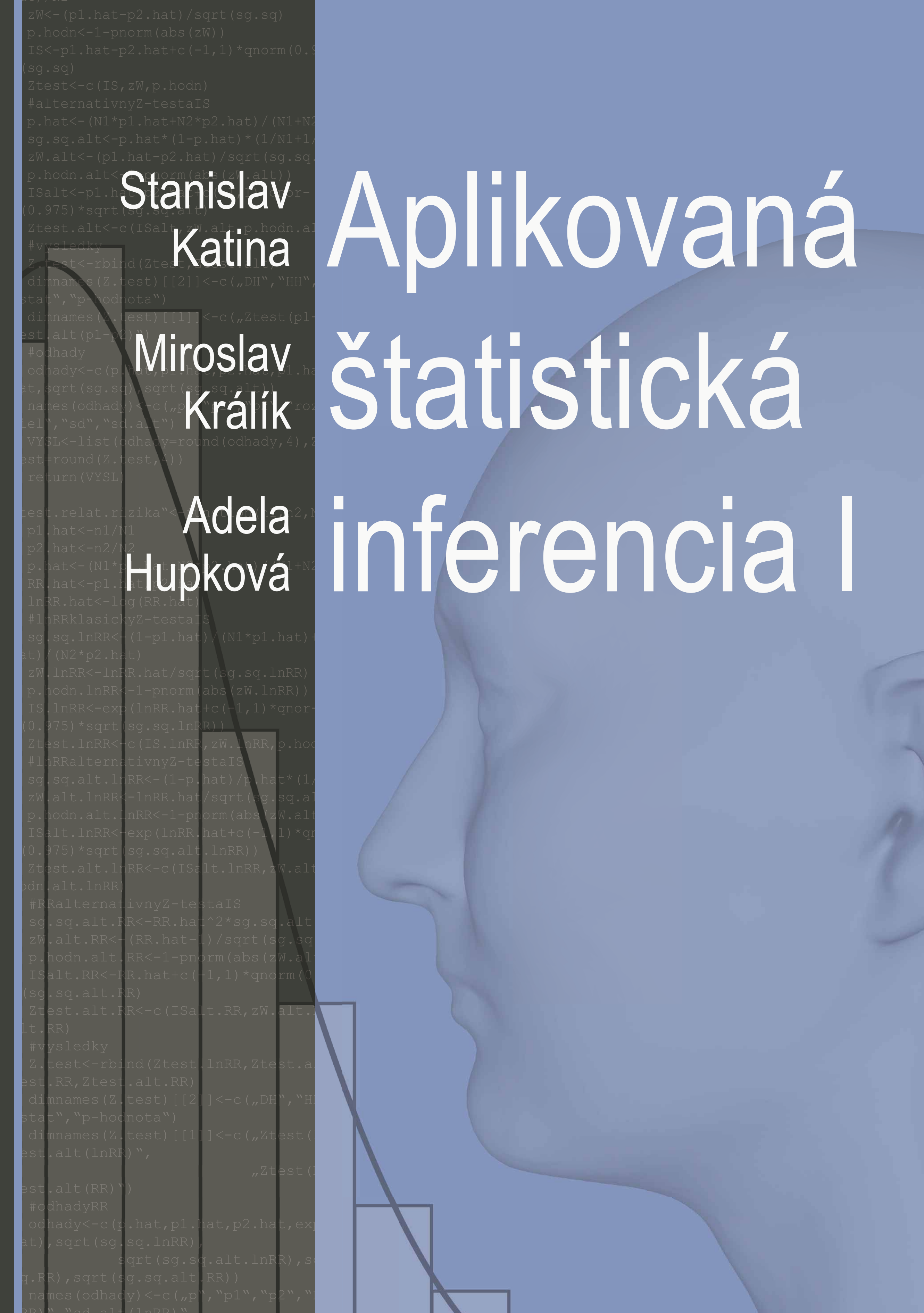
Stanislav Katina

Miroslav Králík

Adela Hupková

Aplikovaná štatistická inferencia I



Aplikovaná štatistická inferencia I

Biologická antropológia očami matematickej štatistiky

Stanislav Katina
Miroslav Králík
Adela Hupková

Masarykova univerzita, 2015

Aplikovaná štatistická inferencia I
Biologická antropológia očami matematickej štatistiky

Autori © Stanislav Katina, Miroslav Králík, Adela Hupková
Recenzenti prof. RNDr. Gejza Wimmer, DrSc.; doc. RNDr. Milan Thurzo, CSc.
Jazyková korektúra Mgr. Lenka Moravčíková
Obálka © Adela Hupková
Sadzba, typografická úprava Stanislav Katina
Vydavateľstvo © Masarykova univerzita, 2015, 1. vydanie

ISBN 978-80-210-7752-2 (brožovaná vazba)
ISBN 978-80-210-7841-3 (online : pdf)

Príspevok jednotlivých autorov:

1. doc. PaedDr. RNDr. Stanislav Katina, Ph.D. – vytvoril koncepciu knihy, navrhol štruktúrované riešenia príkladov a štruktúrované informácie o dátach, vyriešil a naprogramoval príklady a obrázky k nim v R a napísal text všetkých kapitol v LaTeX-u, vytvoril tabuľky funkcií s ich definíciami, vytvoril register, zoznam grafov, tabuliek, matematických symbolov a skratiek;

2. doc. RNDr. Miroslav Králík, Ph.D. – napísal antropologicky a metodicky orientované časti textu, pripravil dáta a štruktúrované informácie o nich, spolupodieľal sa na definícii hypotéz v príkladoch a interpretácii výsledkov štatistických testov a vyvodzovaní záverov;

3. Mgr. Adela Hupková – podieľala sa na tvorbe antropologicky a metodicky orientovaných častí textu, ich preklade do slovenčiny a prepise do LaTeX-u, na príprave dát, štruktúrovaných informácií a grafických podkladov a vytvorení tabuliek funkcií.

Autori by sa radi poďakovali:

- recenzentom prof. RNDr. Gejzovi Wimmerovi, DrSc. a doc. RNDr. Milanovi Turzovi, CSc. za starostlivé prečítanie a korekciu textu knihy;

- Mgr. Ivete Selingerovej za korekcie kapitol 2 a 4, Mgr. Kateřine Konečnej a Mgr. Dagmar Lajdovej za korekcie kapitol 4, 6, 7 a 8;

- Mgr. Tomáši Mořkovskému za všestrannú organizačnú podporu.

Prípravu a vydanie tejto publikácie podporil projekt Formování mezinárodního týmu pro výzkum evoluční antropologie moravských populací (FITEAMP, CZ.1.07/2.3.00/20.0181) z Operačného programu Vzdelávanie pre konkurencieschopnosť, riešený na Ústave antropológie Prírodovedeckej fakulty Masarykovej univerzity.

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Obsah

Úvod . . . xi

1 Vedecké štúdie **1**
1.1 Ciele vedeckej štúdie a sledované premenné . . . 3
1.2 Typy vedeckých štúdií . . . 6
1.2.1 Observačné štúdie vo fyzickej a klinickej antropológii . . . 7
1.2.2 Systematický prehľad . . . 13
1.2.3 Štúdie kostrových pozostatkov . . . 14
1.3 Štatistické hľadiská plánovania štúdií . . . 15
1.3.1 Typy výberov . . . 15
1.3.2 Náhodný výber v realite antropologického výskumu . . . 18
1.3.3 Hlavné atribúty dát . . . 18
1.3.4 Metódy zberu dát . . . 18
1.3.5 Zisťovanie presnosti merania . . . 21
1.3.6 Štatistické znaky . . . 23
1.3.7 Triedenie dát . . . 24
1.3.8 Znáhodnenie a zaslepenie . . . 25
1.3.9 Databáza . . . 25
1.3.10 Rozsah náhodného výberu . . . 27
1.3.11 Kontroly a ich hodnotenie . . . 28
1.3.12 Nábor (získavanie) subjektov vo fyzickej a klinickej antropológii . . . 30
1.4 Úloha bioštatistika pri tvorbe vedeckých štúdií . . . 31

2 Model rozdelenia pravdepodobnosti a štatistický model **33**
2.1 *Simulačný experiment ako nástroj štúdia teoretických vlastností modelov . . . 56
2.2 *Štatistika . . . 59
2.3 *Funkcia vierohodnosti . . . 62
2.4 *Maximalizácia funkcie vierohodnosti . . . 69
2.5 Kritériá klasifikácie štatistických modelov . . . 74
2.6 Praktické dôsledky odchýlok od normality . . . 75

3 Charakteristiky polohy a variability a štatistická grafika **77**
3.1 Charakteristiky polohy . . . 80
3.2 Charakteristiky variability . . . 82
3.3 Detekcia odľahlých pozorovaní . . . 85
3.4 Z-skóre . . . 86
3.5 Príklady na charakteristiky polohy a variability . . . 87
3.6 Štatistická grafika . . . 89
3.6.1 Stĺpcový diagram . . . 89
3.6.2 Spojnicový graf, polygón početnosti a frekvenčná krivka . . . 91
3.6.3 Bodový graf . . . 91
3.6.4 Kruhový diagram . . . 94
3.6.5 Histogram . . . 96
3.6.6 Empirická distribučná funkcia . . . 99
3.6.7 Krabicový diagram . . . 99
3.6.8 Kvantilový diagram . . . 101
3.7 Príklady zo štatistickej grafiky . . . 102

4 Testovanie hypotéz **105**
4.1 *Asymptotické vlastnosti odhadov . . . 109
4.2 Interval spoľahlivosti Waldovho typu . . . 113

4.3 Testovanie H_0 oproti H_1 115
4.4 *Tri typy testovacích štatistík 120
4.5 Vierohodnostné intervaly spoľahlivosti 126

5 Testy dobrej zhody 131
5.1 χ^2 test dobrej zhody 131
5.2 Kolmogorov-Smirnovov test dobrej zhody 137

6 Testovanie hypotéz o jednom parametri 143
6.1 Asymptotické testy o strednej hodnote 143
6.2 Asymptotické testy o rozptyle 159
6.3 Asymptotické testy o korelačnom koeficiente 164
6.4 Asymptotické testy o pravdepodobnosti 174

7 Testovanie hypotéz o dvoch parametroch 189
7.1 Asymptotické testy o rozdiele stredných hodnôt 189
7.2 Asymptotické testy o podiele rozptylov 199
7.3 Asymptotické testy o rozdiele korelačných koeficientov 203
7.4 Asymptotické testy o dvoch pravdepodobnostiach 208

8 Testovanie hypotéz o viacerých parametroch 227
8.1 Asymptotické testy o stredných hodnotách 229
8.1.1 Jednofaktorový ANOVA model s fixnými efektami pri rovnakých rozptyloch . . . 229
8.1.2 Metódy mnohonásobného porovnávania 235
8.1.3 Jednofaktorový ANOVA model s fixnými efektami pri rôznych rozptyloch 248
8.2 Asymptotické testy o rozptyloch 256
8.3 Asymptotické testy o korelačných koeficientoch 260
8.4 Asymptotické testy o pravdepodobnostiach 265

9 Antropologické dátové subory 273
9.1 Dátový súbor – jednovýberový test o strednej hodnote 273
9.2 Dátový súbor – párový test o strednej hodnote 273
9.3 Dátový súbor – párový test o strednej hodnote 275
9.4 Dátový súbor – jednovýberový test o rozptyle 275
9.5 Dátový súbor – jednovýberový test o korelačnom koeficiente 276
9.6 Dátový súbor – jednovýberový test o lineárno-uhlovom korelačnom koeficiente 277
9.7 Dátový súbor – jednovýberový test o uhlovom korelačnom koeficiente 277
9.8 Dátový súbor – jednovýberový test o pravdepodobnosti 278
9.9 Dátový súbor – jednovýberový test o pravdepodobnosti 279
9.10 Dátový súbor – dvojvýberový test o rozdiele stredných hodnôt 279
9.11 Dátový súbor – dvojvýberový test o rozdiele stredných hodnôt 280
9.12 Dátový súbor – dvojvýberový test o podiele rozptylov 280
9.13 Dátový súbor – dvojvýberový test o rozdiele korelačných koeficientov 281
9.14 Dátový súbor – dvojvýberový test o rozdiele pravdepodobností 282
9.15 Dátový súbor – viacvýberový test o stredných hodnotách, nominálna premenná 283
9.16 Dátový súbor – viacvýberový test o stredných hodnotách, nominálna premenná 283
9.17 Dátový súbor – viacvýberový test o stredných hodnotách, ordinálna premenná 285
9.18 Dátový súbor – viacvýberový test o rozptyloch 285
9.19 Dátový súbor – viacvýberový test o korelačných koeficientoch 286
9.20 Dátový súbor – viacvýberový test o pravdepodobnostiach 287
9.21 Dátový súbor – viacvýberový test o pravdepodobnostiach 287
9.22 Dátový súbor – viacvýberový test o pravdepodobnostiach 288
9.23 Dátový súbor – viacvýberový test o pravdepodobnostiach 290
9.24 Dátový súbor – homogenita vektorov pravdepodobností 290

10 Funkcie 293
Literatúra 299
Register 303

Zoznam obrázkov

2.1 Miery normálneho rozdelenia; krivka hustoty s vyfarbeným obsahom pod touto krivkou medzi príslušnými kvantilmi na osi x; obsah je rovný pravdepodobnosti výskytu subjektov s danou výškou v rozpätí týchto kvantilov . 37
2.2 Upravené miery normálneho rozdelenia; krivka hustoty s vyfarbeným obsahom pod touto krivkou medzi príslušnými kvantilmi na osi x; obsah je rovný pravdepodobnosti výskytu subjektov s danou normovanou výškou v rozpätí týchto kvantilov . . . 38
2.3 Aproximácia binomického rozdelenia normálnym pre $p = 0.515$ a $N = 5, 10$ a 50; spojnicový graf superponovaný hustotou (prvý riadok) a distribučnou funkciou (druhý riadok) . . . 41
2.4 Aproximácia binomického rozdelenia normálnym pre $p = 0.1$ a $N = 5, 10$ a 50; spojnicový graf superponovaný hustotou (prvý riadok) a distribučnou funkciou (druhý riadok) . . . 42
2.5 Hustoty dvojrozmerného normálneho rozdelenia pri rôznych parametroch (prvý riadok – kontúrový graf, druhý riadok – perspektívny trojrozmerný graf v podobe plochy); čím je ρ odlišnejšie od nuly, tým viac sa kontúry líšia od kruhov (menia sa na elipsy); so zväčšujúcim sa rozdielom medzi σ_1 a σ_2 sa zväčšuje rozdiel rozptýlenia koncetrických kruhov v smere jednotlivých osí (hovoríme, že rozdiel variability premenných X_1 a X_2 sa zväčšuje) 43
2.6 Hustota dvojrozmerného normálneho rozdelenia s parametrom $\widehat{\boldsymbol{\theta}}$, ktorý je odhadnutý z dát (vľavo) a superimpozícia kontúr hustoty dvojrozmerného normálneho rozdelenia s parametrom $\widehat{\boldsymbol{\theta}}$, ktorý je odhadnutý z dát a dvojrozmerného jadrového odhadu hustoty (vpravo) . . . 44
2.7 Hustoty dvojrozmerného normálneho rozdelenia (prvý riadok $n = 50$; druhý riadok $n = 1000$) . . . 45
2.8 Spoločná hustota dvojrozmerného normálneho rozdelenia (vľavo), hustota zmesi dvoch dvojrozmerných normálnych rozdelení (uprostred) a dvojrozmerný jadrový odhad superponovaný hustotou zmesi dvoch dvojrozmerných normálnych rozdelení (vpravo) – simulačná štúdia . . . 46
2.9 Spoločná hustota dvojrozmerného normálneho rozdelenia (vľavo), hustota zmesi dvoch dvojrozmerných normálnych rozdelení (uprostred) a dvojrozmerný jadrový odhad superponovaný hustotou zmesi dvoch dvojrozmerných normálnych rozdelení (vpravo) – reálne dáta . . . 47
2.10 Hustoty normálneho rozdelenia a zošikmeného normálneho rozdelenia pri rôznych parametroch (prvý riadok); hustoty dvojrozmerného zošikmeného normálneho rozdelenia (druhý riadok vľavo a uprostred) a dvojrozmerného normálneho rozdelenia (druhý riadok vpravo) pri rôznych parametroch . . . 48
2.11 Pravdepodobnostná funkcia a distribučná funkcia $Bin(5, 0.5)$ . . . 49
2.12 Histogram vygenerovaných pseudonáhodných čísel superponovaný spojnicovým grafom teoretickej pravdepodobnostnej funkcie X . . . 58
2.13 Histogram vygenerovaných priemerov superponovaný teoretickou krivkou hustoty $\overline{X}_n$ . . . 59
2.14 Histogram vygenerovaných rozdielov priemerov superponovaný teoretickou krivkou hustoty rozdelenia rozdielu výberových aritmetických priemerov . . . 60
2.15 Histogramy vygenerovaných testovacích štatistík v relatívnej škále superponované s teoretickými krivkami hustoty F; $X \sim N(0, 1)$ (ľavý stĺpec) a $X \sim [(1-p)N(\mu, \sigma^2) + pN(\mu, \sigma_1^2)]$ (pravý stĺpec); $n = 15$ (horný riadok), $n = 100$ (dolný riadok) . . . 62
2.16 Porovnanie škálovaného logaritmu funkcie vierohodnosti (plná čiara) s jeho kvadratickou aproximáciou (čiarkovaná čiara) v prvom riadku a porovnanie škálovanej skóre funkcie a priamky s nulovým interceptom a jednotkovým sklonom v druhom riadku . . . 66
2.17 Profilová funkcia vierohodnosti pre μ (vľavo), σ^2 (uprostred) a funkcia vierohodnosti pre oba parametre (vpravo); $X \sim N(4, 1)$); maximálne vierohodné odhady strednej hodnoty a rozptylu sú označené zvislou čiarkovanou čiarou (vľavo a uprotred) a maximálne vierohodný odhad vektora parametrov je označený $\bullet$ (vpravo) . . . 72
2.18 Funkcia vierohodnosti pre $X \sim Bin(N, p)$ ($p = 0.1, 0.5, 0.9$ a $N = 20$); odhady $\widehat{p}$ sú označené zvislou čiarkovanou čiarou . . . 73
2.19 Logaritmus štandardizovanej funkcie vierohodnosti multinomického rozdelenia v parametroch p_1 a p_2 (Európska populácia) s maximom označeným $\bullet$ . . . 74

3.1 Stĺpcové diagramy – početnosti (prvý riadok) pre vlasy (vľavo), pre oči (vpravo); pravdepodobnosti (druhý riadok) pre vlasy (vľavo), pre oči (vpravo) . . . 92
3.2 Základné typy bodov (dolný riadok) a farieb (horný riadok) . . . 93
3.3 Základné typy čiar – zvislo, vodorovne a v uhle 45° (zľava doprava) . . . 94
3.4 Rozptylový graf pre dĺžku a šírku kališných lístkov . . . 94
3.5 Kruhový diagram (dáta oči vs. vlasy) . . . 95
3.6 Kruhový diagram (odtiene sivej) . . . 96
3.7 Histogram so superponovanou krivkou hustoty normálneho rozdelenia (plná čiara) a hustoty vypočítanej z dát (empirická hustota; čiarkovaná čiara); pod histogramom je tzv. „koberec“ . . . 97
3.8 Dva histogramy s priloženými bázami (základňami) . . . 98
3.9 Hustota so superponovanou hustotou normálneho rozdelenia v podobe 95% pásu spoľahlivosti . . . 98
3.10 Emirická distribučná funkcia (schodovitá krivka) superponovaná krivkou distribučnej funkcie normálneho rozdelenia (hladká krivka) . . . 100
3.11 Bodový graf s marginálnymi krabicovými diagramami (vľavo) a s histogramami (vpravo) . . . 101
3.12 qq-diagram normovanej spojitej premennej výška 10-ročných dievčat (mm) so superponovanou obálkou normálneho rozdelenia . . . 102
3.13 Základná štvorica grafov pre spojitú premennú výška 10-ročných dievčat (mm) – hustota (vľavo hore), kumulatívna distribučná funkcia (vpravo hore), krabicový diagram (vľavo dole), kvantilovy diagram s Atkinsonovou obálkou (vpravo dole) . . . 103
3.14 Základná trojica grafov pre pôrodnú hmotnosť birth.W (dáta two-samples-means-birth.txt) – hustoty (vľavo), kumulatívne distribučné funkcie (uprostred), krabicové diagramy (vpravo) . . . 104

4.1 Grafické znázornenie hustôt normálneho rozdelenia, t-rozdelenia, χ^2-rozdelenia a F-rozdelenia pri rôznych stupňoch voľnosti . . . 107
4.2 Grafické znázornenie významu pravdepodobnosti pod krivkou rozdelenia medzi dvoma kvantilmi (normálne rozdelenie) . . . 108
4.3 Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou normálneho rozdelenia; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom . . . 109
4.4 Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou t-rozdelenia s $df = 10$; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom . . . 110
4.5 Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou χ^2-rozdelenia s $df = 10$; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom . . . 111
4.6 Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou F-rozdelenia s $df_1 = 20$ a $df_2 = 20$; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom . . . 112
4.7 MC experiment pre IS (IS, ktoré neobsahujú $\mu = 0$, sú označené hrubou čiarou); vľavo $X \sim N(0,1)$ a vpravo $X \sim [0.9N(0,1) + (1-0.9)N(0,4)]$ . . . 114
4.8 Silofunkcie asymptotického testu o μ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 125
4.9 Schematický nákres silofunkcií pre $n = 9$ a $n = 36$ . . . 126
4.10 Funkcia vierohodnosti (plnou čiarou) pre p (vľavo) a $g(p)$ vpravo spolu so superponovanými kvadratickými aproximáciami (čiarkovanou čiarou); horizontálna priamka predstavuje hodnotu *cut-off* . . . 129

5.1 Histogram superponovaný s očakávanými hodnotami SAT skóre . . . 133
5.2 Simulované rozdelenie D_n (vľavo) a histogram superponovaný s očakávanými početnosťami výšok (vpravo) . . . 139
5.3 Simulované hustoty rozdelenia D_n pre jednoduchú a zloženú hypotézu . . . 140
5.4 Pásy spoľahlivosti normálneho rozdelenia – pre hustotu (vľavo), distribučnú funkciu (uprostred) a kvantilovú priamku (vpravo) . . . 141

6.1 Silofunkcie asymptotického testu o μ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 145
6.2 Rozptylový graf $\overline{x}_i, s_i$, $i = 1, 2, \ldots, M$, $M = 100000$ pre $n = 5$ (vľavo), $n = 50$ (v strede) a $n = 100$ (vpravo) . . . 146
6.3 Distribučná funkcia (vľavo) a hustoty (vpravo) necentrálneho t-rozdelenia pri rôznych parametroch necentrality λ vyjadreného pomocou δ . . . 147
6.4 Hustota centrálneho a necentrálneho t-rozdelenia; vľavo – hustota necentrálneho t-rozdelenia je superponovaná histogramom simulácií pre $X \sim N(4, 2.5^2)$ a vpravo – pre $X \sim [pN(4, 2.5^2) + (1-p)N(4, 4.5^2)]$, kde $p = 0.9$ . . . 148
6.5 Empirická (krivka s chybovými úsečkami vo vybraných bodoch) a teoretická (hladká krivka) silofunkcia t-testu; simulácie; vľavo – $X \sim N(500, 100^2)$ a vpravo – $X \sim [pN(500, 100^2) + (1-p)N(500, 200^2)]$, kde $p = 0.9$ . . . 149
6.6 Silofunkcie asymptotického testu o σ^2 pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 161
6.7 Silofunkcie asymptotického testu o ρ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 167
6.8 Silofunkcie asymptotického testu o p pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 176
6.9 Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p; $N = 30$ (vľavo), 100 (uprostred) a $N = 1000$ (vpravo) . . . 183
6.10 Pravdepodobnosť pokrytia skóre 95% empirického DIS pre p pri rôznych N . . . 184
6.11 Pravdepodobnosť pokrytia vierohodnostného 95% empirického DIS pre p pri rôznych N . . . 184
6.12 Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p pri $N = 30$; spätne transformovaný DIS pre šancu (vľavo), spätne transformovaný DIS pre logaritmus šance (uprostred) a spätne transformovaný DIS pre arcsin $\sqrt{p}$ (vpravo) . . . 185
6.13 Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p pri $N = 100$; spätne transformovaný DIS pre šancu (vľavo), spätne transformovaný DIS pre logaritmus šance (uprostred) a spätne transformovaný DIS pre arcsin $\sqrt{p}$ (vpravo) . . . 185
6.14 Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p pri $N = 1000$; spätne transformovaný DIS pre šancu (vľavo), spätne transformovaný DIS pre logaritmus šance (uprostred) a spätne transformovaný DIS pre arcsin $\sqrt{p}$ (vpravo) . . . 186

7.1 Silofunkcie Waldovho testu o $\mu_1 - \mu_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 190
7.2 Silofunkcie Waldovho testu o $\mu_1 - \mu_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 192
7.3 Hustota centrálneho a necentrálneho t-rozdelenia; vľavo – hustota necentrálneho t-rozdelenia je superponovaná histogramom simulácií pre $X_1 \sim N(4, 2.5^2)$, $X_2 \sim N(2, 2.5^2)$ a vpravo – pre $X_1 \sim [pN(4, 2.5^2) + (1-p)N(4, 4.5^2)]$, $X_2 \sim N(2, 4.5^2)$, kde $p = 0.9$ . . . 193
7.4 Nelineárny vzťah rozdielu rizík, pomeru šancí a pomeru rizík (vľavo $\mathrm{OR}_{\max} = 0.669$; vpravo $\mathrm{OR}_{\max} = 112.235$ a $\mathrm{OR}_{\min} = 1.494$) . . . 210
7.5 Nelineárny vzťah rozdielu rizík, pomeru šancí a pomeru rizík (vľavo $\mathrm{OR}_{\max} = 0.111$; vpravo $\mathrm{OR}_{\max} = 1003.004$ a $\mathrm{OR}_{\min} = 9.000$) . . . 211
7.6 Silofunkcie Waldovho testu o $p_1 - p_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 211
7.7 Silofunkcie alternatívneho Waldovho testu o $p_1 - p_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo) . . . 212
7.8 Funkcia vierohodnosti a profilová funkcia vierohodnosti binomického rozdelenia pre θ; $n_1 = 5, n_2 = 1$ (prvý riadok), $n_1 = 6, n_2 = 0$ (druhý riadok) . . . 223
7.9 Funkcia vierohodnosti a profilová funkcia vierohodnosti binomického rozdelenia pre logaritmus pomeru šancí zlomenia sa zuba . . . 224
7.10 Profilová funkcia vierohodnosti Poissonovho rozdelenia pre θ a $\theta_j, j = 1, 2, 3, 4$ . . . 225
7.11 Profilová funkcia vierohodnosti pre θ . . . 226

8.1 Silofunkcie ANOVA F-testu pri rôznych J a K – vľavo $J = 3$, uprostred $JK = 12$ a vpravo $JK = 100$ . . . 231
8.2 Hustota centrálneho a necentrálneho F-rozdelenia superponovaná histogramom simulácií pre $Y_j \sim N(\mu_j, \sigma_e^2)$ . . . 232
8.3 Histogramy vygenerovaných testovacích štatistík za platnosti alternatívnej hypotézy v relatívnej škále superponované s teoretickými krivkami hustoty F_W (vľavo) a U_{LR} (vpravo) . . . 235
8.4 Histogramy vygenerovaných testovacích štatistík (za platnosti nulovej hypotézy) v relatívnej škále superponované s teoretickými krivkami hustoty F_W (vľavo) a U_{LR} (vpravo) . . . 236
8.5 Odhadnuté hustoty vybraných t-štatistík T_{LSD} (pre rozdiel populácií A a B) a maximálnych t-štatistík (T_{HSD}) spolu s korešpondujúcimi teoretickými kritickými hodnotami $t_{n-J}(0.025)$, resp. a $q_{J,n-J}(0.05)$ . . . 241

8.6 Kritické hodnoty – Scheffeho $hF_{h,df_e}(\alpha)$, Bonferroniho $F_{1,df_e}(\alpha/h)$, Tukeyho HSD $q^2_{J,n-J}(\alpha)$ a Fisherove LSD $F_{1,df_e}(\alpha)$, $J = 5, 7, 10$, pre rôzne $n < 100$, kde $n - J > 0$ (prvý riadok); upravené hladiny významnosti α_k – Bonferroniho $\alpha_k = \alpha/h$, Holmove $\alpha_k = \alpha/(h-k+1)$, Benjamini-Hochbergove $\alpha_k = k\alpha/h$, Benjamini-Yekutieliho $\alpha_k = k\alpha/(h\sum_{l=1}^{h} 1/l)$, $J = 5, 7, 10$, pre vzrastajúce $k = 1, 2, \ldots, h$ (druhý riadok) . . . 242
8.7 Rozptylové grafy ANOVA modelov – $\mathcal{F}_{H_0}$ (vľavo) a $\mathcal{F}_{H_1}$ (vpravo) . . . 246
8.8 Waldove simultánne 95% empirické IS Tukeyho typu pre rozdiely stredných hodnôt . . . 247
8.9 Krabicové diagramy pôrodnej hmotnosti pre dievčatá a chlapcov . . . 250
8.10 Krabicové diagramy pôrodnej hmotnosti pre dievčatá a chlapcov podľa počtu predchádzajúcich detí biologickej matky . . . 251
8.11 Krabicové diagramy pôrodnej hmotnosti pre dievčatá a chlapcov podľa vzdelania biologickej matky . . . 251
8.12 Waldove simultánne 95% empirické intervaly spoľahlivosti Tukeyho typu pre rozdiely stredných hodnôt pôrodnej hmotnosti (pohlavie dieťaťa vs počet predchádzajúcich detí biologickej matky) – za predpokladu homogenity rozptylov (vľavo) a nehomogenity rozptylov (vpravo); rozdiely v dĺžke toho istého intervalu spoľahlivosti pri rôznych predpokladoch sa líšia od 0.63 po 14.52 v závislosti od odchýlky hodnôt rozptylov použitých na ich výpočet a od rozdielov stupňov voľnosti df_w (newinsorizované dáta) . . . 252
8.13 Waldove simultánne 95% empirické intervaly spoľahlivosti Tukeyho typu pre rozdiely stredných hodnôt pôrodnej hmotnosti (pohlavie dieťaťa vs vzdelanie biologickej matky) – za predpokladu homogenity rozptylov (vľavo) a nehomogenity rozptylov (vpravo); rozdiely v dĺžke toho istého intervalu spoľahlivosti pri rôznych predpokladoch sa líšia od 2.51 po 54.40 v závislosti od odchýlky hodnôt rozptylov použitých na ich výpočet a od rozdielov stupňov voľnosti df_w (newinsorizované dáta) . . . 253
8.14 Krabicové diagramy pre ženy a mužov – dĺžka hlavy (vľavo hore), šírka dolnej čeľuste (vľavo dole), šírka hlavy (vpravo hore) a šírka tváre (vpravo dole) . . . 254
8.15 Krabicové diagramy pre ženy a mužov podľa sexuálnej orientácie – dĺžka hlavy (vľavo hore), šírka dolnej čeľuste (vľavo dole), šírka hlavy (vpravo hore) a šírka tváre (vpravo dole) . . . 256
8.16 Waldove simultánne 95% empirické intervaly spoľahlivosti Tukeyho typu pre rozdiely stredných hodnôt – dĺžka hlavy (vľavo hore), šírka dolnej čeľuste (vľavo dole), šírka hlavy (vpravo hore) a šírka tváre (vpravo dole); newinsorizované dáta . . . 257
8.17 Histogramy vygenerovaných testovacích štatistík (za platnosti nulovej hypotézy) v relatívnej škále superponované s teoretickými krivkami hustoty U_{LR} (vľavo) a U_{B} (vpravo) . . . 260
8.18 Histogramy vygenerovaných testovacích štatistík (za platnosti nulovej hypotézy) v relatívnej škále superponované s teoretickými krivkami hustoty U_{LR} (vľavo) a χ^2_{F} (vpravo) . . . 263
8.19 Logaritmus štandardizovanej funkcie vierohodnosti s maximom označeným • a vierohodnostnou 95% empirickou elipsou spoľahlivosti pre $\boldsymbol{\theta}$ . . . 268
8.20 Stĺpcový diagram relatívnych početností zakončení dlaňových línií (v %) pre každú farbu vlasov. . . . 271

9.1 Znázornenie premenných najväčšia dĺžka mozgovne (skull.L) a najväčšia šírka mozgovne (skull.B) . . . 274
9.2 Znázornenie premennej vertikálny priemer v strede dĺžky tela kľúčnej kosti (simd) . . . 274
9.3 Znázornenie premennej dĺžka kosti kľúčnej z ľavej strany (length.L) . . . 275
9.4 Znázornenie premenných najväčšia výška mozgovne (skull.pH) a morfologická výška tváre (face.H) . . . 276
9.5 Znázornenie premenných výška lebky (skull.H), výška lebečnej bázy (base.H), šírka lebečnej bázy (base.B) a uhol, ktorý zvierajú línie prechádzajúce oboma bodmi *porion* s vrcholom v bode *basion* (base.A) . . . 278
9.6 Znázornenie premenných uhol v bode *nasion* (front.A) a uhol tvárového trojuholníka v bode *prosthion* (prog.A) . 279
9.7 Znázornenie premennej výška lebky (skull.H) . . . 281
9.8 Znázornenie premenných dĺžka dolnej končatiny (lowex.L) a dĺžka trupu (tru.L) . . . 282
9.9 Znázornenie premennej výška hornej časti tváre (upface.H) . . . 283
9.10 Znázornenie premenných dĺžka hlavy (head.L), šírka hlavy (head.W), šírka dolnej čeľuste (bigo.W) a šírka tváre bizyg.W . . . 284
9.11 Znázornenie premennej najväčšia dĺžka kosti kľúčnej z pravej strany (cla.L) . . . 285
9.12 Znázornenie premenných výška nosa (nose.H), šírka nosa (nose.B) a interorbitálna šírka (interorb.B) . . . 286
9.13 Kostný reliéf na vnútornej strane *os pubis* . . . 288
9.14 Číslovanie pozícií na okraji dlane a príklady vysokého (Hi), stredného (Mi) a nízkeho (Lo) zakončenia troch hlavných dlaňových línií (D, C a B) . . . 289

Zoznam tabuliek

1.1 Štatistiky frekvencie vs. typ štúdie (# znamená počet) 11

2.1 Príklady minimálnych N pre fixované p potrebných na aproximáciu 40
2.2 Kontingenčná tabuľka 2×3 pravdepodobností p_j pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (multinomické rozdelenie) 50
2.3 Kontingenčná tabuľka 2×3 očakávaných početností Np_j pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (multinomické rozdelenie) 50
2.4 Kontingenčná tabuľka 2×3 pravdepodobností $p_{j|i}$ pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (súčinové multinomické rozdelenie) 51
2.5 Kontingenčná tabuľka 2×3 očakáných početností $N_i p_{j|i}$ pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (súčinové multinomické rozdelenie) 51
2.6 Kontingenčná tabuľka 3×3 pravdepodobností p_j pre tri farby vlasov a tri farby očí (multinomické rozdelenie) . 52
2.7 Pozorované a očakávané početnosti m_n (zaokrúhlené na nula desatinných miest) Pruských armádnych jednotiek, v ktorých nastalo n úmrtí zapríčinených kopnutím koňom 53
2.8 Pozorované početnosti rodín m_n s n chlapcami 55
2.9 Očakávané početnosti rodín m_n (zaokrúhlené na nula desatinných miest) s n chlapcami (binomické rozdelenie) . 55
2.10 Pozorované početnosti robotníkov m_n s n úrazmi v továrni 56
2.11 Očakávané početnosti robotníkov m_n (zaokrúhlené na nula desatinných miest) s n úrazmi v továrni (Poissonovo rozdelenie) 56
2.12 Simulované a teoretické relatívne početnosti úspechov 57
2.13 Teoretické hodnoty stredných hodnôt a rozptylov S^2 a F a ich odhady zo simulačnej štúdie pri $n = 15$ a $n = 100$ 61
2.14 Očakávané početnosti robotníkov m_n (zaokrúhlené na nula desatinných miest) s n úrazmi v továrni (negatívne binomické rozdelenie) 74

3.1 Zoradné realizácie x_i a ich poradia r_i pre výšky 10-ročných dievčat 81
3.2 Rozsah, aritmetický priemer a smerodajná odchýlka pre surové, urezané a winsorizované dáta (výšky 10-ročných dievčat) 88
3.3 Vybrané charakteristiky polohy a variability pre surové dáta (výšky 10-ročných dievčat) 89
3.4 Vybrané charakteristiky polohy a variability pre najväčšiu dĺžku lebky 89
3.5 Kontingenčná tabuľka 2×3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov 90
3.6 Kontingenčná tabuľka 2×3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi pravdepodobnosťami (multinomické rozdelenie) 90
3.7 Kontingenčná tabuľka 2×3 početností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi početnosťami (multinomické rozdelenie) 91
3.8 Kontingenčná tabuľka 2×3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi stĺpcovými početnosťami (súčinové multinomické rozdelenie; po stĺpcoch) 91
3.9 Kontingenčná tabuľka 2×3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi riadkovými početnosťami (súčinové multinomické rozdelenie; po riadkoch) 91

5.1 Očakávané pravdepodobnosti a početnosti a pozorované početnosti pre SAT skóre 132
5.2 Pozorované početnosti m_n päťsekundových intervalov v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom, v ktorých sa plod ovce n-krát pohol 135
5.3 Očakávané početnosti m_n päťsekundových intervalov v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom (zaokrúhlené na nula desatinných miest), v ktorých sa plod ovce n-krát pohol (Poissonovo rozdelenie) 135
5.4 Očakávané početnosti m_n päťsekundových intervalov v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom (zaokrúhlené na nula desatinných miest), v ktorých sa plod ovce n-krát pohol (ZIP rozdelenie) 136

6.1 Minimálne rozsahy n pri vybraných rozdieloch ρ a ρ_0, kde $\rho_0 = 0$ 167
6.2 Minimálne rozsahy n pri vybraných rozdieloch ρ a ρ_0 spolu s rozdielom $z_R - \xi_0$, ktorý je funkciou ρ a ρ_0 168
6.3 Minimálne rozsahy N pre rôzne p v súvislosti s Haldovou podmienkou 175
6.4 Minimálne rozsahy N pre rôzne rozdiely p a p_0 (kde $p_0 = 0$) v porovnaní s minimálnym rozsahom vypočítaným pomocou Haldovej podmienky (tučne sú zvýraznené tie N, ktoré spĺňajú obe kritériá) 182
6.5 Minimálne rozsahy N pre rozdiel $p - p_0 = 0.1$ pri rôznych p a p_0 v porovnaní s minimálnym rozsahom vypočítaným pomocou Haldovej podmienky (tučne sú zvýraznené tie N, ktoré spĺňajú obe kritériá) 182

7.1 Minimálne rozsahy n pre rozdiel $\rho_1 - \rho_2 = 0.2$ pri rôznych ρ a ρ_0 spolu s rozdielom $\xi_1 - \xi_2$, ktorý je funkciou ρ_1 a ρ_2 205
7.2 Početnosti subjektov s rozšírenými a lokalizovanými metastázami 223
7.3 Odhad relatívneho rizika úmrtia a 95% IS relatívneho rizika úmrtia celkovo a pre každú vekovú skupinu 225
7.4 Početnosti subjektov s infarktom myokardu a mozgovou mŕtvicou v skupine A a B 226

8.1 Experimentálna chyba α_e ako funkcia α_c a h 239
8.2 Koncentrácia stroncia Sr (mg/ml) v piatich vodných celkoch 245
8.3 Výsledky Tukey HSD metódy – rozdiely aritmetických priemerov $\overline{y}_{i\cdot} - \overline{y}_{j\cdot}$, dolná a horná hranica Waldových simultánnych 95% empirických IS Tukeyho typu pre $\mu_i - \mu_j$ (DH a HH), adjustované p-hodnoty $\widetilde{p}_k$ 247
8.4 Adjustované p-hodnoty pre Tukeyho HSD (THSD) metódu, Bonferroniho (B) metódu, Holm-Bonferroniho metódu (HB step-down), Hochberg-Bonferroniho metódu (HB step-up), Benjamini-Hochbergovu (BH) metódu a Benjamini-Yekutieliho (BY) metódu 248

8.5 Pozorované početnosti a pravdepodobnosti a očakávané početnosti a pravdepodobnosti (vypočítané pomocou funkcie vierohodnosti) pre Prahu 268

10.1 Prehľad základných funkcií – charakteristiky polohy a variability a matematické funkcie 293
10.2 Prehľad funkcií súvisiacich s dátovým manažmentom a funkcia library() 294
10.3 Prehľad štatistických a numericko-matematických funkcií 295
10.4 Prehľad funkcií kresliacich rôzne druhy grafov a funkcií s grafmi súvisiacich 296
10.5 Prehľad funkcií súvisiacich s grafmi – popis obrázka a jeho parametre 296
10.6 Prehľad funkcií súvisiacich s rozdeleniami pravdepodobnosti 297

Úvod

Aplikácia štatistických metód a štatisticky orientovaný spôsob uvažovania nebol v biologickej antropológii zaužívaný od jej začiatkov v európskom novoveku, a ani v súčasnosti nepatrí k optimálne rozvinutým oblastiam antropológie. Celý rad štatistických metód pritom vznikol v súvislosti s potrebami vedných odborov zaoberajúcich sa človekom (napr. lineárna regresná analýza vznikla kvôli odhadovaniu výšky postavy človeka na základe dĺžky kostí jeho skeletu, Pearson (1899)). Jedným z dôvodov je po storočia pretrvávajúci typologický prístup k štúdiu človeka, v dôsledku čoho sa na dlhý čas zakonzervoval rasový koncept, ktorý brzdil akékoľvek štúdium ľudskej rozmanitosti a prispôsobivosti. Antropologické štúdie devätnásteho storočia, napísané často na mnoho desiatkach až stovkách strán, do najmenšieho detailu popisovali niekoľko ľudí či niekoľko ľudských kostí, a bez akýchkoľvek údajov o spôsobe vzniku vzorky vyvodzovali ďalekosiahle závery napr. o „rasových“ rozdieloch (cf. Broca, 1862). Aj niektorí zakladatelia a priekopníci biologickej antropológie v storočí dvadsiatom sa využitiu štatistiky bránili. Dokonca i znamenitý Aleš Hrdlička (1869–1943), americký antropológ českého pôvodu, zakladateľ a dlhodobý editor dnes celosvetovo uznávaného odborného časopisu *American Journal of Physical Anthropology*, štatistike neprial a nepodporoval použitie pokročilých štatistických metód v článkoch publikovaných v tomto časopise (Ashley Montagu, 1944). Situácia sa začala meniť v polovici 20. storočia, kedy sa biologická antropológia začala orientovať ekologickým smerom a skúmať ľudské populácie a variabilitu ľudských vlastností v nich i medzi nimi ako výsledok dlhého procesu prispôsobovania sa rôznym podmienkam prostredia (Mielke a kol., 2011, str. 3–22). Populačné vzorky, spôsob ich vzniku a ich vlastnosti (populačné mikroevolučné procesy) sa stali pre interpretáciu výsledkov a vyvodzovanie záverov zásadné, rovnako tak ako sa štatistika stala dôležitým nástrojom zabezpečujúcim platnosť (štatistickú validitu) antropologických záverov. Už prvá česko-slovenská učebnica antropológie má kapitolu zoznamujúcu s podstatou základov štatistického uvažovania (Suchý, 1967), nasledovaná vznikom niekoľkých ďalších prác o štatistike a bioštatistike, vhodnej pre antropológov, biológov a lekárov (napr. Zvára, 1999, 2001; Komenda, 2000). Aplikácia štatistických metód je dnes súčasťou takmer každej odbornej práce v biologickej antropológii, pričom pokroky v rozvoji štatistických metód umožňujú výber vhodnej metódy prispôsobenej skúmaným javom a komplexné usporiadanie štatistických porovnávaní. Výučba základných štatistických metód patrí medzi štandardné súčasti štúdia biologickej antropológie. Napomáha adeptom naučiť sa pristupovať k riešeným problémom koncepčne od samotného začiatku, t.j. počiatočného nápadu a návrhu výskumu až po rozhodovanie o závažnosti nimi zistených výsledkov. Ako v prípade každého nástroja aj v prípade štatistických metód platí, že pri správnom použití dobre spĺňajú svoj účel, ale pri nesprávnom použití môžu výsledky výskumu v neznámej miere strácať svoju platnosť.

Dnešný antropologický výskum sa bez uplatnenia štatistického uvažovanie nezaobíde. I keď sa väčšina biológov, lekárov ani antropológov štatistikmi nikdy nestanú (až na výnimočné personálne únie), výsledky výskumov i možnosti ich publikácie na správnom usporiadaní a štatistickom spracovaní dát silno závisia. Kľúčovým pre obojstranne úspešnú spoluprácu preto je na strane týchto odborov porozumieť základom štatistických postupov, vedieť sa vyjadrovať aspoň základnou štatistickou terminológiou a vedieť štatistikovi vysvetliť svoj výskumný zámer; na strane štatistika potom pochopiť podstatu zamerania jednotlivých odborov, charakter ich metodológie a typické problémy, s ktorými sa po štatistickej stránke stretávajú – a vytvoriť si vzťah obojstranného porozumenia a prospešnosti. Jedným z cieľov tejto knihy je podporiť kladný prístup mladých ľudí k využívaniu štatistických metód v antropológii a ďalších odboroch a umožniť ich aplikáciu v prostredí programu R, univerzálneho výpočtového prostredia, voľne dostupného všetkým (R Development Core Team, 2013).

Táto kniha nie je ani „kuchárskou knihou“ (pretože matematická štatistika nie je knihou receptov alebo návodov), nie je ani „teoretickou matematickou štatistikou“ (pretože teória je vysvetlená prostredníctvom príkladov pochádzajúcich z reality praxe biologickej antropológie) a nie je ani „učebnicou prostredia R“ (pretože program R je v knihe použitý len ako nástroj, prostredníctvom ktorého sú riešené reálne situácie). Čím teda kniha je? **Kniha je unikátnou kombináciou sily teoretického základu matematickej štatistiky implementovaného v prostredí R s cieľom pochopiť a riešiť praktické situácie z biologickej antropológie (ako aj biológie a medicíny).** Prečo práve R? Pretože R naozaj stojí za to. Nie je to len „užívateľsky neprispôsobivý“ komerčný klikací softvér slúžiaci na riešenie vybraných problémov z aplikácií s obmedzeným množstvom použiteľných štatistických metód (pokiaľ sa nejaká metóda v takomto sofvéri nenachádza, je nutné ju hľadať inde). Je to voľne šíriteľné komplexné prostredie slúžiace na štatistické výpočty, prostredie schopné prispôsobiť sa požiadavkám užívateľa. Je to interaktívne prostredie s príťažlivou statickou, ako aj animačnou grafikou. Je to prostredie, v ktorom je možné riešiť nekonečné množstvo praktických, ako aj teoretických štatistických problémov. R je „všetko v jednom“ a naozaj stojí za to!

Kniha je napísaná na základe dlhoročných skúseností prvého autora z prednášania bioštatistických predmetov na štyroch univerzitách (Univerzita Komenského, University of Vienna, University of Glasgow a Masarykova univerzita) pre študentov antropológie, medicíny, aplikovanej matematiky a štatistiky. Je jeho víziou komplexného prístupu k matematickej štatistike pozostávajúceho z dvoch trojpiliérových systémov

1. (a) matematicko-štatistická teória, (b) implementácia v R a (c) aplikácia na reálne situácie a
2. (a) plánovanie bio-medicínskych štúdií, (b) databáza a dátový manažment a (c) interpretácie.

Je členená do deviatich kapitol. Prvá kapitola Vedecké štúdie obsahuje informácie o plánovaní bio-medicínskych štúdií, ktoré by mal vedieť prakticky používať každý vedec pracujúci s reálnymi dátami. Zahŕňa témy ako typy vedeckých štúdií, teória a prax náhodného výberu, presnosť merania, znáhodnenie a zaslepenie, pravidlá tvorby bio-medicínskych databáz („Achilovej päty“ výskumu), plánovanie rozsahu náhodného výberu („piliéra“ plánovania) a pod. Druhá kapitola podáva prehľad filozofie pojmov „model rozdelenia pravdepodobnosti“ a „štatistický model“, z ktorých sú postavené „pevné základy“ výskumu na reálnych dátach. Nachádza sa v nej prehľad vybraných, najčastejšie sa vyskytujúcich modelov ako model binomického rozdelenia, Poissonovho rozdelnia, multinomického a súčinového multinomického rozdelenia, negatívne-binomického rozdelenia, model normálneho rozdelenia a pod. Parametre týchto modelov sú podkladom „štatistického výkladového slovníka“, ktorý slúži na interpretácie výsledkov. Tretia kapitola Charakteristiky polohy a variability a štatistická grafika obsahuje metódy základnej číselnej a grafickej exploratórnej analýzy. Charakteristiky polohy a variability zjednodušene charakterizujú dáta. Štatistická grafika je „obrazom v ráme“, prostredníctvom ktorého sú dáta a výsledky štatistických analýz prezentované. Štvrtá kapitola Testovanie hypotéz komplexne zahŕňa najdôležitejší „pojmotvorný štatistický aparát“, bez znalosti ktorého nie je možné robiť žiadnu štatistickú analýzu, t.j. ide o „piliér“ štatistických analýz a ich interpretácií. Ďalšie štyri kapitoly sú uceleným prehľadom najčastejšie používaných štatistických metód. Ide o systematiký prehľad jednovýberových, dvojvýberových a viacvýberových praktických situácií. Tieto kapitoly sú základom štatistickej inferencie, prostredníctvom ktorej sú zovšeobecnené a interpretované výsledky analýz. Teoretický úvod vo väčšine prípadov obsahuje (1) teoretické predpoklady, (2) matematickú definíciu hypotéz, (3) testovaciu štatistiku (Waldovu testovaciu štatistiku, testovaciu štatistiku pomerom vierohodnosti alebo skóre testovaciu štatistiku), (4) jej rozdelenie za platnosti nulovej a alternatívnej hypotézy, (5) definíciu kritického oboru a silofunkcie, (6) definíciu p-hodnoty, (7) definíciu empirických intervalov spoľahlivosti, (8) algoritmus výpočtu minimálneho rozsahu súboru a (9) odvodenie vybraných typov testovacích štatistík. Po teoretickom úvode nasledujú riešené didaktické príklady na simulovaných alebo reálnych dátach (vysvetľujúce vlastnosti testov alebo demonštrujúcimi ich použitie) nasledované **riešenými reálnymi príkladmi. Ich riešenia sú štruktúrované do špecifických odsekov špeciálne vyvinutých pre túto knihu** – (1) slovná a matematická formulácia hypotéz, (2) testovacia štatistika, (3) zamietacia oblasť, (4) empirický dvojstranný interval spoľahlivosti, (5) štatistický záver, (6) slovný záver a (7) antropologický slovný záver. **Podobne sú v štrukturovaných odsekoch prezentované aj informácie o dátach** – (1) hodnotený súbor, (2) súbor dát, (3) popis premenných, (4) biologické súvislosti a (5) ciele. Zadania príkladov vhodné pre antropológov (biológov alebo lekárov) sú podfarbené sivou farbou. Matematicky náročnejšie kapitoly sú označené hviezdičkou. Definície, vety a dôležitá teória sú vľavo zvýraznené hrubou vertikálnou čiarou. Riešenia príkladov v R sú zvýraznené iným typom písma, jednotlivé riadky kódu sú očíslované a poznámky odlíšené sivou farbou. Kniha spolu obsahuje 273 príkladov (približne 90 % riešených), 1633 riadkov R-kódu, 107 obrázkov, 48 tabuliek a 24 antropologických dátových súborov. Súčasťou knihy je aj tabuľkový prehľad použitých a naprogramovaných funkcií a ich stručné definície.

1 Vedecké štúdie

V procese poznania má teória nezastupiteľné miesto a stojí vždy na počiatku. Reprezentuje tradičný pohľad, ktorý v danej chvíli vysvetľuje väčšinu javov a ich súvislostí. Teória, ktorú sa spravidla naučíme, predstavuje sféru javov možných a overených, v jej rámci sa pohybujeme pri formulovaní pracovnej hypotézy. Posun v poznaní však umožňuje až novinka (inovácia), ktorá do jestvujúcej teórie (tradície) nezapadá. Veda je v tomto pohľade nikdy nekončiace posúvanie hranice medzi tradíciou (teóriou, „sedliackym rozumom“, sférou vecí možných a uznávaných) a informáciou (novinkou). Vedecká metóda má zaistiť, aby sa posúvanie tejto hranice dalo v reálnom živote/svete uskutočniť, a aby nesmerovalo nežiaducim smerom vplyvom šumu (náhody, nesystematických javov). Teóriu/tradíciu potrebujeme poznať preto, aby sme dokázali v nových empirických dátach spoznať chyby, metódu potrebujeme na to, aby sme dokázali odlíšiť novinky od chýb.

Veda prináša systematizované poznanie, predstavuje aplikáciu štandardných metód na štandardné (a neštandardné) problémy. Zmyslom vedeckej metódy je predovšetkým zaistiť, aby vlastný postup získavania dát a ich interpretácie nevnášal do výsledkov skutočnosti, ktoré v reálnom svete neexistujú, a aby naše výsledky mali atribúty vedecky podložených výsledkov. Na rozdiel od iných (viac či menej legitímnych) spôsobov ľudského poznania (subjektívne myslenie, tušenie, viera) sa veda vyznačuje tým, že jej metódy a poznatky sú interpersonálne vymeniteľné (naučiteľné) a opakovateľné. Základným nárokom na vedecký výsledok je podmienka, aby rovnakým postupom mohol ktokoľvek a kedykoľvek v budúcnosti dosiahnuť rovnaké výsledky (čo je ideálna situácia).

Základnými vlastnosťami každej vedeckej štúdie by mala byť teoretická zakotvenosť, inovatívnosť, interpersonálna vymeniteľnosť a opakovateľnosť. Každá vedecká štúdia by mala byť pred realizáciou jasne a zrozumiteľne opísaná. Jej opis musí obsahovať dostatočné detaily v podobe tzv. **plánu (dizajnu) vedeckej štúdie** (**protokolu vedeckej štúdie**) s rozsahom 5 až 25 strán. Takýto plán musí zahŕňať jeho skrátenú, prehľadnú a kompaktnú verziu v podobe **anotácie**[1] v rozsahu 200−400 slov alebo **téz**, z ktorých možno získať rýchly prehľad o šúdii v netechnickom jazyku zrozumiteľnom aj mimo vedeckého odboru, do ktorého štúdia patrí (Hackshaw, 2011). Plán a anotáciu možno usporiadať do nasledovných odstavcov, s každým v rozsahu jednej až troch viet, obsahujúcich len kľúčové fakty (pre každý typ štúdie je potrebné vybrať vhodnú podmnožinu odstavcov):

1. **Biologické súvislosti/Pozadie štúdie a jej význam** – informácie o cieľoch, význame a dôležitosti štúdie a o poznatkoch z danej problematiky.

2. **Dizajn štúdie** – informácie o zvolenom type štúdie (pozri kapitolu 1.2 Typy vedeckých štúdií), o počte sledovaných subjektov, v tejto časti sa často vyskytuje aj zoznam použitých štatistických metód (môže však byť uvedený aj samostatne; pozri ďalej).

3. **Subjekty/skúmané osoby/účastníci štúdie** – informácie o skúmaných subjektoch (zdroj subjektov – aké subjekty, odkiaľ a ako budú vyberané) a ich počte (počet prípadov a kontrol spolu s opisom párovania) a o základných vyšetreniach alebo vykonaných zákrokoch (v antropológii napr. meranie živého človeka, meranie kostí z kostrových sérií, morfometrická analýza obrazu (fotografie, 2D snímky, 3D modely)).

4. **Ciele** – často nazývané aj **Účel/Účely**, príp. **Biologicky formulované hypotézy** – informácie o tom, akú vedeckú otázku/otázky bude štúdia riešiť; či budú primárne alebo sekundárne.

5. **Premenné** – zoznam sledovaných premenných v zmysluplných skupinách, ich počet a definície
   - **Závislé premenné**[2] – primárne a sekundárne – presne korešpondujúce s cieľom, ich počet a typ, rozdelenie do zmysluplných skupín.

[1] Doslovný preklad anglického termínu abstract nie je presný, pretože pod abstraktom chápeme stručné zhrnutie už hotového textu (napr. článku), ale pod anotáciou rozumieme stručné zhrnutie plánovaného projektu/výskumu.

[2] Často štatisticky modelujeme (napr. lineárnym regresným modelom) kauzálny vzťah nezávislých a závislých premenných, kde sa nezávislé premenné nazývajú aj prediktory, pretože na základe nich odhadujeme hodnoty závislej premennej vnútri rozsahu hodnôt prediktorov alebo predpovedáme (predikujeme) hodnoty závislej premennej mimo rozsahu prediktorov.

- **Nezávislé premenné/Prediktory** – označenie typu meraných premenných, stanovenie metódy/metód a frekvencie meraní.

6. **Klinické vyšetrenia subjektov/skúmaných osôb** – pokiaľ ide o klinickú štúdiu (pozri ďalej).

7. **Zákroky** – údaje o tom, kedy a aký typ randomizácie (pozri ďalej) subjektov do skupín bol použitý.

8. **Trvanie** a prípadne ďalšie časové aspekty štúdie.

9. **Opis štatistickej analýzy** – štatistický model/modely, model/modely rozdelenia pravdepodobnosti, matematicky formulované hypotézy[3], výpočet minimálneho rozsahu súboru a pod. (nebýva súčasťou téz, ale vlastného plánu štúdie).

Definícia 1 (subjekt vs. výskumník) ***Subjekt*** *(skúmaný) je osoba/človek, štatistická jednotka, ktorá sa zúčastní výskumu (v experimente je i cieľom intervencie) a je zdrojom hodnotených údajov/dát o sledovaných premenných, ktoré sa získavajú (napr. meraním alebo vyšetrením) a štatisticky spracovávajú* ***výskumníkom*** *(skúmajúcim). Subjektom sa niekedy hovorí aj participanti; subjektom vypĺňajúcim dotazník aj respondenti.*

Príklad 1 (závislé a nezávislé premenné) *Lorencová a Beneš (1976) uskutočnili prierezovú antropologickú štúdiu (pozri kapitolu 1.2 Typy vedeckých štúdií) zameranú na zistenie faktorov ovplyvňujúcich u mužov (adolescentov a dospelých) zloženie tela v priebehu života. Nameranými* ***primárnymi závislými premennými*** *boli hrúbky troch kožných rias (tricepsovej, subskapulárnej, suprailiakálnej) merané kalibrom typu Harpenden, z ktorých boli následne podľa rovnice* $y = 1,87 + 0,398x$ *(Pařízková, 1973) vypočítané* ***sekundárne závislé premenné*** *– percentuálny obsah tuku v tele a percentuálne zastúpenie netukového tkaniva ako doplnok do 100 % (x – súčet hrúbok troch kožných rias v mm, y – % tuku). Autori sledovali ovplyvnenie percenta tuku* ***nezávislými premennými****, medzi ktoré zaradili vek (resp. vekové kategórie od 16 do 64 rokov), frekvenciu príjmu stravy za deň (2 a menej, 3, 4 a viac), zamestnanie z hľadiska fyzickej námahy (učni, študenti, nemanuálne pracujúci, ľahko manuálne pracujúci a ťažko fyzicky pracujúci) a pravidelnú športovú aktivitu (pravidelne športujúci, nepravidelne športujúci, nešportujúci). Z výsledkov štúdie vyplýva, že percentuálny obsah tuku v tele sa s vekom zvyšuje a v rozmedzí 4. až 6. decénia je relatívne stály. Ďalej zistili negatívnu závislosť percenta tuku na frekvencii jedla a telesnej záťaže, t.j. muži s frekventovanejším príjmom stravy (väčším počtom menších dávok) a vyšším kalorickým výdajom (ťažko fyzicky pracujúci a aktívne športujúci) mali nižší percentuálny obsah telesného tuku.*

Veľmi účelným grafickým znázornením abstraktu a téz je **sieťový diagram**, prehľadne ukazujúci hlavné odstavce štúdie, ako aj:

- hlavné (dve až tri) kritériá výberu subjektov,
- základné vyšetrenia alebo zákroky na subjektoch na začiatku štúdie; v prípade randomizácie subjektov je potrebné uviesť, kedy a aká randomizácia sa vykonala,
- čo sa udeje so subjektmi počas štúdie (protokol sledovania).

Typická sekvencia štúdií je od **opisnej** (deskriptívnej), cez **analytickú** (opis príčinno-následného vzťahu), až po (prevažne v medicíne) **klinickú štúdiu** sledujúcu efekt nejakého zákroku (intervencie).

Príklad 2 (sekvencia vedeckých štúdií) *Opisná štúdia: Aká je priemerná týždenná porcia rýb v strave ľudí s anamnézou koronárnej choroby srdca? Analytická štúdia: Aká je asociácia medzi konzumáciou rýb a rizikom opakovaného infarktu myokardu u ľudí s anamnézou koronárnej choroby srdca? Klinická štúdia: Redukuje liečba tabletami rybieho oleja mortalitu ľudí s anamnézou koronárnej choroby srdca?*

[3] Formulujeme nulovú hypotézu voči alternatívnej, kde sa často snažíme nulovú hypotézu vyvrátiť (princíp falzifikácie); formulácia je striktne previazaná s biomedicínskymi cieľmi; viac pozri v kapitole 4 Testovanie hypotéz.

Príklad 3 (biologická formulácia hypotézy) *Hypotéza:* 50 *až* 69*-ročná žena s anamnézou koronárnej choroby srdca, ktorá užíva tablety s rybím olejom, má nižšie riziko výskytu opakovaného infarktu myokardu oproti tým ženám rovnakého veku, ktoré tieto tablety neužívajú.*

Okrem anotácie (téz) a protokolu štúdie musí každá štúdia obsahovať aj tzv. **operačný manuál**, čo je zhrnutie špecifických inštrukcií, dotazníkov a iných častí vytvorených na zaistenie uniformného a štandardného prístupu a postupu v priebehu celej štúdie s dobrou kontrolou jej kvality.

Plán vedeckej štúdie vychádza z jednej či viacerých pracovných hypotéz, ktorých východiskom sú aktuálne poznatky o skúmanej téme a jej širokom ukotvení v rámci daného odboru, resp. vedeckej paradigmy, t.j. teórie. Každá hypotéza preto vychádza z teórie prijímanej určitou časťou vedeckej komunity a zameriava sa na rozvinutie/naplnenie teórie v oblastiach dosiaľ neznámych a dosiaľ nezodpovedaných otázok. Hypotéza by mala byť navrhnutá tak, aby bolo v podmienkach reálneho výskumu možné učiniť platný pokus o jej vyvrátenie. Žiadnu hypotézu nemožno s definitívnou platnosťou potvrdiť (verifikovať) ani akokoľvek vysokým počtom pozitívnych dôkazov, keďže niekedy v budúcnosti môže nastať situácia, keď sa hypotéza nepotvrdí. Naproti tomu s definitívnou istotou môžeme zistiť, že hypotéza neplatí. Novinku a posun v poznatkoch preto predstavujú predovšetkým situácie, v ktorých sa pracovnú hypotézu podarilo vyvrátiť. Hypotézy by teda mali byť navrhnuté tak, aby sa ich prípadným vyvrátením dosiahol postup v sledovanej otázke.

Definícia 2 (metóda) *(Antropometrická) metóda je plánovitý postup, ktorý vedie k vytýčenému cieľu a má okrem iného zaistiť validitu vedeckej štúdie; tento pojem sa používa vo všeobecnejšom zmysle.*

Definícia 3 (validita vedeckej štúdie) *Validita vedeckej štúdie označuje platnosť, správnosť a hodnovernosť vedeckej štúdie, vrátane jej výsledkov a záverov.*

Validitu štúdie možno rozdeliť na niekoľko úrovní (oblastí/typov). **Konštrukčná validita** odkazuje na to, či použitým postupom skutočne sledujeme (meriame) vlastnosť, ktorú merať chceme. **Interná validita** zabezpečuje, že zaznamenaný rozdiel je skutočne spôsobený sledovaným faktorom. Inými slovami, že zaznamenaná súvislosť je skutočne spôsobená nami predpokladaným kauzálnym vzťahom. **Externá validita** sa týka šírky platnosti výsledkov, t.j. či sú výsledky platné nielen pre použitú vzorku, ale aj pre inú vzorku na inom mieste a v inom čase atď. **Validita štatistických záverov** zaisťuje, že zistený (väčšinou číselný) rozdiel je skutočným systematickým rozdielom a nielen náhodne spôsobenou výchylkou. Použitá metóda by mala zaistiť validitu výskumu na všetkých štyroch úrovniach. Zabezpečiť to je nevyhnutnou súčasťou správne navrhnutej a uskutočnenej štúdie.

Definícia 4 (metodika) *Metodika predstavuje konkrétny návod, algoritmus realizácie použitej metódy.*

Príkladom metodiky je napr. metodika metrického odhadu pohlavia jedinca podľa jeho panvovej kosti.

Definícia 5 (technika) *Technika predstavuje konštrukciu a použitie nástrojov pri výskume.*

1.1 Ciele vedeckej štúdie a sledované premenné

Ciele, účel/účely a hypotézy by mali byť formulované a napísané jasne a zrozumiteľne v netechnickom jazyku. Príkladmi cieľov v klinickej a fyzickej antropológii sú:

- hodnotenie rizikových faktorov, príčin choroby alebo smrti,
- hodnotenie charakteristík, postojov, skúseností alebo správania sa skúmanej skupiny osôb,
- hodnotenie metód prevencie, detegovania alebo liečenia choroby, príp. prevencie predčasnej smrti,
- konanie laboratórnych experiementov na biologických objektoch, zvieratách a pod. s cieľom preskúmať efekty stimulov alebo expozícií (dávok) nejakých látok/liečiv, identifikácie asociácií ako súčasť výskumu liekov a pod.,

- hľadanie vzťahov medzi biologickými premennými, ako napr. hodnotenie genetických a iných biomarkerov asociovaných s chorobou alebo predčasnou smrťou a pod.,
- hodnotenie vplyvu regionálnych, environmentálnych a socio-ekonomických faktorov na telesnú stavbu človeka súčasných alebo minulých populácií,
- hodnotenie vplyvu civilizačných trendov (miera pohybovej aktivity, prevládajúce typy pohybov, množstvo a zloženie potravy) na telesné zloženie, metabolizmus a zdravotný stav človeka,
- antropometrické charakterizovanie súčasných populácií a tvorba referenčných údajov (noriem) a ich aplikácia v ergonómii,
- hľadanie vzťahov medzi biologickými (veľkostnými a tvarovými) a socio-kultúrnymi premennými, napr. sledovanie fyziologických rastových zmien u detí a dospievajúcich, ako aj včasné odhalenie rastových porúch,
- vývin metód aplikovateľných v kriminalistickej a forenznej praxi pri identifikácii jedinca.

Každý cieľ by mal byť opísaný jednou stručnou vetou nezameniteľne s opisom sledovaných závislých premenných, t.j. premenných slúžiacich na opis cieľov. Do štúdie je potrebné zahrnúť realistické a primárne dôležité ciele priamo naviazané na závislé premenné, ktoré by nemali byť príliš komplexné a mali by byť realizovateľné v zmysluplnom časovom horizonte.

Príklad 4 (cieľ vs. závislá premenná) *Príklady cieľov a priamo naviazaných závislých premenných (v podobe „cieľ – závislá premenná") v klinickej a fyzickej antropológii:*

- *zistenie vplyvu statínovej terapie na ľudí bez anamnézy srdcovej choroby – hladina cholesterolu v sére;*
- *zistenie vplyvu statínovej terapie na ľudí s anamnézou srdcovej choroby – percentuálne zastúpenie ľudí s druhým infarktom myokardu;*
- *zhodnotenie efektívnosti novej terapie astmy v porovnaní so štandardnou terapiou – percentuálne zastúpenie pacientov, ktorí trpia silnými astmatickými záchvatmi;*
- *zhodnotenie efektívnosti testu A v porovnaní s testom B v prípade identifikácie ženy čakajúcej dieťa s Downovým syndrómom – percentuálne zastúpenie falošne pozitívnych a falošne negatívnych výsledkov;*
- *zistenie vzťahu tlaku krvi a veku – systolický/diastolický tlak a vek každého sledovaného subjektu;*
- *hodnotenie efektu chemoterapie pri liečení nádoru pľúc – čas prežívania pacientov od randomizácie po smrť z rôznych príčin alebo po dátum, kedy bol pacient ešte nažive (prípadne do cenzurujúceho dátumu, t.j. dátumu, kedy bola štúdia ukončená);*
- *hodnotenie efektu bezpečnej dávky nového lieku na králikoch pred testovaním na ľuďoch – nepriaznivé udalosti (akákoľvek nepriaznivá zmena zdravotného stavu alebo škodlivý účinok zistený napr. krvnými testami);*
- *zistenie rozdielov medzi ľuďmi trpiacimi schizofréniou, bipolárnou poruchou alebo velokardiofaciálnym syndrómom a zdravými ľuďmi – charakteristiky tvaru ľudskej tváre;*
- *hodnotenie miery pohlavného dimorfizmu vo výške postavy – výška postavy mužov a žien;*
- *vekové zmeny v množstve a rozložení telesného tuku v oblasti trupu – hodnota WHR;*
- *zistenie vplyvu laterality na vznik medzistranových rozdielov na hornej končatine – rozmery hornej končatiny.*

Rozlišuje sa hlavný cieľ/ciele a vedľajší cieľ/ciele. **Hlavné ciele** a na ne naviazané závislé premenné sa môžu použiť po skončení štúdie ako indikátory zmeny (rutinnej) klinickej praxe, zdravotníckej politiky, prípadne ako indikátory zvýšenia stavu vedomostí o danej chorobe alebo problematike a pod. **Vedľajšie ciele** často obsahujú doplnkové informácie.

Príklad 5 (hlavný cieľ a vedľajšie ciele) *Hlavný cieľ: hodnotenie efektu nového lieku pri liečení nádoru pľúc vo vzťahu k celkovému prežívaniu. Vedľajšie ciele: (1) hodnotenie efektu nového lieku na zmenšovanie nádoru, (2) sledovanie, či nový liek ovplyvňuje opakovaný výskyt nádoru, (3) zistenie bezpečnosti (toxického profilu) novej liečby, (4) zistenie, či nová liečba pozitívne ovplyvňuje kvalitu života pacienta.*

Výstižne formulované ciele v podobe **vedeckých otázok** možno správne formulovať len vďaka kvalitnému systematickému prehľadu (pozri kapitolu 1.2.2 Systematický prehľad). Vedecká otázka musí mať päť základných ***FINER*** (*Feasible*, *Interesting*, *Novel*, *Ethical*, *Relevant*) atribútov, t.j. musí byť:

1. *reálna* – v zmysle dostatočného počtu subjektov, dostatočných odborných znalostí, finančnej prístupnosti a časovej prijateľnosti, a musí poskytnúť odpoveď;
2. *zaujímavá* – mala by v istom zmysle fascinovať vedcov alebo vzbudzovať ich záujem;
3. *nová* – mala by buď potvrdzovať správnosť predošlých poznatkov a rozširovať ich o nové zistenia, alebo dokazovať ich nesprávnosť;
4. *etická* – prístupná k skúmaniu odpovedí po schválení etickou komisiou inštitúcie;
5. *relevantná* – v zmysle vedeckých znalostí pre teóriu a smerovanie výskumu, ako aj aplikáciu poznatkov v praxi (napr. v klinickej praxi, v sociálnej a zdravotnej politike a pod.).

Príklad 6 (ciele v podobe vedeckých otázok) *Mali by ľudia jesť viac rýb? Ako často jedia ľudia istého regiónu ryby? Znižuje jedenie rýb riziko kardiovaskulárnych chorôb? Spôsobuje zvýšená konzumácia rýb u starších ľudí riziko otravy ortuťou? Majú doplnky stravy v podobe rybieho oleja v tabletách rovnaký efekt na kardiovaskulárne choroby ako vlastné ryby? Ktoré typy doplnkov stravy v podobe rybích olejov nespôsobujú, že ľudský pach sa podobá rybaciemu?*

Správne identifikované ciele a na ne naviazané závislé premenné pomáhajú pri rôznych rozhodnutiach týkajúcich sa dát – napr. ako získať potrebné dáta, ako ich štatisticky analyzovať, ako interpretovať výsledky a nakoniec, ako získané výsledky publikovať.

Sledované premenné musia prejsť

- **internou validizáciou** (**kontrolou kvality**), t.j. *kontrolou kvality databázy*, *kontrolou validity merania* a následným *výpočtom reliability merania*, ktorého súčasťou je výpočet
  - **intraindividuálnej chyby** (rozdiel mezi dvoma (alebo viacerými) opakovanými meraniami rovnakého výskumníka na tej istej vzorke),
  - **interindividuálnej chyby** (rozdiel mezi dvoma (alebo viacerými) meraniami dvoch (alebo viacerých) výskumníkov na tej istej vzorke);
- **externou validizáciou**, ktorá predstavuje napr. zlepšenie plánu štúdie, zväčšenie rozsahu náhodného výberu a použitie stratégií zlepšenia **presnosti** aj **správnosti a výstižnosti merania**. Patria sem:
  - stratégie zlepšenia presnosti: (1) štandardizácia meracích metód v operačnom manuáli, (2) školenie výskumníka, (3) nastavenie prístroja alebo zvýšenie jednoznačnosti dotazníka, (4) automatizácia meracieho prístroja alebo vyplnenie dotazníka samotnými subjektmi (respondentmi), (5) opakovanie meraní a následný výpočet priemernej hodnoty merania;
  - stratégie zlepšenia správnosti a výstižnosti merania (navyše k bodom (1) až (4) v stratégiách zlepšenia presnosti): (1) redukcia systematickej chyby napr. kalibráciou prístroja, (2) porovnanie meraní so „zlatým štandardom“, napr. výpočtom senzitivity a špecificity pre dichotomické merania alebo výpočtom koeficientu kappa pre kategorické premenné s viac ako dvojpoložkovou škálou, (3) zaslepenie príslušnosti subjektu do nejakej skupiny (z hľadiska subjektu aj z hľadiska výskumníka).

Definícia 6 (presnosť merania) *Presnosť merania je stupeň, do ktorého má premenná pri viacnásobnom meraní približne rovnakú hodnotu. Presnosť sa najlepšie hodnotí porovnaním opakovaných meraní. Zvyšovaním presnosti merania sa zvyšuje sila štúdie. Presnosť merania najviac ohrozuje náhodná chyba či už na strane výskumníka, alebo subjektu, ale aj chyba meracieho prístroja.*

Definícia 7 (správnosť a výstižnosť merania) *Správnosť a výstižnosť merania je stupeň, do akého premenná reprezentuje to, čo reprezentovať má. Správnosť a výstižnosť merania sa najlepšie hodnotí porovnaním so „zlatým štandardom“. Zvyšovaním správnosti a výstižnosti merania sa zvyšuje validita záverov štúdie. Správnosť a výstižnosť merania najviac ohrozuje systematická chyba či už na strane výskumníka, alebo subjektu, ale tiež meracieho prístroja.*

1.2 Typy vedeckých štúdií

Rozlišujeme dva hlavné typy vedeckých štúdií – experimentálne a observačné. Na zodpovedanie cieľov musí byť použitá vhodnejšia z nich. Oba typy štúdií môžu byť zahrnuté v systematickom prehľade.

Experimentálna vedecká štúdia sú štúdie, v ktorých sú niektoré alebo všetky subjekty experimentálne podrobené zmeneným a kontrolovaným podmienkam prostredia (dávke žiarenia a pod.) alebo liečbe, s ktorými by normálne neprišli do styku. Tu rozlišujeme:

- klinické pokusy na ľuďoch,
- laboratórne experimenty.

Observačné vedecké štúdie sú štúdie, v ktorých sú študované subjekty sledované v ich prirodzených podmienkach bez zámeru do subjektov akokoľvek zasahovať. Rozlišujú sa:

- kvalitatívne štúdie,
- pacientsky audit,
- prípadové štúdie,
- prierezové štúdie,
- (retrospektívne) štúdie prípadov a kontrol,
- kohortové (retrospektívne/prospektívne) štúdie, v biologickej antropológii nazývané aj longitudinálne štúdie,
- sekvenčné alebo semilongitudinálne štúdie.

Špeciálnym typom štúdií v antropológii sú **štúdie kostrových pozostatkov** – kostrových sérií, ojedinelých kostrových nálezov, historických osobností, múmií, atď.

Príklad 7 (štyri typy štúdií) *Príklady štyroch základných plánov štúdií skúmajúcich, či príjem rýb v strave redukuje riziko vypuknutia srdcovocievnej choroby.*
Klinická (randomizovaná zaslepená) štúdia: Dve skupiny sú vytvorené náhodným (randomizovaným, znáhodneným) a zaslepeným spôsobom. Výskumník náhodne priradí subjekty do skupiny, ktorá prijíma v strave rybí olej a do skupiny, ktorá prijíma placebo, potom sleduje obe skupiny niekoľko rokov z hľadiska výskytu srdcovocievnych chorôb.
Prierezová štúdia: Subjekty sú hodnotené v jednom časovom bode. Výskumník zaznamená subjektom udávaný príjem rýb v strave v minulosti a v súčasnosti a odpovede dáva do vzťahu s prepuknutím srdcovocievnych chorôb a so súčasným výskytom vápnika v krvi.
Štúdia prípadov a kontrol: Vybraní jedinci sú rozdelení do dvoch skupín na základe výskytu ochorenia (závislá premenná). Výskumník hodnotí skupinu pacientov so srdcovocievnymi chorobami (prípady) a porovnáva ju so skupinou bez srdcovocievnych chorôb (kontroly) pýtajúc sa na minulý príjem rýb v strave.

Kohortová štúdia: Subjekty sú sledované v čase. Výskumník zaznamenáva príjem rýb v strave na začiatku štúdie a následne príjem rýb opakovane zaznamenáva počas niekoľkých návštev zisťujúc, či sa u jedincov konzumjúcich viac rýb vyskytovalo menej srdcovocievnych chorôb.

Systematický prehľad (zahŕňajúci **meta-analýzu**) je štúdia, ktorej cieľom je identifikovať všetky publikované a nepublikované štúdie v príslušnej oblasti, ako aj kombinovať a interpretovať ich výsledky. Tento typ štúdie sa najčastejšie používa pri klinických pokusoch, kohortových štúdiách a štúdiách prípadov a kontrol.

Špeciálnym typom štúdie je tzv. **translačná štúdia**, ktorá predstavuje premostenie medzi laboratórnym výskumom a klinickou štúdiou (ozn. ako **T1**), alebo klinickou štúdiou a praxou (ozn. ako **T2**).

Väčšinu štúdií realizovaných v antropológii predstavujú štúdie observačné. Experimentálny výzkum ovplyvňujúci biologické procesy u zdravých ľudí podlieha schváleniu etickej komise a je takmer celkom obmedzený na oblasť klinických štúdií (hlavne to platí pre invazívne vyšetrenia s narušením integrity tela, užívaním testovaných liečiv atď.). Laboratórne experimenty na ľuďoch, dobrovoľníkoch sú však možné v psychológii, behaviorálnych a kognitívnych vedách, pokiaľ nepredstavujú pre dobrovoľníkov zdravotné alebo iné riziká. Možné sú aj napr. tafonomické alebo biomechanické experimenty s kostnými a zubnými tkanivami.

1.2.1 Observačné štúdie vo fyzickej a klinickej antropológii

Observačná štúdia je taká štúdia, v ktorej sa študované subjekty sledujú v ich prirodzených podmienkach a dáta sa získavajú z nemocničných alebo ambulantných záznamov, regionálnych alebo národných registrov (napr. register pacientov s cystickou fibrózou, register operácií bedrového a kolenného kĺbu a pod.), z dotazníkov alebo interview s ľuďmi, z databáz RTG snímok alebo CT skenov. Pri zbere dát je dôležité dbať na presnosť a komplexnosť záznamov, pričom treba rozlišovať, či informácie pochádzajú zo záznamov vypĺňaných samotným pacientom alebo vyšetrujúcim lekárom, či pochádzajú z lekárskych záznamov alebo z priamej diagnózy klinickým expertom, či boli merania vykonané študentom v zácviku alebo skúseným antropológom.

Kvalitatívne štúdie majú zvyčajne málo subjektov (menej ako 50) a sú založené na pozorovaní alebo interview s kladením otázok otvoreného typu. Môžu slúžiť na vývin dizajnu ďalšej štúdie, vývin meracích prostriedkov v podobe dotazníka alebo na zistenie názoru na zdravotnú starostlivosť, príp. liečbu.

Audit pacientov predstavuje extrakciu údajov z minulosti z už existujúcich záznamov pacientov v nemocniciach, slúži na zistenie zákrokov, rizikových faktorov choroby alebo predčasnej smrti. Tu treba mať na zreteli výberovú odchýlku.

Prípadové štúdie, nazývané aj *prípadové správy* sú popisné či iné štúdie, v ktorých je vzorka obmedzená na jeden alebo niekoľko zaujímavých alebo niečím významných prípadov nejakého javu. Dôvodom môže byť ich neexistencia (súčasný výskyt) vo väčšom počte (napr. veľmi vzácna choroba) alebo ťažké štúdium vo väčšom počte (etické obmedzenie, finančné nároky, ap.). Príkladom môže byť *paleopatologická štúdia jedného skeletu s výskytom početných prejavov nejakej choroby na rôznych kostiach* (napr. syfilis). Tiež však môže ísť o *zámernú výskumnú stratégiu*, kedy na rozdiel od relatívne povrchného štúdia jedného alebo niekoľkých znakov u veľkého počtu prípadov (čo je bežná populačná štúdia, niekedy označovaná ako *cross-case study* a predstavujúca väčšinu štúdií citovaných a popisovaných v tejto knihe) študujeme na jednom alebo niekoľkých prípadoch veľké množstvo znakov a súvislostí s cieľom pochopenia komplexity ich vzťahov a ich vnútornej organizácie. Vlastná štúdia nespočíva v sledovaní štatistického rozdielu medzi skupinami alebo súvislostí medzi premennými, ale v detailnom, mnohostrannom popise študovaných prípadov a (bio)logickom vysvetlení povahy javu vo všetkých rozlíšených rovinách/uhloch pohľadu. Existujú tiež metódy, ktorými možno kombinovať výsledky viacerých prípadových štúdií. V širšom slova zmysle možno za prípadovú štúdiu považovať *súčasť alebo prípravu bežnej populačnej štúdie.* Napríklad v longitudinálnej štúdii rastu najskôr vytvoríme rastové krivky pre každé dieťa (t.j. toľko jednotlivých prípadových štúdií, koľko máme longitudálne zaznamenaných detí) a potom štatisticky hodnotíme závislosti parametrov rastových kriviek na nejakom faktore v populačnej štúdii. Pre rozdelenie typov prípadových štúdií a ich dizajn pozri napr. Gerring (2007). Špeciálnym prípadom prípadovej štúdie je aj štúdia porovnania prípadu so starostlivo zvolenou *referenčnou populáciou* (Šefčáková a Katina, 2008; Šefčáková a kol., 2011).

Prierezové štúdie, niekedy nazývané aj *prevalenčné*, sa používajú v prípadoch prieskumu, monitorovania alebo dohľadu. Svoje pomenovanie dostali na základe toho, že vybrané subjekty predstavujú prierez populáciou alebo jej časťou na danom mieste a čase v kontexte skúmaného javu. Používajú sa napr. na zistenie vzťahu rizikových faktorov a chorôb. Údaje možno získavať z lekárskych vyšetrení. Často sa využíva validovaný dotazník alebo dotazník špecificky vytvorený pre danú štúdiu (vývoj tohto dotazníka musí byť presne dokumentovaný a dotazník musí byť validovaný). Základnou úlohou je nájsť rozdiely v *prevalencii* medzi exponovanou a kontrolnou (referenčnou) skupinou. Takáto štúdia informuje o prevalencii v určitom čase. Stanovenie expozícií a meranie osobných charakteristík alebo biologických markerov môže byť *prospektívne* na základe budúcich výskumov alebo *retrospektívne* na základe už existujúcich záznamov. V týchto štúdiách ide o *reprezentatívny výber*, čo umožňuje odhadovať pravdepodobnosť výskytu choroby alebo vystavenie sa expozícii v celej populácii a následne testovať isté hypotézy. Overuje sa napr., či pravdepodobnosť výskytu choroby závisí od expozície. V praxi sú časté náhodné výbery veľkého rozsahu a predmetom záujmu je vzťah vybraných faktorov s expozíciou alebo chorobou.

Prevalenčné štúdie často slúžia na porovnanie prevalencie choroby na rôznych miestach alebo u rôznych populačných skupín vzhľadom na expozíciu. Tieto štúdie môžu skúmať napr. rozdiely v prevalencii respiračných problémov v rôznych skupinách populácie (u ľudí s rôznou expozíciou), môžu sledovať chronické stavy, napr. artritídu, diabetes, hypertenziu, peptický vred, srdcovocievne choroby a mozgové príhody, alebo tiež študovať hygienu práce pri expozícii chemikáliám, prachu a plynom, príp. koncentrácie rôznych látok na pracoviskách (napr. ortute) a pod. Prevalenčné štúdie môžeme využiť pri overovaní prvých pracovných hypotéz ako východisko pre kohortové štúdie. Nie sú vhodné na sledovanie akútneho účinku nových chemikálií a iných škodlivín, ktoré vyžaduje aj kohortovú štúdiu.

V biologickej antropológii predstavujú jeden z najčastejších typov výskumu živého človeka prierezové štúdie, ktoré sa uplatňujú pri zisťovaní výskytu či miery prejavu akéhokoľvek biologického znaku/javu v ľudských populáciách. Uskutočňujú sa u detí aj dospelých. V prípade ľudí rôzneho veku sa získavajú rozdiely medzi vekovými kategóriami, z ktorých možno vytvárať *nepravé rastové krivky* (nemožno ich zamieňať za *pravé rastové krivky* vytvorené na základe longitudinálnych dát). Ak sa na základe prierezovej štúdie hodnotia vekové rozdiely (detí alebo dospelých), vnútorná validita štúdie môže byť ohrozená tým, že vekové skupiny sa môžu líšiť v mnohých iných vlastnostiach nesúvisiacich s vekom, ktoré vyplývajú z rozdielov medzi generáciami. Tento jav sa označuje ako *efekt kohorty.*

Príklad 8 (efekt kohorty; znižovanie výšky postavy s vekom) *V prierezovej štúdii zaznamenaný rozdiel vo výške postavy medzi dvadsaťročnými a osemdesiatročnými ľuďmi bol spôsobený efektom kohorty v populácii, ktorá prekonala sekulárnu zmenu vo výške postavy. Pred šesťdesiatimi rokmi ľudia dosahovali pri ukončení rastu postavy do výšky nižšieho vzrastu (Cardoso, 2008) než dnešní dvadsaťroční ľudia a túto výšku postavy si potom niesli životom ako kohorta. To však neznamená, že v priebehu starnutia nedochádza k menším degeneratívnym zmenám kostí a kĺbov a úbytku výšky, väčšia časť rozdielov je však daná rozdielmi medzi generáciami (Eveleth a Tanner, 1990, s. 208–223).*

Priemerná výška postavy v danej kohorte sa v priebehu starnutia môže meniť pod vplyvom selekcie, ktorá nie je neutrálna voči výške postavy. V každej vekovej kategórii totiž mortalita klesá s výškou postavy. Vyššieho veku sa teda dožíva väčší podiel jedincov s vyššou než s nižšou postavou (Waaler, 1984).

Príklad 9 (efekt kohorty; vekové rozdiely v inteligencii (IQ)) *V minulosti viedol efekt kohorty k nadhodnoteným predstavám o znižovaní inteligencie v priebehu starnutia (Schaie, 2000).*

Prierezová štúdia použitá na porovnanie vekových skupín neumožňuje zachytenie vnútornej štruktúry vekových zmien. Nedovoľuje napríklad odpovedať na otázku, či majú ľudia vo veku dvoch rokov s telesnou výškou na 10. percentile aj v desiatom roku stále výšku odpovedajúcu 10. percentilu populácie.

Štúdie prípadov a kontrol sa zvyčajne používajú na zistenie rizikových faktorov choroby alebo predčasnej smrti. Sledujú sa subjekty, napr. pacienti a k nim prislúchajúce kontroly, retrospektívne alebo prospektívne.

Do ***retrospektívnych štúdií prípadov a kontrol*** sa vyberajú pacienti (chorí) a k nim zdraví jedinci. Pozerajú sa spätne do minulosti v určitom časovom intervale a porovnávajú sa napr. početnosti neexponovaných a exponovaných pacientov pomocou premenných charakterizujúcich napr. životný štýl, zvyky alebo takých premenných, ako biochemické alebo genetické markery. Cieľom je nájsť premenné (rizikové faktory), ktoré pomôžu vysvetliť, prečo prípady dostali chorobu a kontroly nie. Ďalej tiež zistiť, či expozícia súvisí s chorobou. Rozsahy skupín chorých a zdravých jedincov nemusia odrážať ich rozdelenie v populácii. Keďže tradičným cieľom týchto štúdií je zistiť prítomnosť alebo neprítomnosť choroby, prípady sú zvyčajne chorí. Avšak prípadmi môžu byť aj pacienti s nejakým druhom postihnutia a kontroly pacienti bez tohoto postihnutia.

Príklad 10 (retrospektívna štúdia prípadov a kontrol) *Vyberajú sa deti postihnuté a nepostihnuté kongenitálnymi malformáciami. Sledujeme, či matky mali alebo nemali rubeolu počas tehotenstva (Hackshaw, 2009).*

Príklad 11 (retrospektívna štúdia prípadov a kontrol) *Vzťah fajčenia a pľúcnych nádorov. Výskum sa realizoval v* 20 *nemocniciach v Londýne na* 709 *subjektoch rovnakého pohlavia s diagnostikovaným pľúcnym nádorom, ktoré boli zoskupené v 5-ročných vekových intervaloch, a im vekovo odpovedajúcich 709 subjektoch toho istého pohlavia bez nádorového ochorenia. Fajčiar sa definoval ako osoba, ktorá vyfajčila v ostatnom roku aspoň jednu cigaretu denne. Ide o retrospektívny dizajn, v ktorom nepoznáme prevalenciu pľúcnych nádorov v populácii.*

Zriedkavé sú ***prospektívne štúdie prípadov a kontrol***, kde sa prípady a kontroly sledujú v budúcnosti.

Príklad 12 (prospektívne štúdie prípadov a kontrol) *Štúdia chrípky ako dôvodu akútneho infarktu myokardu (AIM; MacIntyre a kol., 2013), kde prípadmi boli hospitalizovaní pacienti s AIM a kontrolami ambulatní pacienti bez AIM v Sydney (Austrália) v časovom období rokov* 2008 *až* 2010*.*

Okrem klasických štúdií prípadov a kontrol existujú aj tzv. ***vnorené štúdie prípadov a kontrol***, kde sa prípady a spárované kontroly vyberajú z kohortovej štúdie (pozri nižšie). Tieto štúdie sa špeciálne vyskytujú pri registroch chorôb, kedy sú prípady reprezentatívnou vzorkou populácie a reprezentatívne kontroly sa k nim získavajú z tej istej populácie.

Príklad 13 (vnorené štúdie prípadov a kontrol) *Štúdia úlohy infekcie dolných dýchacích ciest u pacientov liečených antibiotikami, ktorí znovu navštívili lekára v priebehu štyroch týždňov po poslednej návšteve, na ktorej im boli predpísané antibiotiká (Macfarlane a kol., 1997). Prípadmi boli pacienti, ktorí znovu navštívili lekára s rovnakými symptómami. Kontrolami boli pacienti, ktorí lekára znovu v priebehu štyroch týždňov už nenavštívili.*

Kohortové štúdie sledujú subjekty nejakej kohorty[4], retrospektívne alebo prospektívne.

[4]Slovo *kohorta* pochádza z latinského *cohort* (útvar bojovníkov s počtom 1/10 rímskej légie, ktorí boli súčasne nasadení do boja), ale dnes sa používa na opísanie určitej skupiny ľudí, ktorú sledujeme v danom časovom intervale.

Do ***retrospektívnej kohortovej štúdie*** sa vyberajú subjekty, ktorých základné merania (ako aj sledovanie) prebehli v minulosti, teda už existujú, a my ich vyhľadávame v lekárskych alebo archívnych záznamoch, regionálnych alebo národných databázach.

Do ***prospektívnej kohortovej štúdie*** (nazývanej aj *longitudinálna* v antropológii alebo *incidenčná* v medicíne) sa vyberajú subjekty, ktorých základné merania prebehnú v budúcnosti. Subjekty sa potom sledujú istý čas *longitudinálne* (niekoľko rokov) a početnosti tých, u ktorých sa vyvinie skúmaná choroba sa porovnávajú medzi neexponovanými a exponovanými nejakému faktoru, pričom možno zisťovať *incidenciu* choroby, príp. *remisiu* (dočasné vymiznutie prejavov choroby) alebo *relaps* (opakovaný výskyt prejavov choroby). V rastovej antropológii (auxológii človeka) sa longitudinálne štúdie využívajú predovšetkým na zisťovanie vývinových a rastových zmien u detí a dospievajúcich (Bogin, 1999).

Príklad 14 (longitudinálne štúdie v ČR) *V Brne sa realizovali dve rozsiahle longitudinálne štúdie. Prvou z nich bola **Brněnská růstová studie**, ktorú viedla docentka Marie Bouchalová (Bouchalová, 1987). Druhú reprezentuje zapojenie do **Evropské dlouhodobé studie těhotenství a dětství (ELSPAC)**, ktorej českú časť vedie docent Lubomír Kukla (Kukla, 2008).*

V longitudinálnych štúdiách sa sledujú zmeny v priebehu času na tej istej vzorke. Každá skúmaná osoba sa sleduje minimálne dvakrát za sebou s určitým časovým odstupom (hovoríme o *párovom dizajne*). Longitudinálna štúdia umožňuje sledovať skutočné zmeny a vytvárať pravé rastové krivky, vrátane *individuálnych rastových kriviek*. Umožňuje tiež sledovať vnútornú štruktúru časových (tu vekových) zmien. Jej nevýhodou je časová náročnosť, zmenšovanie vzorky (po čase skúmané osoby strácajú záujem pokračovať v štúdii) a *riziko metodickej nekonzistencie* (zmena metód, prístrojov, výskumníkov, spoločenských cieľov alebo dokonca etického rozmeru študovaného problému). Ohrozenie vnútornej validity štúdie spočíva vo vplyve histórie a celospoločenských zmien, ktoré ovplyvňujú celú populáciu. V longitudinálnej štúdii však túto zmenu zaznamenáme ako zmenu s vekom osôb sledovanej vekovej kohorty.

Príklad 15 (vplyv histórie a celospoločenských zmien v longitudinálnej štúdii) *V longitudinálnej štúdii sa v nedávnej histórii západnej civilizácie zaznamenalo zníženie počtu fajčiarov cigariet medzi 30. až 50. rokom života. Toto zníženie však nemusí byť len dôsledkom zákonitých systematických zmien v prístupe ľudí ku svojmu zdraviu v priebehu dospelosti a starnutia (stále zodpovednejšie, resp. opatrnejšie chovanie), ale môže byť aj dôsledkom celospolečenských zmien prebiehajúcich synchrónne so starnutím ľudí sledovanej kohorty (stále väčší dôraz spoločnosti na zdravý životný štýl, resp. politický tlak na obmedzovanie fajčenia, ekonomické zmeny relatívne zvyšujúce cenu a znižujúce dostupnosť tabakových výrobkov a pod.). Bez prierezovej kontroly týchto faktorov nie je možné rozhodnúť, ktorý z nich má na sledovaný jav podstatnejší vplyv. Vzhľadom na to, že z prierezových demografických štúdiách vyplýva (Sovinová a kol., 2012), že v mladších vekových kategóriách bolo percentuálne zastúpenie fajčiarov v roku 2011 takmer rovnaké ako v roku 1999 a zníženie bolo tak v roku 1999 ako aj v roku 2011 jasne zreteľné až vo vyššej vekovej kategórii (55–64 rokov), pravdepodobnejšie vysvetlenie pre uvedené obdobie sú skutočné individuálne zmeny súvisiace s vekom (zdravotné komplikácie alebo finančné problémy vo vyššom veku). Vo väčšom časovom rozsahu (napr. medzi 50. rokmi 20. storočia a začiatkom 21. storočia) však väčší vplyv môžu mať spoločenské zmeny postoja k fajčeniu.*

Porovnanie štúdií prípadov a kontrol a kohortových štúdií. Vo všeobecnosti sa považujú kohortové štúdie za reliabilnejšie než štúdie prípadov a kontrol, pretože *„spomienková“* (napr. ľudia so sledovanou chorobou si na zvyky v minulosti spomínajú lepšie než ľudia bez tejto choroby) a *selekčná výchylka* ich ovplyvňujú v menšej miere. Kohortové štúdie však zvyčajne trvajú niekoľko rokov, čo spôsobuje, že sú finančne často náročné. Štúdie prípadov a kontrol sú vhodnejšie, ak je študovaná choroba zriedkavá alebo trvanie výskumu z nejakého dôvodu časovo obmedzené, pretože v zmysluplnom časovom horizonte umožňujú získať väčší počet subjektov. Kohortové štúdie sa využívajú v prípadoch častejšie sa vyskytujúcich chorôb (s výskytom nad 10 % prípadov), kým štúdie prípadov a kontrol sa využívajú pri štúdiu príčin vzácnejšie sa vyskytujúcich chorôb (s výskytom pod 10 %). Na minimalizáciu rôznych druhov výchyliek (napr. „spomienkovej“, selekčnej, informačnej výchylky merania alebo výchylky spôsobenej chýbajúcimi pozorovaniami) sa hodia rôzne typy štúdií, takže je potrebné zahrnúť otázku výchylky do plánu štúdie.

Výber medzi kohortovou štúdiou a štúdiou prípadov a kontrol závisí napr. na zložitosti získania subjektov (napr. zriedkavo sa vyskytujúce choroby a pod.), ovplyvniteľnosti výsledkov výchylkou a mätúcimi premennými, finančnom zabezpečení a trvaní štúdie.

Kohortová štúdia a štúdia prípadov a kontrol sa líšia aj použitým **efektom sledovaného rizika**. Pri kohortovej štúdii je to zvyčajne relatívne riziko a pri štúdii prípadov a kontrol pomer šancí.

Definícia 8 (mätúce premenné) *Mätúce premenné definujeme ako cudzie premenné, ktorých prítomnosť ovplyvňuje študované premené, takže výsledky nereflektujú aktuálny vzťah medzi závislou a nezávislou premennou.*

Mätúce premenné možno povoliť pri dizajne štúdií prípadov a kontrol, kde sú prípady párované s kontrolami, ale nie pri kohortových štúdiách, lebo by nežiaduco ovplyvnili štatistické modely počas štatistických analýz. Plán štúdie musí obsahovať zdôvodnenie voľby prípadných mätúcich premenných a možnosť kvantifikácie ich vplyvu nejakou štandardnou, presnou a komplexnou cestou (napr. uvedenie etnicity, presné zaznamenanie množstva skonzumovaného alkoholu).

Napokon je potrebné zdôrazniť, že incidenciu alebo riziko výskytu ochorenia možno zistiť z kohortových štúdií, ale nie zo štúdií prípadov a kontrol.

Definícia 9 (štatistiky frekvencie) ***Frekvencia*** *(počet, početnosť) v užšom slova zmysle vyjadruje počet nejakých udalostí (prípadov), kým v širšom slova zmysle vyjadruje početnosť výskytu nejakého znaku (pozri kapitolu 1.3.7 Triedenie dát).* ***Incidencia*** *vyjadruje počet nových prípadov s daným znakom v populácii za jednotku času.* ***Prevalencia*** *označuje celkový výskyt (počet) žijúcich jedincov s daným znakom v populácii teraz alebo kedykoľvek v minulosti.* ***Abundancia*** *vyjadruje početnosť prípadov (napr. v ekológii počet jedincov daného živočíšneho druhu na danej lokalite).*

Tabuľka 1.1: Štatistiky frekvencie vs. typ štúdie (# znamená počet)

typ štúdie	štatistika	definícia
kohortová	**incidencia**	$\frac{\#\text{subjektov, ktorí } \textbf{dostali} \text{ chorobu}}{\#\text{subjektov v riziku} \times \text{čas v riziku}}$
	kumulatívna incidencia	$\frac{\#\text{subjektov, ktorí } \textbf{dostali} \text{ chorobu}}{\#\text{subjektov v riziku}}$
prierezová	**prevalencia**	$\frac{\#\text{subjektov, ktorí } \textbf{majú} \text{ chorobu}}{\#\text{subjektov v riziku}}$
	kumulatívna prevalencia	$\frac{\#\text{subjektov, ktorí } \textbf{kedykoľvek nadobudli} \text{ chorobu}}{\#\text{subjektov v riziku}}$

Všetky štyri štatistiky frekvencie sa používajú v absolútnej škále (počet, početnosť) alebo v relatívnej škále (škálované počtom subjektov v súbore). Je potrebné si uvedomiť, že hodnoty incidencie sú ovplyvnené časovými jednotkami. Definícia štatistík frekvencie závisí aj na type štúdie (pozri tabuľku 1.1). **Sekvenčné alebo semilongitudinálne štúdie** predstavujú kombináciu prierezovej a longitudinálnej štúdie, ktorá kombinuje výhody a súčasne obmedzuje ohrozenie vnútornej validity výskumu oboch prístupov. Prebieha kratší čas než štúdia longitudinálna a zatiaľ čo prierezová zložka umožňuje kontrolu vplyvu histórie, longitudinálna zložka umožňuje kontrolu efektu kohorty.

Plán observačnej štúdie. Najdôležitejšími prvkami observačnej štúdie, ktoré musia byť zahrnuté v jej pláne, sú:

- v prípade študovaných subjektov (skúmaných osôb, účastníkov výskumu)
  - **výberová schéma**, t.j. špecifikácia miesta, odkiaľ budú subjekty pochádzať, v štúdii prípadov a kontrol špeciálne kontroly (z populácie, pacientov nemocnice alebo z príbuzných prípadov);
  - **selekčné kritériá**, **kritériá vhodnosti** alebo **zoznam prijímacích a vylučovacích kritérií** – prijímacie kritériá musia byť presné a podrobné (presná špecifikácia populácie relevantnej na riešenie vedeckej otázky a presná špecifikácia vhodnosti štúdie, napr. *demografické charakteristiky* ako pohlavie a vek, *klinické charakteristiky* ako dobré všeobecné zdravie a stály sexuálny partner, *časové charakteristiky* ako časový interval zberu dát a pod.), vylučovacie kritériá musia byť úsporné (špecifikácia subjektov, ktoré nebudú študované pre vysoké riziko straty zo sledovania, pre nemožnosť zabezpečiť kvalitné dáta, napr. v dôsledku jazykovej bariéry, pre vysoké riziko škodlivých/vedľajších/nepriaznivých účinkov, napr. v dôsledku prekonaného infarktu myokardu alebo mozgovej príhody);

  - **typ náhodného výberu**, t.j. selekcia z výberovej schémy;
  - jasná **definícia prípadov a kontrol** (v štúdii prípadov a kontrol);
  - **párovacie kritéria a faktory**, t.j. ako budú kontroly párované s prípadmi (v štúdii prípadov a kontrol);

- v prípade náboru subjektov a ich sledovania
  - **časový interval** použitý na identifikáciu subjektov;
  - **dĺžka náboru a sledovania v budúcnosti** (v kohortových štúdiách);
  - **dĺžka sledovania do minulosti** t.j. ako hlboko do minulosti budú pacienti sledovaní v súvislosti s detailami ich životného štýlu alebo zvykov (v štúdii prípadov a kontrol);

- v prípade zberu dát a vyšetrení
  - **detaily zberu dát**, t.j. interview, typ dotazníkov (samovyplňovacie alebo vyplňované niekým iným), dáta z pacientskych záznamov, dáta z regionálnych alebo národných databáz a pod.;
  - **detaily vyšetrení**, t.j. typ vyšetrenia, napr. klinické vyšetrenie, snímkovanie (RTG, CT sken, MRI sken, 2D fotografia tváre, 3D laserová snímka, 3D stereofotogrametrická snímka, snímka z elektrónového mikroskopu a pod.), krvný test, odber moču, slín a pod.; kto vykoná vyšetrenia, frekvencia vyšetrení;
  - **spôsob skladovania biologického materiálu** (krv, moč, sliny) v rámci danej štúdie alebo z hľadiska jeho ďalšieho využitia a analýz v budúcich štúdiách;
  - **spôsob merania mätúcich premenných/faktorov**;
  - **spôsob minimalizácie výchyliek** rôzneho typu, napr. „spomienkovej", selekčnej, informačnej výchylky, výchylky merania alebo výchylky spôsobenej chýbajúcimi pozorovaniami;
  - **spôsob určenia diagnózy**.

Párovacie kritériá v štúdii prípadov a kontrol. Častým nedostatkom protokolu je nedostatočný opis párovania prípadov a kontrol. Pojem párovanie sa navyše často používa veľmi nepresne. Mal by indikovať pojem **spárovaná dvojica**, čo však nemusí byť zámerom výskumníka alebo plánovaného štatistického prístupu. Návrh, v ktorom sú subjekty párované na základe pohlavia, etnicity, veku alebo BMI, je nepresný (v spojitosti s realitou) a v praxi nerealizovateľný. Výskumník by mal zabezpečiť, aby sledované skupiny boli **vyvážené** v závislosti od sledovaných charakteristík, t.j. ide o **frekvenčné párovanie**. Pre spojité premenné (ako napr. vek) je potrebné špecifikovať rozsah, ktorý charakterizuje toto párovanie spolu s priemerom a smerodajnou odchýlkou. Vyhnúť sa treba párovaniu pomocou mätúcich premenných, ktorých vzťah ku závislej premennej nie je dostatočne preskúmaný.

Definícia 10 (párovanie do dvojíc) *Párovanie do dvojíc je výber kontroly pre každý prípad tak, aby obe mali rovnakú hodnotu mätúcej premennej. Ak je medzi skupinami štatisticky významný rozdiel v premenných, ktoré charakterizujú párovanie, je potrebné ich vplyv odpočítať pomocou vhodného štatistického modelu. Toto je jednoznačne úlohou bioštatistika.*

Príklad 16 (štatistické hľadisko párovania do dvojíc) *Majme párovanie do dvojíc pomocou premenných pohlavie a vek. Keďže párovanie prípadov bolo vykonané v rámci rodiny, kde kontroly boli rodinní príslušníci rôzneho veku a pohlavia (rodičia, súrodenci a pod.), je potrebné použiť štatistický model odstraňujúci negatívny efekt párovania, ktorý vplyv veku a pohlavia eliminuje.*

Definícia 11 (frekvenčné párovanie) *Frekvenčné párovanie je taký výber kontrol ku prípadom, v ktorom sú mätúce premenné v oboch skupinách zo štatistického hľadiska identické.*

Definícia 12 (štatistické hľadisko frekvenčného párovania) *Zo štatistického hľadiska nesmie byť štatisticky významný rozdiel medzi tými premennými, ktoré charakterizujú frekvenčné párovanie.*

Výberová schéma musí byť jednoznačne definovaná a musí zahŕňať tzv. *kritériá vhodnosti*, hlavne v podobe vekových ohraničení (napr. zahrnutie starších ľudí do štúdie a zdôvodnenie rozhodnutia).

Príklad 17 (kritériá vhodnosti) *Ak je cieľom štúdie hľadanie vzťahu (asociácie) medzi fajčením a nádorom pľúc, potom prijímacími kritériami sú napr. vek* 18–80 *rokov, nefajčiari a fajčiari schopní doručiť informovaný súhlas a vylučovacími kritériami sú napr. súdobo liečení alebo jedinci s anamnézou nádoru, vystavení nejakým látkam alebo situáciám, ktoré sú známe svojou asociáciou s vysokým rizikom vzniku nádoru.*

Príklad 18 (výberová schéma – prierezová štúdia) *Štúdia a ciele: prierezová štúdia odhadu prevalencie konzumácie alkoholu a fajčenia medzi študentmi antropológie. Výberová schéma: register všetkých študentov antropológie univerzity. Typ výberu: všetci študenti v registri.*

Príklad 19 (výberová schéma – štúdia prípadov a kontrol) *Štúdia a ciele: štúdia prípadov a kontrol na hodnotenie asociácie medzi vrodenými poruchami a fajčením počas tehotenstva. Výberová schéma: prípady – deti s vrodenou poruchou vybrané z národného registra vrodených porúch, kontroly – deti bez vrodenej poruchy vybrané z registra pôrodov vo vybranom geografickom regióne. Typ výberu: prípady – všetky deti s vrodenou poruchou z národného registra narodené medzi rokmi 2000 a 2010, kontroly – náhodný výber detí bez vrodenej poruchy narodené medzi rokmi 2000 a 2010 so spárovaným vekom matky a rokom narodenia dieťaťa.*

Príklad 20 (výberová schéma – kohortová štúdia) *Štúdia a ciele: kohortová štúdia hodnotenia asociácie medzi cvičením a rizikom vzniku srdcovocievnej choroby. Výberová schéma: všetci pacienti starší ako* 40 *rokov registrovaní na klinike v príslušnom geografickom regióne. Typ výberu: náhodný výber* 5000 *pacientov.*

1.2.2 Systematický prehľad

Úvod každej štúdie v podstate predstavuje krátky (cielený) systematický prehľad témy, ktorá súvisí so študovanou problematikou. Tento prehľad akumuluje vedomosti v sledovanej oblasti na základe už publikovaných a nepublikovaných výsledkov, ktorých závery sa syntetizujú (kombinujú) do všeobecnejšej podoby. Efekty každej štúdie sa kvantitatívne kombinujú pomocou štatistických metód, ktoré súborne nazývame **meta-analýza**. Základom takejto štúdie je identifikácia publikovanej a nepublikovanej literatúry v sledovanej tematike. Kumulácia vedomostí môže mať tieto dôsledky:

- **Potvrdenie doterajšej praxe** získaním presnejšieho odhadu efektu alebo rizikového faktora a následné častejšie používanie v praxi a pod. Keďže kombinovaná analýza má vyšší rozsah výberu, sila použitého testu v rámci nejakého štatistického modelu bude vyššia. Pomôže to aj vyhnúť sa neistote prípadného falošného výsledku v individuálnej analýze.
- **Zmena doterajších praktík**, čo v praxi vedie napr. k zmene typu liečby, k zmene názoru na nejakú vedeckú problematiku, k identifikácii nového rizikového faktora vedúceho k vzniku ochorenia alebo k predčasnej smrti (na základe observačných štúdií), ďalej k zmene doterajšej štandardnej klinickej praxe, alebo môže viesť k vývinu novej zdravotnej politiky a pod.

Problémom systematického prehľadu pri observačných štúdiách môže byť kombinácia individuálnych výchyliek a mätúcich premenných štúdií zahrnutých do meta-analýzy. To môže viesť ku falošným záverom alebo asociáciám, ktoré sa zdajú byť presnejšie (napr. získaním užšieho intervalu spoľahlivosti). V klinických pokusoch možno tieto výchylky účinne kontrolovať.

Existujú dva typy meta-analýzy

1. **meta-analýza na základe sumárnych výsledkov** získaných z výsledkov publikovaných a nepublikovaných individuálnych štúdií;
2. **meta-analýza na základe originálnych (pôvodných) dát** získaných od autorov individuálnych štúdií; v medicínskom výskume sa nazýva aj *meta-analýza dát od individuálnych pacientov.*

Meta-analýza sa v súčasnosti uplatňuje predovšetkým v psychológii, sociológii a v biomedicínskom výskume. Vzhľadom na komparatívnu a syntetickú povahu antropológie je meta-analýza žiaducou metódou aj v tomto odbore. Napriek tomu sa však meta-analýza v antropológii doteraz príliš neuplatnila (príklady meta-analytických štúdií v antropológii možno nájsť v práci Russella Bernarda 2006, s. 107–108) a na svoj väčší rozvoj ešte len čaká.

Príklad 21 (dôležitosť meta-analýzy) *Dôležitosť meta-analýzy pri formulovaní hypotéz v antropológii demonštruje výsledok štúdie Whitea (White, 1980, cit. Russell Bernard, 2006, s. 108), v ktorej na základe stovky publikovaných štúdií autor zistil, že sociálny status zodpovedá iba za 6.3 % celkového rozptylu v školskom prospechu študentov. Autor tiež zistil, že pôvodné korelácie naprieč štúdiami kolísali od hodnoty* -0.14 *po hodnotu* $+0.97$. *Tento rozsah kolísania demonštruje nutnosť kontroly ďalších faktorov, ktoré môžu prospech ovplyvňovať.*

1.2.3 Štúdie kostrových pozostatkov

Zvláštnym typom súborov/vzoriek v antropológii sú súbory kostrových pozostatkov, tzv. kostrové série. Podľa charakteru výberu a kontextových informácií ich možno rozdeliť na niekoľko skupín. Ich spoločnou vlastnosťou je fakt, že obsahujú kosti zomretých ľudí[5]. Vzorky pozostávajúce z nájdených kostrových pozostatkov sa značne odlišujú od vzoriek vytvorených na základe definovaného výberu ľudí zo živej populácie. Tieto súbory totiž nevznikajú štandardnými postupmi výberu vzorky z populácie, ale pod vplyvom rôznych, väčšinou nekontrolovaných (často nekontrolovateľných) faktorov.

Prvým typom sú **dokumentované antropologické zbierky**[6], ktoré obsahujú relatívne kompletné skelety a údaje o hlavných biologických znakoch jedincov (dožitý vek, pohlavie, a i.). Tieto zbierky majú veľký význam pri tvorbe antropologických metód odhadujúcich telesné znaky zo skeletov z archeologických nálezov. Jednotlivé dokumentované zbierky sa však líšia spôsobom výberu vzorky a rozsahom kontextových dát. Jednu podskupinu predstavujú **anatomické zbierky** vytvorené na základe kostí jedincov z anatomických pitiev, druhou sú **forenzné zbierky** (väčšinou menšieho rozsahu), treťou sú **zbierky z exhumácií pochovaných tiel** a štvrtou sú zväčša nedostupné **zbierky kostrových pozostatkov dokumentovaných obetí vojnových konfliktov**. Novo vznikajú aj **zámerne vytvárané „pohrebiská“**[7], zložené z hrobov dobrovoľníkov, ktorí ešte za života odkázali svoje telo po smrti vede. Rozdiely v selekcii vzoriek sa premietajú do parametrov súborov. Kým anatomické zbierky bývajú zo sociálneho či zdravotného hľadiska *selektívne*, *neselektívne* exhumácie dokumentovaných hrobov z rozsiahlejších bežných civilných cintorínov predstavujú pravdepodobne najprirodzenejšie kostrové súbory (hoci aj tieto môžu byť selektívne zo sociálneho hľadiska).

Druhým typom sú **kostrové súbory z archeologických výskumov**, na základe ktorých kostrová antropológia (humánna osteológia, bioarcheológia, historická antropológia) študuje život minulých populácií. Získanie týchto súborov je dnes väčšinou viazané na záchranné archeologické výskumy obmedzeného rozsahu. Preto sa v nich výrazne prejavuje *neúplnosť dát* v miere danej stavom zachovalosti a kompletnosti kostí/skeletov (závisiacej na geologických a pôdnych podmienkach danej lokality, na rozdieloch v odolnosti skeletu rôznych jedincov – deti vs. dospelí, ženy vs. muži – voči tafonomickým vplyvom, na pohrebnom ríte, na exkavačnej a rekonštrukčnej metodike a i.), rozsahom pohrebísk, pomerom rozsahu výskumu a rozsahu pohrebiska (možnosť lokálnych nehomogénností v zložení hrobov na pohrebiskách) a dostupných kontextových údajov a ich (ne)úplnosti (napr. rozdiely v pohrebnom ríte a hrobovej výbave, znalosť väzby na sídlisko atď.). Extrémnym prípadom sú **súbory kremácií**, v ktorých sú kosti poškodené žiarom (príp. aj rozbité zámerne v súvislosti s pohrebným rítom) a extrémne fragmentárne. V prípade takýchto sérií nemožno určiť ani základné biologické kategórie.

Okrem obmedzenej reprezentatívnosti a celistvosti dát získaných z kostrových súborov sa s výskumom minulých populácií spájajú ďalšie okolnosti, ktoré sa označujú ako **osteologický paradox** (Wood a kol., 1992). Tento jav možno sledovať v troch problémových oblastiach, ktoré sprevádzajú archeologické kostrové súbory:

1. **demografická nestabilita** – výskyt a proporcie výskytu skeletov na pohrebiskách závisia od demografických zmien v danej populácii, napr. vyšší výskyt detských skeletov môže znamenať väčšiu

[5] Výnimkou sú série založené na záznamoch z neinvazívnych zobrazovacích metód. Tam však ide o nenáhodnosť a selektívnosť iného druhu – pacientov.

[6] Pozri napr. http://skeletal.highfantastical.com.

[7] Jantzova zbierka v Tennesee (WM Bass Donated Skeletal Collection): http://fac.utk.edu/collection.html.

detskú mortalitu v demograficky stabilnej populácii alebo napr. nižšiu detskú mortalitu a vyššiu natalitu v demograficky rastúcej populácii;

2. **selektívna mortalita** – k dispozícii sú iba skelety ľudí, ktorí zomreli, nikdy ale nemáme prístup k ľuďom, ktorí boli v danej vekovej kategórii rizikom ohrození, ale prežili; rôzne vekové kategórie v archeologických kostrových súboroch sú potom zložené z pozostatkov ľudí odlišných biologických vlastností;

3. **stratená heterogenita rizík** – v nájdených kostrových súboroch sa vyskytuje zmes ľudí líšiacich sa nielen príčinou smrti, ale aj citlivosťou na rizikové faktory (napr. choroby, úrazy a pod.); bez poznania príčin smrti je ťažké vytvoriť relevantné biologické kategórie vzájomne si podobných jedincov a následne študovať napríklad rozdiely v pôsobení genetických a sociálnych faktorov atď.

Celkovo sú archeologické súbory v rôznej miere **nereprezentatívne**, zaťažené mnohými neznámymi tendenciami, takže nemusia spĺňať nároky štatistických metód na charakter vstupných dát (napr. náhodné vzorkovanie, normalita rozdelenia). Uvedené okolnosti výrazne limitujú možnosti štatistických porovnávaní a testov a predovšetkým interpretácií, ktoré z nich vyplývajú. Pri hodnotení archeologických kostrových súborov preto treba – podľa možnosti – do použitého modelu zakomponovať predstavu o demografických procesoch v danej populácii, ako aj o environmentálnych, socio-ekonomických, socio-kultúrnych a zdravotných podmienkach. To kladie ešte väčšie nároky na spoluprácu antropológa a štatistika od samého počiatku príprav výskumu.

1.3 Štatistické hľadiská plánovania štúdií

Zo štatistického hľadiska sa treba pri plánovaní pokusov zamerať na typ výberu, rozsah náhodného výberu, na metódy zberu dát a spôsoby merania, na hlavné atribúty dát a ich triedenie. Každému z týchto okruhov bude venovaná pozornosť v nasledovných kapitolách.

1.3.1 Typy výberov

Štatistický (základný) súbor je konečná množina prvkov (napr. osôb, jedincov, subjektov), na ktorej sledujeme určité znaky, veličiny, vlastnosti a pod. Ak uskutočníme nejaký výber zo štatistického súboru, hovoríme o **výberovom súbore**, ktorý reprezentuje štatistický súbor výberom určitého počtu štatistických jednotiek. **Štatistické jednotky** tvoria napr. subjekty, respondenti dotazníka, pričom po uskutočnení výberu hovoríme o **výberových jednotkách**. Na základe výberových jednotiek vo výberovom súbore si vytvárame predstavu o štatistických jednotkách v základnom súbore. Kvalita tejto predstavy v podobe nejakého štatistického modelu alebo modelu rozdelenia pravdepodobnosti striktne závisí od správnosti zvoleného typu výberu.

Príklad 22 (náhodný výber vs. zovšeobecnenie na populáciu) *Často nemôžeme z rôznych príčin sledovať celú populáciu. Nemôžeme napr. do výskumu zaradiť všetky tehotné ženy alebo všetky osoby žijúce v skúmanej geografickej oblasti. Ak je našim cieľom napr. zisťovať vzťah prírastku hmotnosti ženy počas tehotenstva a hmotnosti novorodenca, môžeme údaje zistiť len od určitej vzorky (výberu) tehotných žien. Získané výsledky sa následne zovšeobecňujú na všetky tehotné ženy. Kvalita zovšeobecnenia pritom priamo závisí od kvality výberu.*

Definícia 13 (zovšeobecnenia záverov štúdie) *Na základe* ***aktuálnych/použitých subjektov*** *v* ***aktuálnom/realizovanom výbere*** *formulujeme závery štúdie pre použité subjekty [zistenia v štúdii], potom zovšeobecníme závery na* ***zamýšlanú populáciu*** *napr. na populáciu dospelých vo veľkých slovenských mestách (****inferencia na základe internej validity****) [pravdivosť štúdie], potom na* ***dosiahnuteľnú populáciu*** *napr. na populáciu dospelých vo všetkých slovenských mestách (****inferencia na základe externej validity****), a nakoniec na* ***cieľovú populáciu*** *– napr. dospelých vo všetkých slovenských mestách a dedinách [všeobecná pravdivosť]. Menej spoľahlivým záverom by bolo napr. zovšeobecnenie na dospelých a deti, na ľudí v iných krajinách a pod. (Hulley a kol., 2007).*

Interná validita je vlastnosť vedeckej štúdie, ktorá reflektuje mieru/rozsah zaručenosti kauzálnych záverov vedeckej štúdie. Externá validita je validita zovšeobecnenej (kauzálnej) štatistickej inferencie (vedeckej štúdie) alebo miera zovšeobecnenia na iné (všeobecnejšie, v istom zmysle hierarchicky nadradené) situácie.

Medzi základné typy výberov patria (1) náhodný výber, (2) selektívny výber, (3) zámerný výber, (4) výber typu snehovej gule a (5) výber typu vhodnosti (komfortu) a príležitosti. Prvému typu výberu hovoríme aj **pravdepodobnostný výber**, ostatné patria do skupiny **nepravdepodobnostných výberov**.

Náhodný výber je výber, v ktorom vyberanie šatistických jednotiek z populácie prebieha celkom náhodne a nezávisle na našom úsudku (Bland, 2009). Rozlišujeme nasledovné typy náhodných výberov:

1. **Jednoduchý náhodný výber** je priamy výber štatistických jednotiek, pričom každá má rovnakú pravdepodobnosť, že bude vybraná (napr. *generovanie pseudo-náhodných čísel* alebo *žrebovanie*, pričom je výhodné, keď sú štatistické jednotky očíslované a je možné použiť matematickú teóriu náhodných čísel).

2. **Mechanický (systematický) výber** je založený na určitom (dopredu stanovenom) usporiadaní prvkov populácie a spočíva vo výbere tých prvkov (subjektov), ktoré sú vzájomne vzdialené o zvolený výberový krok. Prvý prvok (subjekt) vyberieme *jednoduchým náhodným výberom*. Pri tomto výbere musíme dať pozor, aby usporiadanie prvkov (subjektov) nesúviselo so sledovaným znakom.

3. **Oblastný (stratifikovaný) výber**. Základný súbor je rozdelený na oblasti podľa určitého hľadiska. Oblasti sú vytvorené tak, aby boli *vnútorne homogénne* (v sledovaných znakoch sa vnútri oblasti príliš neodlišujú) a *medzi sebou heterogénne* (v sledovaných znakoch sa môžu, ale nemusia odlišovať). V jednotlivých oblastiach sa uskutoční *jednoduchý náhodný výber* alebo *mechanický výber*. Percento vybraných jednotiek môže byť buď vo všetkých oblastiach rovnaké, alebo medzi oblasťami odlišné (napr. z ekonomických dôvodov musíme vybrať menší počet jednotiek, výber jednotiek je náročný; navyše percento vybraných jednotiek môže reprezentovať prirodzené rozdelenie sledovanej premennej v populácii). Konečný výberový súbor vytvoríme spojením všetkých oblastí.

4. **Skupinový výber**. Pokiaľ je štatistický (základný) súbor pomerne rozsiahly (stotisíce alebo milióny osôb), dá sa jednoduchý náhodný výber uskutočniť veľmi ťažko; najprv sa náhodne vyberú *skupiny (agregáty) jednotiek* (nie jednotlivé jednotky), ktoré tvoria buď prirodzené alebo umelé agregáty. Je žiaduce, aby boli jednotlivé agregáty, pokiaľ je to možné, rovnako veľké a vnútri každej skupiny *heterogénne*. Variabilita medzi skupinami musí byť čo najmenšia, t.j. blízka *homogénnej* (opačne než pri oblastnom výbere). Pri tomto výbere rozlišujeme dve možnosti.
   - *Výber všetkých jednotiek v agregáte.*
   - *Dvojstupňový a viacstupňový výber*. Je založený na hierarchickom opise prvkov základného súboru, ku ktorým sa dostávame postupne cez vyššie výberové jednotky. Každá výberová jednotka je skupinou výberových jednotiek nižšieho rádu, kde rozlišujeme *jednotky prvého stupňa (primárne jednotky)*, potom *jednotky druhého stupňa (sekundárne jednotky)* atď., až nakoniec máme *základné jednotky* štatistického súboru. Postupné výbery prebiehajú často *jednoduchým náhodným výberom*, ale použiť sa môže aj *mechanický* alebo *oblastný výber*.

Príklad 23 (kartotéka pacientov) *Z abecedne usporiadanej kartotéky (databázy) pacientov u praktického lekára vyberáme s krokom* 10 *a prvú kartu vyžrebujeme (napr. deviatu kartu), takže výber bude tvorený pacientmi, ktorých karty boli uložené v poradí* 9, 19, 29, 39, 49 *atď.*

Príklad 24 (zamestnanie pacientov) *Zisťujeme zamestnanie pacientov vedených v stomatologickej ambulancii. Pacienti prichádzajú do ambulancie v určitom časovom slede, ktorému zodpovedá usporiadanie ich zdravotných záznamov. Výber s krokom rovným dennému počtu pacientov (za predpokladu, že do ambulancie príde každý deň rovnaký počet pacientov) vedie k značne selektívnemu výberu tých pacientov, ktorí prichádzajú v určitú časť dňa, čo môže súvisieť s typom ich zamestnania.*

Príklad 25 (oblasti) *Ak robíme výber obyvateľov Slovenskej alebo Českej republiky, oblasťami sú územné celky, vekové skupiny alebo socioekonomický status.*

Príklad 26 (hierarchia) *Vyššie výberové jednotky → prvky základného súboru = mestá – bloky – domy – domácnosti, okresy – podniky – dielne – zamestnanci.*

Príklad 27 (agregát) *Agregát môže byť malý – rodina, škola, podnik, zdravotný obvod, ale i väčší – obec, okres.*

Selektívny výber je výber, v ktorom vybraná vzorka nepokrýva požadovanú cieľovú vzorku, čo dáva skreslený obraz o študovanej populácii.

Príklad 28 *Vzorka* 15 *až* 16*-ročných chlapcov, prvoligových basketbalistov, z ktorej by sme chceli robiť inferenciu o výške chlapcov tohto veku v celej slovenskej populácii.*

Zámerný výber je výber, v ktorom o výbere jednotky rozhodujú okrem náhody častokrát nekontrolovateľné činitele (subjektívny názor vyberajúceho, ochota, resp. neochota odpovedať na kladené otázky a pod.). Všeobecne sa tvrdí, že tento typ výberu sa opiera o „expertné“ stanovisko a rôzne „odhady“ ako získať reprezentatívny výber. Takto získané výberové súbory sú často ovplyvnené subjektívnym pohľadom „experta“, ale aj ďalšími faktormi ovplyvňujúcimi tvorbu výberu a presnosť zovšeobecňujúcich záverov sa skôr opiera o „expertný“ pohľad než o metodológiu.

Výber typu snehovej gule predstavuje typ nenáhodného výberu subjektov, v ktorom práve hodnotené subjekty odporučia alebo privedú ďalšie subjekty z radov svojich známych (Russell Bernard, 2006, s. 192–194). Tento (rovnako ako aj iné typy nenáhodného výberu subjektov) je predmetom kritiky, jeho výsledkom je však vzorka zložená z ľudí tvoriacich prirodzenú sieť sociálne prepojených osôb. To má význam pri vytváraní vzorky u skrytých populácií. Ide o osoby riedko rozmiestené na veľkom území, osoby sociálne stigmatizované (napr. osoby HIV pozitívne) alebo osoby zámerne sa skrývajúce (napr. narkomani, zločinci a pod.). Interpretácia výsledkov musí vždy vychádzať z charakteru vzorky, ktorý je daný typom vzorkovania (výberu).

Výber typu vhodnosti (komfortu) a príležitosti predstavuje typ nenáhodného výberu subjektov, v ktorom vyberáme subjekty ochotné a ľahšie dostupné (Walker a Almond, 2010). Najčastejšie sa tento typ výberu používa v pilotných štúdiách a v prieskumoch verejnej mienky. Tento výber je selektívny, preto je veľmi dôležité zozbierať aj demografické dáta, ktoré môžu poukazovať na výberové výchylky.

Príklad 29 (dobrovoľníci vo výbere) *Dobrovoľníkov by sme nemali považovať za prvky náhodného výberu. Ak by meranie prebiehalo napr. kolektívne počas telocviku, ako dobrovoľníci by sa mohli prihlásiť iba ľudia, ktorí nemajú problém so svojou telesnou hmotnosťou. Vo vzorke by tak chýbali jedinci obézni alebo anorektickí, ktorí sa meraniu vyhli z obavy pred sociálnou kritikou vyplývajúcou zo zistených hodnôt. Pri zachovaní dobrovoľnosti i s meraním v intimite mimo zrak ostatných je potrebné s týmto počítať. Výskum preto musí obsahovať presnú definíciu prijímacích a vylučovacích kritérí).*

Pri realizácii uvedených štatistických výberov dochádza k dvom druhom chýb.

1. **Chyby reprezentatívnosti** sú rozdiely v hodnotách charakteristík základného a výberového súboru, ktoré sú zapríčinené tým, že sa skúma iba časť základného súboru. Ak bol výberový súbor vytvorený korektne, tak rozdiely v charakteristikách základného a výberového súboru predstavujú *náhodné chyby.*

2. **Chyby registrácie** sú rozdiely medzi presnými hodnotami znaku u vybraných jednotiek a zaregistrovanými hodnotami[8].

[8]Chyba merania nie je chybou registrácie (pozri kapitolu 1.3.3 Hlavné atribúty dát).

1.3.2 Náhodný výber v realite antropologického výskumu

Ideálny náhodný výber, zodpovedajúci vyššie uvedenej definícii, je v reálnych podmienkach antropologického výskumu (ľubovoľného zamerania či materiálu) skôr výnimkou než normou. Zriedka totiž máme k dispozícii všetkých jedincov danej populácie (tj. celú populáciu), z ktorej náhodne vyberáme študovanú vzorku (pozri kapitolu 1.3.1 Typy výberov). V štúdiách živého človeka väčšinou ide o **štúdie na dobrovoľníkoch**. Závisí od charakteru skúmaného javu a použitej metodiky, do akej miery bude dobrovolná účasť v šúdii meniť charakter výberu oproti náhodnému. Nenáhodnosť môžeme obmedziť z hľadiska geografického pôvodu – napr. v geograficky stratifikovanom výbere, kde ide o výber dobrovoľníkov proporcionálny ku množstvu jedincov potrebných charakteristík v danej geografickej oblasti. Dáta štatisticky analyzujeme tak, akoby o náhodný výber šlo. Zovšeobecnenia na celú populáciu sú však problematické.

Výnimkou môžu byť v antropológii iba **štúdie kompletných dát** (databáza tohto typu má charakter **registra**), t.j. štúdie celej populácie určitej geografickej oblasti, kedy nemáme k dispozícii vzorku, ale všetky prípady daného charakteru (napr. všetkých novorodencov v danej spádovej oblasti za dané obdobie).

V **kostrových súboroch** (pozri kapitolu 1.2.3 Štúdie kostrových pozostatkov, komentár k osteologickému paradoxu) nemožno väčšinou náhodný výber zaručiť, a to ani v prípade dokumentovaných kostrových súborov. Pri kostrových súboroch z pohrebiska dáta štatisticky analyzujeme tak, akoby o náhodný výber šlo. Pritom predpokladáme, že množstvo rôznych náhodných a vzájomne nesúvisiacich a často protichodných faktorov (aj keď samostatne nenáhodných) v rôznych obdobiach a na rôznych hierarchických úrovniach spôsobí, že výsledná vzorka sa blíži náhodnému výberu z populácie. Nemusí tomu tak však byť, lebo synergické pôsobenie niekoľkých významnejších faktorov môže spôsobiť systematický posun aj napriek ich náhodnosti. Napr. mladé dospelé ženy s maskulínnejšími kosťami panvy môžu hypoteticky byť v kostrovej sérii zomrelých dospelých žien nadhodnotené už v čase úmrtia (z dôvodu problémov pri pôrode), tafonomické procesy potom môžu túto nerovnováhu ešte zosilniť (skelety väčších, maskulínnejších žien se lepšie zachovajú). Výsledkom môže byť nižší sexuálny dimorfismus v danej sérii oproti reálnej situáci v danej populácii za života.

1.3.3 Hlavné atribúty dát

Štatistické spracovanie dát vyžaduje získanie kvalitných dát, a to už v terénnej fáze výskumu. Kvalitné dáta spájajú nasledovné atribúty:

1. **Objektivita**. Pod objektivitou informácií rozumieme vlastnosť pravdivo odrážať alebo zachytávať skúmané javy resp. ich znaky. Postup skúmania je objektívny vtedy, keď získané fakty podávajú verný obraz o skúmanom jave. Objektivitu zabezpečí taký postup, ktorý hodnoty znaku určí jednoznačne, takže výsledky nie sú závislé od osoby vyhodnocovateľa.

2. **Reliabilita (spoľahlivosť)**. Výskumná metóda je reliabilná vtedy, keď zisťuje (meria) tak, že pri opätovnom použití tejto metódy na rovnakých dátach (aj pri zisťovaní rôznymi výskumníkmi) za inak rovnakých podmienok dostaneme v podstate tie isté výsledky.

3. **Validita (platnosť)**. Je zabezpečená vtedy, keď otázky v dotazníku alebo miery na lebke a pod. zisťujú a merajú tie vlastnosti, ktoré skutočne chceme zisťovať.

4. **Reprezentatívnosť**. Pri niektorých typoch štúdií je potrebné zabezpečiť, aby dáta tvorili reprezentatívnu vzorku, teda vzorku, ktorá odráža skutočné rozdelenie znaku v populácii, napr. vekové, podľa pohlavia alebo sociálneho statusu a pod. Reprezentatívnosť môže znamenať aj rovnomerné zastúpenie alebo komplexnosť v zastúpení nejakých skupín z populácie. Interpretácia tohto pojmu závisí aj od typu štúdie.

1.3.4 Metódy zberu dát

Medzi metódy zberu dát patria (1) observačné metódy, (2) rozhovor/dotazník a (3) dokumentácia.

Observačné metódy predstavujú priame pozorovanie, antropologické merania, klinické, biochemické alebo mikrobiologické vyšetrenia a pod. Spôsoby merania lineárnych, oblúkových, obvodových a uhlových rozmerov v *klasickej (tradičnej) morfometrii* sú previazané s chybami merania. Na meranie sa používajú tieto nástroje (Katina a kol., 2011; Hrdlička, 1920, 1939/1952; Martin, 1914/1928; Martin a Saller, 1957–1966; Knussmann, 1988; Fetter, 1967; Preedy, 2012):

- *pomocné nástroje* – kraniofor (na nastavenie a upevnenie lebky orientovanej vo frankfurtskej horizontále), kovové ihlice (na nastavenie príslušných priamok na lebke),
- *meracie nástroje* – dotykové meradlo (kefalometer, na meranie lineárnych rozmerov, napr. M1 = dĺžka lebky), posuvné meradlo (na meranie lineárnych rozmerov, napr. M52 = výška očnice), koordinátové (hĺbkové) meradlo (na meranie projekčných mier a hĺbok na lebke, napr. M20 = nadušná bregmatická výška), uhlomer (na meranie uhlov, napr. M73 = uhol profilu nosa), mandibulometer (na meranie rozmerov sánky, napr. M68 = dĺžka sánky), pásové meradlo (na meranie oblúkových a obvodových mier, napr. M27 = mediánny parietálny oblúk alebo M23 = horizontálny obvod lebky cez glabellu); antropometer (na meranie výšky postavy a telesných proporcií), dynamometer (na meranie sily stisku), bioimpedančná váha (na stanovenie hmotnosti a telesného zloženia), kaliper na meranie kožných rias (na stanovenie obsahu telesného tuku).

Príklad 30 (príklady skupín meraní) *Merania môžeme zaradiť do nasledovných skupín*
Lekárska anamnéza: údaje o diagnóze, liečení, operáciách, symptómoch.
Psychosociálne faktory: rodinná anamnéza, depresia.
Antropometrické premenné: dĺžkové, šírkové, obvodové, objemové miery, uhly, indexy.
Biochemické markery/premenné: cholesterol v sére, fibrinogén v plazme.
Genetické/molekulárne testy: jednoduchý nukleotidový polymorfizmus, typ ľudského leukocytového antigénu.
Obrazový materiál: hustota kosti, vápnik v krvi.
Elektromechanický materiál: arytmia, kongenitálna srdcová choroba a pod.

Rozhovor a dotazník zhromažďujú údaje pomocou zámerne kladených otázok. Získané údaje môžu byť skreslené nepochopením otázok, ich nejednoznačnosťou, zlou formuláciou, zlým záznamom odpovedí a pri rozhovore aj vplyvom sociálnej interakcie. Medzi najpoužívanejšie techniky patrí **metóda štandardizovaného rozhovoru pomocou dotazníka** realizovaná školenými anketármi, koordinátormi. Pri tvorbe dotazníka je potrebné vytvoriť na meranie zisťovaných javov vhodné škály reprezentované predpísanými odpoveďami v otázkach. Základné typy škál sú (1) *nominálna (neusporiadaná)*, (2) *ordinálna (usporiadaná)* a (3) *kvantitatívna (numerická)*. Škála znaku musí by zvolená tak, aby dobre rozdeľovala znak do kategórií. Nemá napr. zmysel zaradiť do dotazníka takú otázku, na ktorú je možná jediná odpoveď. Otázky v dotazníku delíme na *zatvorené* (je možné vybrať si jednu z predvoleného konečného počtu odpovedí) a *otvorené* (odpoveď je potrebné napísať slovne a rozsah slov nie je ohraničený). Špecifickú skupinu tvoria *otázky s viacerými možnosťami odpovede*, ktoré však nie sú štatistickým znakom v zmysle definície – *štatistický znak je jednoznačná transformácia* ($R \Rightarrow O$, kde R označuje množinu respondentov a O označuje množinu odpovedí na danú otázku). Pri ich spracovaní treba využiť špecifické techniky, takže vhodnejšie je vyhnúť sa v dotazníku tomuto typu otázok. Z definície štatistického znaku, ako jednoznačnej transformácie, dostávame **základné štatistické požiadavky**, ktoré je potrebné rešpektovať *pri konštrukcii škál otázok* dotazníka:

- *jednoznačnosť* – zaručuje disjunktný rozklad množiny objektov, t.j. jeden objekt nemôže mať priradenú viac než jednu hodnotu daného znaku;
- *úplnosť* – zaručuje zahrnutie všetkých možných hodnôt, t.j. škála znaku (otázky) musí byť konštruovaná tak, aby zahŕňala všetky možnosti;
- *nezávislosť* – pri všetkých podmienkach musí platiť rovnaká škála (ráno a večer, v Prahe, ako aj v Bratislave) v rámci daného výskumu.

Všetky tri body platia nielen pre škály otázok dotazníka, ale aj všeobecne pre akýkoľvek spôsob získavania dát, t.j. nie je možné mať dve alternatívne namerané hodnoty pre jeden objekt, nesmie sa stať, že nejaký objekt nebude možné zmerať a za všetkých podmienok je potrebné na rovnakom objekte namerať to isté (t.j. akákoľvek metóda získavania dát musí byť *vyčerpávajúca*, *exkluzívna* a *nezávislá*).

Príklad 31 (úplná vs. neúplná škála) *Majme znak „vzdelanie“ s kategóriami – základné, stredoškolské bez maturity, stredoškolské s maturitou a vysokoškolské. Ak by sme v dotazníku vynechali čo i len jeden variant, nemáme úplnú škálu.*

Formálna úprava dotazníka musí by prehľadná, aby sa minimalizovali chyby pri kódovaní, ako aj pri ich zázname do PC. Okrem takejto úpravy musí dotazník obsahovať *identifikačné znaky*, ktoré sú pomôckou pri kontrolách, t.j. číslo dotazníka či číslo koordinátora. Okrem meritórnych znakov (otázok) musí dotazník obsahovať patričné demografické znaky podľa skúmanej témy, napr.

- *kraj*
  - Slovenská republika: Bratislavský kraj, Trnavský kraj, Trenčiansky kraj, Nitriansky kraj, Žilinský kraj, Banskobystrický kraj, Prešovský kraj, Košický kraj;
  - Česká republika: Středočeský kraj, Jihočeský kraj, Plzeňský kraj, Karlovarský kraj, Ústecký kraj, Liberecký kraj, Královéhradecký kraj, Pardubický kraj, Vysočina, Jihomoravský kraj, Olomoucký kraj, Moravskoslezský kraj, Zlínský kraj, Hlavní město Praha;
- *veľkostná skupina obce* – zaužívaná kategorizácia: do 1999 obyvateľov (1), 2000−9999 (2), 10000 −49999 (3), 50000−99999 (4), nad 100000 (5);
- *okresy* – vhodne kódované číslami $(1, 2, \ldots, k)$ alebo jednoduchými skratkami ich názvov; podľa potreby možno kategorizáciu usporiadať podľa rozlohy okresu alebo počtu obyvateľov.

Ďalšie demografické ukazovatele možno prípadne doplniť podľa potreby.

V prípade, že kategórie znaku majú malú početnosť, zlučujú sa podľa logiky problému. V dotazníku sa môžu vyskytnúť otázky, na ktoré neodpovedajú všetci respondenti. Takýmto prípadom sa treba striktne vyhýbať, lebo spracovanie takýchto otázok vnáša do štatistickej analýzy určité problémy, ktorým možno predchádzať úplným vyplnením dotazníka[9]. Vhodné je v jednotlivých otázkach zachytávať jedno hľadisko, pretože pri kombinácii viacerých hľadísk sa musia tieto kombinácie variantov znakov „násobiť“.

Príklad 32 (úplná škála) *Ovplyvnila alebo neovplyvnila prebiehajúca verejná diskusia o vstupe štátu XX do inštitúcie YY Váš postoj k ZZ?*
1) ovplyvnila, nechcel som sa ho zúčastniť, ale zúčastním sa ho;
2) ovplyvnila, chcel som sa ho zúčastniť, ale nezúčastním sa ho;
3) neovplyvnila, chcel som sa ho zúčastniť a zúčastním sa ho;
4) neovplyvnila, nechcel som sa ho zúčastniť a nezúčastním sa ho;
5) neviem posúdiť.

Príklad 33 (úplná škála) *Očakávate zmenu v YY?*
1) určite áno;
2) asi áno;
3) neviem;
4) asi nie;
5) určite nie.

Príklad 34 (úplná škála) *Podporujete vstup XX do YY?*
1) vstup som podporoval/a a podporujem ho aj naďalej;
2) vstup som podporoval/a, ale teraz ho nepodporujem;
3) vstup som nepodporoval/a, ale teraz ho podporujem;
4) vstup som nepodporoval/a a naďalej ho nepodporujem.

[9] Neúplne vyplnený dotazník predstavuje z hľadiska reprezentatívnosti dát rovnakú situáciu ako úplné odmietnutie účasti jedinca v štúdii (t.j. jedinec sa nestane dobrovoľníkom).

Niekedy môže byť výhodné dva alebo viac znakov kombinovať do jedného bez straty informácie, napr. dve otázky do jednej.

Príklad 35 (kombinácia znakov) *Znak XA nadobúda hodnoty* 0 *(v úrade nemajú PC) a* 1 *(v úrade majú PC), znak XB (vek PC) je X rokov; je možné kombinovať do jednej otázky XX, kde kód 0 znamená, že PC nemajú a ostatné odpovede sú už len nominálne, teda vek PC. Takto sa redukuje počet otázok (premenných) a eliminujú sa chýbajúce dáta.*

V kvantitatívnom výskume je potrebné vyhýbať sa všetkým typom **otvorených otázok**, ako *otvoreným otázkam so stručnou odpoveďou*, tak aj *otvoreným otázkam so širokou odpoveďou.* Najvýhodnejšími typmi otázok z pohľadu ďalšieho matematicko-štatistického spracovania sú *otázky dichotomické (binárne)*, *otázky priraďovacie* (či už priraďujeme do usporiadanej alebo neusporiadanej kategorizácie) a *otázky s numerickou odpoveďou.* Všetky takéto vhodné otázky môžeme zahrnúť pod jeden pojem **uzavreté otázky**. Treba ich formulovať *stručne, jasne, zrozumiteľne, úplne* a *jednoznačne.* Navrhovať treba *navzájom nezávislé otázky* (porovnaj s príkladom 35), čo znamená, že odpoveď na nasledujúcu otázku by nemala závisieť od odpovede na predchádzajúcu otázku.

Dokumentácia je jednoduchý a často jediný spôsob ako získať informácie z minulosti. V lekárskom výskume (alebo aj v klinickej antropológii) rozoznávame tri typy dokumentácie, a to

1. **pôvodná zdravotnícka dokumentácia** (ako je *záznam o zdraví a chorobe („karta“), hlásenie o narodení dieťaťa* alebo *list o prehliadke mŕtveho*);
2. **údaje rutinnej zdravotníckej štatistiky a rutinných štatistík iných odvetví**, napr. demografie, poisťovníctva, ekonomiky, geografie a meteorológie (*údaje o zdravotnom stave obyvateľstva* – o celkovej chorobnosti, chorobnosti spojenej s pracovnou neschopnosťou, príčinách úmrtia a príčinách invalidity, ako aj *o zdravotníckych službách* – o sieti zdravotníckych zariadení, pracovníkoch v zdravotníctve, zdravotníckych školách o lekárskych a farmaceutických fakultách);
3. **historická dokumentácia**, napr. cirkevné matriky; uplatňuje sa v biologickej antropológii a príbuzných odboroch (napr. behaviorálne-ekologické štúdie v rámci teórie životnej histórie); umožňujú sledovať zaznamenané udalosti z celého života ľudí aj v niekoľkých po sebe idúcich generáciách.

1.3.5 Zisťovanie presnosti merania

Na zistenie spoľahlivosti nameraných dát je vhodné, a v odborných prácach dnes už nevyhnutné, číselne charakterizovať **presnosť merania** (napr. Goto a Mascie-Taylor, 2007). Výsledky merania sú vždy ovplyvnené kombináciou systematických a náhodných chýb merania. **Systematické chyby** spôsobené meraním môžeme rozdeliť na tieto skupiny (Katina a kol., 2011):

1. **chyby spôsobené externými/environmentálnymi faktormi** – denná doba, intenzita svetla, vlhkosť prostredia a oblečenie;
2. **chyby prístroja** – presnosť merania prístroja;
3. **chyby merania** – chyby z odlišnej aplikácie techniky merania (rôzne pochopenie definície meranej miery), intraindividuálna a interindividuálna chyba (iné držanie prístroja, iný tlak aplikovaný pri meraní, iná orientácia lebky pri meraní a pod.);
4. **kalibrácie meracieho prístroja** (často sa používa aj anglický pojem *zero error*), napr. MicroScribe G2, kde je potrebné opakovane kalibrovať prístroj na tri fixne stanovené landmarky pri otočení tej istej lebky a pod.

Zmiešaním chýb (1) až (4) vzniká tzv. **kombinovaná systematická chyba**, ktorú nie je možné objektívne hodnotiť. Problematické tiež je, keď sa kombinujú miery (na výpočet indexov, ako aj v štatistických výpočtoch) merané inými meracími prístrojmi s rôznou presnosťou merania (zvyčajne od zlomku milimetra do troch milimetrov). Pri meraniach všeobecne môžeme hovoriť aj o **náhodnej**

chybe[10] (šume), ktorá je dôsledkom nesprávneho náhodného výberu. Pri antropologických meraniach na historických populáciách však ide o špecifický problém, lebo pri výskume kostrových sérií sa merajú všetky nájdené lebky. Keďže tento druh výberu nie je možné ovplyvniť, nedá sa hovoriť o náhodnom výbere v pravom slova zmysle. Výber je potom ovplyvnený len dostatočnou zachovanosťou lebiek a veľkosťou kostrovej série, ktorá nemusí dostatočne reprezentovať populáciu žijúcu v danom čase a priestore.

Obidva typy chýb, systematické ako aj náhodné, sa vo výsledkoch kombinujú. Pokiaľ meranie nie je dostatočne presné, náhodný šum výrazne ovplyvňuje možnosť zachytenia štatisticky významného rozdielu. Autor tak čitateľovi dáva vedieť, že napríklad medzi populáciami nie je v strednej hodnote sledovaného znaku rozdiel, aj keď v skutočnosti medzi nimi rozdiel je. Čitateľ je tak uvedený do omylu, pretože nie je možné odlíšiť túto situáciu od situácie skutočnej zhody medzi dvoma strednými hodnotami, v dôsledku čoho svoj ďalší výskum orientuje nesprávnym smerom. Presnejšia štúdia rozdiel nájde, čím vznikne rozpor medzi oboma výsledkami, pričom nie je možné zistiť, či ide o skutočný rozdiel medzi dvoma štúdiami (v jednej sa dve skupiny nelíšia a v druhej sa líšia) alebo ide len o dôsledok odlišnej presnosti merania a miery šumu v dátach oboch štúdií. Je preto v záujme samotného autora nielen uvádzať v publikáciách údaje o presnosti merania (ako intraindividuálnu, tak aj interindividuálnu chybu; pozri kapitolu 1.1 Ciele vedeckej štúdie a sledované premenné), ale predovšetkým mať jasnú predstavu o tom, čo a ako presne meria. Len tak sa môže rozhodnúť, či vôbec meranie použije na analýzu a ako bude výsledky interpretovať.

Na vyjadrenie presnosti merania sa najčastejšie používajú absolútne a relatívne ukazovatele vzťahu medzi rozptylom nameraných hodnôt znaku a rozptylom spôsobeným rozdielom medzi dvoma opakovanými meraniami (Ulijaszek a Lourie, 1994). Najvhodnejšie je dvakrát zmerať všetky prípady danej vzorky, v niektorých špeciálnych metódach aj viackrát (napr. pri hodnotení fluktuačnej asymetrie sa používa opakované meranie priamo ako faktor vstupujúci do analýzy rozptylu a porovnáva sa jeho vplyv s efektom fluktuačnej asymetrie). V mnohým prípadoch však z praktických, finančných či etických príčin chybu merania opakovaním merania všetkých prípadov stanoviť nemožno. Vtedy je potrebné stanoviť chybu merania aspoň na časti vzorky. Rôzni autori odporúčajú rôznu minimálnu veľkosť vzorky na stanovenie chyby merania, napr. 50 prípadov (Mueller a Martorell, 1988), pokiaľ sa štúdia nezaoberá porovnávaním viacerých skupín, u ktorých sa presnosť merania môže líšiť (napr. ťarchavé a neťarchavé ženy), v takom prípade je potrebné stanoviť chybu merania v každej skupine zvlášť.

Jedným z absolútnych ukazovateľov je **technická chyba merania**

$$\mathrm{TEM} = \sqrt{\frac{\sum_{i=1}^{n} x_{d,i}^2}{2n}},$$

kde $x_{d,i} = x_{i1} - x_{i2}$ je rozdiel medzi meraniami. TEM vyjadruje priemernú chybu merania, ktorá pripadá na jedno meranie. Pokiaľ sa hodnota TEM rovná napríklad 1.118 mm, znamená to, že pri každom meraní na jedno meranie tohto znaku pripadá priemerne chyba 1.118 mm. Jednotkou TEM sú pôvodné jednotky, v ktorých bola premenná meraná. Rôzne TEM hodnoty možno porovnávať iba v rámci tejto premennej, t.j. nemožno porovnávať číselné hodnoty TEM u výšky postavy a sily stisku ruky.

Ďalším ukazovateľom je **relatívna technická chyba merania**

$$\mathrm{TEM_{rel}} = \frac{\mathrm{TEM}}{\overline{x}} \times 100,$$

kde $\overline{x}$ je celková priemerná hodnota. $\mathrm{TEM_{rel}}$ predstavuje percentuálne vyjadrený podiel TEM k priemernej hodnote danej premennej. Pokiaľ je hodnota $\mathrm{TEM_{rel}}$ napr. 2.42 %, pripadá na každé meranie chyba predstavujúca 2.42 % z priemernej hodnoty. Tento ukazovateľ teda nezávisí len na samotnej presnosti merania, ale tiež na priemernej hodnote. Tou môže byť populačný priemer tabelový (publikovaný), priemer vypočítaný z hodnôt nameraných v danom výskume alebo priemerná hodnota z takto testovaných opakovaných meraní. Na rozdiel od TEM, porovnávanie $\mathrm{TEM_{rel}}$ medzi rôznymi veličinami je zmysluplné. Vyššie hodnoty $\mathrm{TEM_{rel}}$ majú v antropológii spravidla rozmery merané na živom človeku a na mäkkých tkanivách a menšie rozmery merané na skelete, čo je dané komplikovanejšou špecifikáciou veličín a štandardizáciou podmienok merania u živého človeka (mäkké tkanivá, cirkadiánne rytmy, psy-

[10] Za náhodné chyby považujeme aj **chyby registrácie** – chyby z odčítania hodnôt z meracieho prístroja, chyby zo zápisu hodnôt do protokolu, chyby z prenosu hodnôt z protokolu do PC.

chický stav, atď.). Keďže je tu chyba vztiahnutá k strednej hodnote, u znakov meraných s rovnakým TEM (dané napr. obdobnou definíciou veličiny a rovnakou presnosťou meradla), ale s rôznou strednou hodnotou (môže to byť napr. dĺžka kosti stehennej a druhej kosti záprstnej, obe merané na osteometrickej doske), bude relatívna chyba veličiny s menšou priemernou hodnotou napriek rovnakej TEM väčšia.

Tretím ukazovateľom je **koeficient reliability**

$$\mathrm{CR} = (1 - \frac{\mathrm{TEM}^2}{s^2}) \times 100,$$

kde s je celková smerodajná odchýlka. CR vyjadruje podiel druhej mocniny TEM k celkovému rozptylu (druhej mocnine smerodajnej odchýlky) hodnôt sledovanej vzorky, odčítaný od hodnoty jedna. Ide teda opäť o ukazovateľ relatívny, ktorý udáva, aký podiel rozptylu v dátach *nie je* tvorený chybou merania. Nadobúda hodnoty od 0 (absolútne nereliabilné meranie) do 1 (absolútne reliabilné meranie). Aj keď nie je stanovený žiadny limit určujúci, akú najnižšiu hodnotu by mal CR dosiahnuť, aby sme mohli namerané údaje použiť, Ulijaszek a Kerr (1999) uvádzajú, že v antropometrii by mal dosahovať hodnoty minimálne 0.95 a viac. Zo spôsobu výpočtu je zrejmé, že CR závisí na hodnote smerodajnej odchýlky. Ak vypočítame CR pre zlúčený súbor mužov a žien dohromady a pre mužov a ženy samostatne, nižšia hodnotu CR (a teda nižšia reliabilita merania) bude pravdepodobne v zlúčenom súbore, kde bude smerodajná odchýlka výrazne vyššia.

Vzhľadom na umocnenie a odmocnenie rozdielov, TEM (ani ďalšie nasledujúce dva z neho vypočítané ukazovatele – $\mathrm{TEM}_{\mathrm{rel}}$ a CR) nijak nezohľadňuje zmysel rozdielu (záporný či kladný) medzi prvým a druhým meraním a nie je teda ukazovateľom systematickej zmeny merania medzi prvým a druhým meraním. Na otestovanie systematickej zmeny, pozri kapitolu 6.1 Asymptotické testy o strednej hodnote. Vzhľadom na mnohé druhy príčin, v dôsledku ktorých systematická chyba medzi prvým a druhým meraním často vzniká, okrem jej testovania by sme sa systematickej chybe mali snažiť predchádzať metodickými opatreniami. Systematický posun nenastane, ak znáhodníme poradie prvého a druhého merania, t.j. nemeriame najskôr všetky prípady prvýkrát a potom druhýkrát, ale v rámci jedného merania zaradíme každý prípad dvakrát v náhodnom poradí. S možným výskytom systematickej chyby merania medzi oddeleným prvým a druhým meraním súvisí aj nevyhnutnosť znáhodnenia (randomizácie) poradia merania prípadov z porovnávaných skupín (pozri kapitolu 1.3.8 Znáhodnenie a zaslepenie). Ak napríklad študujeme sexuálny dimorfizmus a porovnávame stredné hodnoty u mužov a u žien, nie je správne zmerať najskôr mužov a potom ženy (alebo dokonca skupiny rozdeliť medzi rôznych výskumníkov), pretože môže vznikať systematická chyba merania (v dôsledku zmeny podmienok prostredia, zmeny v chápaní rozmeru na mužskom a ženskom tele v čase, rovnako tak aj odlišnosti medzi dvoma osobami v chápaní definície rozmeru a praktických skúsenostiach, napr. v tlaku vyvíjanom na meradlo pri meraní mäkkých tkanív ľudského tela). Systematická chyba sa potom neošetrená stáva súčasťou študovaného sexuálneho dimorfizmu. V ideálnom prípade preto treba poradie mužov a žien randomizovať a meranie uskutočňovať jedným výskumníkom. Keďže v rozsiahlejších štúdiách to väčšinou nie je možné, pri vačšom počte merajúcich osôb je nutné ešte pred začiatkom merania testovať interindividuálnu chybu u všetkých výskumníkov na rovnakej testovacej vzorke.

Nevhodné na testovanie presnosti merania je použitie korelačného koeficientu (Halligan, 2002) alebo regresného modelu. Hodnoty opakovaných meraní totiž môžu veľmi silno korelovať aj v prípade nereliabilného merania zaťaženého systematickou chybou merania. Ak by sme napríklad pri druhom meraní u všetkých prípadov namerali presne dvakrát väčšiu hodnotu než pri meraní prvom, meranie by bolo veľmi nereliabilné a s veľkým systematickým rozdielom medzi prvým a druhým meraním, korelačný koeficient by však bol napriek tomu rovný jednej.

1.3.6 Štatistické znaky

Na štatistických jednotkách skúmame **štatistické znaky** (**premenné**), ktoré sa delia do nasledovných kategórií (hovoríme aj o **škále merania**):

1. **kvantitatívne (metrické, kardinálne)**
   - **spojité** (merané na nekonečný počet desatinných miest), **diskrétne** (merané na konečný počet desatinných miest; často sa im hovorí aj diskretizované spojité znaky, lebo znak nemôžeme zmerať s absolútnou presnosťou; napr. dĺžkové miery lebky) a **intervalové** (jav, ktorý

nevieme presne pozorovať, napr. termíny a vekové kategórie v rokoch používané pri odhade veku podľa kostry[11] (Szilvássy, 1988): infans I $\langle 0, 7)$, infans II $\langle 7, 14)$, juvenis $\langle 14, 20)$, adultus I $\langle 20, 30)$, adultus II $\langle 30, 40)$, maturus I $\langle 40, 50)$, maturus II $\langle 50, 60)$, senilis $\langle 60, \text{vek}_{\max}\rangle)$ [vysoká informačná hodnota];

- **frekvencie/početnosti výskytu nejakej udalosti** [vysoká informačná hodnota], kde ide napr. o počet vyfajčených cigariet za deň a pod.;

2. **kvalitatívne**
   - **nominálne znaky** (**neusporiadaná kategorizácia**) – ide o príslušnosť sledovaného objektu k určitej triede objektov, napr. vyučovacie predmety; naboženstvá ako katolíci, protestanti, židia, moslimovia; typ sídla – dom, apartmán; pohlavie[12] – žena, muž; krvný typ – A, B, AB a 0 [nízka informačná hodnota, ak ide o intervalovú premennú vytvorenú z nejakej spojitej premennej; inak vysoká informačná hodnota];
   - **ordinálne znaky** (**usporiadaná kategorizácia**) – ide o znak, ktorého hodnoty môžeme prirodzene usporiadať, napr. stupnica známok – $0, 1, \ldots, 5$; sociálna skupina – nízka, stredná, horná; politická filozofia – liberálna, stredná, konzervatívna; stupeň bolesti (podľa WHO) – 0 (žiadna), 1 (mierna bolesť), 2 (vnímaná a obťažujúca), 3 (silná a stresujúca), 4 (veľmi silná a neznesiteľná), 5 (zničujúca) [stredne vysoká informačná hodnota];
   - **alternatívne znaky** (**binárne**) – typ nie/áno, kedy môže ísť o neusporiadanú kategorizáciu, napr. pohlavie – 0 (žena), 1 (muž) alebo usporiadanú kategorizáciu – 0 (nemá znak), 1 (má znak) [nízka až stredne vysoká informačná hodnota].

Frekvencie/početnosti hovoria o **absolútnej škále znaku**, kým pravdepodobnosti a percentá o **relatívnej škále znaku**. Pri zaznamenávaní detailov o sledovaných premenných je potrebné zaznamenať nielen *typ premennej* (škálu merania), ale aj jej *presné pomenovanie* a *skratku v databáze*, *jednotky*, *stručný popis* a *možný rozsah hodnôt*.

Realizácie štatistických znakov (dáta) môžu mať rôznu formu – **skalár**, **vektor**, **matica**, **pole** alebo **dátová tabuľka** (kombinuje číselné a textové údaje), z ktorých možno získať napr. **incidenčnú maticu** (matica neprítomnosti a prítomnosti – nuly a jednotky), **abundančnú maticu** (matica abundancií; v každej bunke sa vyskytuje počet) alebo **tabuľky početností** atď.

1.3.7 Triedenie dát

Triedenie dát slúži na rozdelenie výberových jednotiek do skupín (tried) podľa dopredu určených triediacich znakov. Podľa počtu triediacich znakov rozlišujeme

- **jednostupňové triedenie** – len jeden triediaci znak, napr. triedenie novorodencov podľa pohlavia;

- **viacstupňové (kombinačné) triedenie** – dva a viac triediacich znakov, napr. triedenie zomretých osôb podľa veku, pohlavia a zamestnania.

S typom triedenia súvisí aj následne použitý štatistický model, v ktorom možno triedenie použiť pre závislé, ako aj nezávislé premenné.

Triedy sú pri triedení určené podľa kvantitatívnych znakov pomocou *triednych intervalov*, ktoré musia pokryť všetky hodnoty sledovaného kvantitatívneho znaku a vzájomne sa nesmú prekrývať (musia byť disjunktné). *Hranice triednych intervalov* sú dané číselne a ich rozdiel predstavuje *dĺžku triedneho intervalu*. *Stred triedneho intervalu* je aritmetický priemer hraníc triedneho intervalu. Rovnako dlhé triedne intervaly sa nazývajú *ekvidištantné*. *Počet triednych intervalov* je najčastejšie od 5 do 20 (niekedy aj menší, podľa potreby), je určený s ohľadom na rozsah súboru a jeho rozpätie (viac pozri v kapitole 3 Charakteristiky polohy a variability a štatistická grafika).

[11] V súčasnosti sa v kostrovej antropológii pochybuje o správnosti zaraďovania jedincov do vekových kategórií v dospelosti v desaťročných intervaloch. Možnosťami reálneho a validného hodnotenia veku dospelých jedincov na základe znakov na kostre sa zaoberali Buk a kol. (2012) a Brůžek (2008).

[12] Pohlavie všeobecne chápeme ako nominálny znak (kódovanie: 0 – žena, 1 – muž); ako ordinálny znak vystupuje len vtedy, keď sledujeme riziko choroby súvisiace s pohlavím, kde nulou označujeme menej rizikové pohlavie.

Príklad 36 (triedenie veku) *Pri sledovaní pacientov po operácii bedra a kolena ich môžeme rozdeliť do nasledovných vekových kategórií (tried): (1)* $\langle x_{\min}, 55)$ *rokov, (2)* $\langle 55, 65)$ *rokov, (3)* $\langle 65, 75)$ *rokov a nakoniec (4)* $\langle 75, x_{\max}\rangle$ *rokov.*

1.3.8 Znáhodnenie a zaslepenie

Znáhodnenie sa používa hlavne pri meraní/zázname dát. Štatistické jednotky (subjekty) sa náhodne poprehadzujú pomocou ich identifikačných čísel tak, aby napr. meranie nebolo ovplyvnené usporiadaním náhodného výberu (poradím subjektov). Znáhodnenie alebo **znáhodnená alokácia**, t.j. náhodné prideľovanie subjektov do skupín, je fundamentálnym princípom klinických štúdií. Cieľom znáhodnenej alokácie je redukcia výchylky, ktorá sa vyskytuje pri nenáhodnom priraďovaní subjektov do skupín. V podstate ide o aplikáciu princípov náhodného výberu (najčastejšie stratifikovaného výberu), pričom detaily implementácie majú byť uvedené v protokole v sekcii Štatistické metódy. Pokiaľ je súčasťou znáhodňovania aj **stratifikácia**[13], je potrebné popísať **prognostické premenné/faktory** a počet hladín/úrovní pre každú z nich (napr. pohlavie, závažnosť ochorenia a pod.). Znáhodnená alokácia prebieha v každej **stratifikačnej skupine** (kombinácii hladín všetkých faktorov). Je potrebné dodržiavať pravidlo použitia malého počtu faktorov a hladín, pretože počet stratifikačných skupín rýchlo stúpa.

Príklad 37 (znáhodnená alokácia; permutačný blokový princíp) *Majme dve skupiny A a B (vo všeobecnosti môžeme mať k skupín), do ktorých chceme náhodne alokovať subjekty.* ***Alokačný pomer*** *(podiel rozsahov týchto dvoch skupín) je* $1:1$. *Alokácia do týchto dvoch skupín prebieha pomocou* $(m \times k)$*-rozmerných blokov (m je nejaké prirodzené číslo; keďže máme 2 skupiny, $k = 2$). Nech $m = 2$. Potom je blok zložený zo štyroch elementov (dva pre každú skupinu). Možné usporiadania (permutácie) elementov v blokoch sú nasledovné – $AABB$, $BBAA$, $ABBA$, $BAAB$, $ABAB$ a $BABA$. Ich poradie je určené náhodným usporiadaním čísel* 1 *až* 6. *Celkový počet písmen A a B zodpovedá rozsahu náhodného výberu.*

Príklad 38 (stratifikovaná znáhodnená alokácia) *Chceme náhodne alokovať subjekty do dvoch skupín v pomere* $1:1$. *Majme premenné pohlavie (hladiny – žena a muž) a závažnosť ochorenia (hladiny –* 1 *a* 2*). Budeme mať potom štyri stratifikačné skupiny (obe závažnosti ochorenia pre ženy ako aj pre mužov) a v každej z nich vykonáme znáhodnenú alokáciu z predchádzajúceho príkladu.*

Zaslepenie/maskovanie sa používa na redukciu výchylky, ktorá by mohla vzniknúť v prípade, že má výskumník informácie o zaradení subjektov do skupín (typ vyšetrenia, typ choroby, prípad alebo kontrola a pod.), pretože tieto informácie môžu ovplyvniť jeho rozhodovanie. Špeciálnym prípadom zaslepenia je **dvojité zaslepenie**, ktoré sa v klinických pokusoch používa v prípadoch, keď ani výskumník, ani subjekt nepoznajú zaradenie subjektu do príslušnej skupiny. V protokole sa takáto situácia opíše nasledovne: „Ide o dvojito zaslepenú znáhodnenú štúdiu, ktorá zisťuje efekt lieku XY a placeba (alebo lieku XZ) na incidenciu infarktu myokardu bez a s následkom smrti (celková mortalita) u n pacientov po prvom infarkte myokardu.“ Je potrebné špecifikovať, čo sa bude znáhodňovať a kroky, ktoré budú vykonané na zabezpečenie znáhodnenia v praxi. Je tiež dôležité zdôrazniť, že ani ďalšie osoby, ako napr. rádiológ, patológ a laboratórny personál, nemajú prístup k zatriedeniu subjektov do skupín. Rovnako tak pri antropologickom meraní, či už klasickom alebo virtuálnom, je potrebné zabezpečiť, aby ani výskumník nemal prístup k zatriedeniu subjektov do skupín. Optimálne je preto použiť identifikačné čísla subjektov v náhodnom poradí.

1.3.9 Databáza

Celistvosť a kvalita štatistickéj analýzy závisí od kvality dát. Každá štúdia musí vychádzať z vysoko kvalitných dát, ale kroky k dosiahnutiu tohto cieľa sú často prehliadané. V angličtine sa používa slovné spojenie „*garbage in – garbage out*“, v slovenčine sa štatistika definuje ako „*presná práca s nepresnými číslami*“. V protokole je potrebné uviesť, akým spôsobom budú dáta zozbierané, akým spôsobom bude

[13] Z latinského *stratum* (singulár) a *strata* (plurál); strátum v štatistike sa prekladá ako vrstva alebo hladina.

zaslepená identifikácia subjektov, akým spôsobom sa budú dáta uchovávať a chrániť. Práca s dátami v elektronickej podobe sa nazýva **dátový manažment**. V prípade veľkých databáz je potrebné kontaktovať informatika špecializujúceho sa na databázy, t.j. **databázového programátora**, ktorý je schopný programovať v jazyku SQL a **bioštatistika**, ktorý je schopný programovať v jazyku R. Bioštatistika je potrebné kontaktovať v každom prípade, pretože sa musí podieľať na **databázovom projekte**, ktorý zahŕňa plánovanie, definíciu a realizáciu databázy, vytváranie dátových formulárov, spôsob zadávania dát, monitorovanie a revíziu databázy. Dobre organizovaná databáza zabezpečí nielen dobré uchovanie dát, ale uľahčí aj hľadanie subjektov, kontrolu kvality, vlastnú štatistickú analýzu a prezentáciu výsledkov.

Pri tvorbe formulárov treba mať na zreteli kto, kedy a kde bude formuláre vypĺňať, kedy budú dáta dostupné a či bude formulár súčasťou databázy. Je vhodné vytvoriť **zoznam všetkých potrebných formulárov**, v ktorom budú uvedené názvy súborov a miesto ich uloženia (Peacock a Peacock, 2011). V každom prípade je dôležité mať

- **zoznam premenných, ich typ, dátum a jednotky merania**; v analýze tvaru aj **zoznam mier, landmarkov a kriviek**; všetko s definíciami premenných, vysvetlením skratiek (forma skratiek musí byť konzistentná s pravidlami vytvárania názvov stĺpcov v použitom štatistickom programe, napr. R), číselným formátom hodnôt (počet desatinných miest, počet cifier, forma dátumu, napr. *dd/mm/rrrr*), s očakávaným rozsahom a očakávanou hodnotou, definíciou kódovania v tzv. vyhľadávacích tabuľkách, s uvedením spôsobu identifikácie chýbajúcich a neposkytnutých dát (v databáze sa nesmú nechávať prázdne bunky, optimálne je definovať tieto hodnoty tak, aby nemohli byť zamenené s reálnymi hodnotami, napr. 999 alebo -9 a pod.), s definíciou pravidiel validácie, spôsobu získania dát (virtuálne z obrazu, fyzickým meraním, akým vyšetrením, akým prístrojom), spôsobom merania, použitím dát (popisná alebo štatistická analýza); sumárne tomu hovoríme **dátový slovník**;

- **záznam o identifikácii subjektov**, ktorý je potrebný na anonymnú identifikáciu subjektov; **kontaktné informácie o subjekte**;

- **výstupný záznam o subjekte** – kedy subjekt opustil štúdiu, status ukončenia a dôvod.

Kontrola kvality databázy zahŕňa kontrolu

- **opakovaného zahrnutia toho istého subjektu** (napr. aj pomocou rodných čísel),

- **rozsahu danej premennej** (čistenie dát, t.j. zhodnotenie a korekcia; napr. pomocou kontroly posunu desatinnej čiarky ako systematickej chyby, zobrazením dát, kontrolou cez stratifikované premenné, hlavne cez klúčové (primárne) premenné a pod.), v angličtine sa používa pojem *out-of-range measurements*, v štatistike sa špecifickej skupine týchto realizácií hovorí **odľahlé pozorovania**, ktoré sa môžu (ale nemusia) vyskytovať v povolenom rozsahu danej premennej,

- **kvality obrazového materiálu** (napr. kontrola rekonštrukcie 3D obrazu, kontrola dodržania svetelných podmienok, pri osobách aj prítomnosti mejkapu, náušníc, vlasov alebo chlpov brániacich vytvoreniu obrazu a pod.),

- **dátový audit** – kontrola splnenia vylučovacích kritérií, t.j. či sa v súbore nachádzajú len tie subjekty, ktoré sa v súbore nachádzať majú v zmysle vylučovacích kritérií.

Bezpečnosť databázy sa často hodnotí pomocou tzv. **bezpečnostnej trojice**:

1. **Dôvernosť** – ochrana dát pred neautorizovaným vniknutím do databázy – v prípade papierovej podoby databázy sa používa uzamknutá skrinka, v prípade elektronickej podoby databázy ide o enkryptovaný súbor v PC, v inej pracovnej stanici, v laptope alebo na inom mobilnom zariadení, ako je tablet, USB-kľúč alebo externý HDD; ochrana tohto súboru heslom (neodporúča sa nechať pracovnú stanicu neuzamknutú, ani používať všeobecné heslá); databázu je potrebné uchovávať bez možnosti identifikovať subjekty, čo je legislatívne ošetrené napr. v Privacy Rule v dokumente HIPPA (Health Insurance Portability and Accountability Act), ktorý hovorí o ochrane dôverných informácií; v neposlednom rade je potrebné zabezpečiť pravidelné zálohovanie na serveri, protivírusovú ochranu a *firewalls* (najčastejšie prostredníctvom IT tímu inštitúcie).

2. **Úplnosť** – zabezpečenie správnosti údajov zabránením nevhodnej zámene dát, či už náhodnej, alebo cielenej (napr. povolením len čítania dokumentu), ďalej limitovanou aktualizáciou, plným prístupom pre príslušných členov riešiteľského kolektívu.

3. **Dostupnosť** – zabezpečenie dostupnosti autorizovaným používateľom (členovia riešiteľského kolektívu, špecificky napr. pre koordinátora projektu, bioštatistika, databázového programátora a pod.).

Uzamknutie databázy. Priebežné čistenie databázy je potrebné vykonávať pravidelne, konečné čistenie na konci štúdie, keď sa predpokladá, že všetky potrebné údaje o subjektoch sú už skompletizované. V tomto období sa vykoná tzv. **zmrazenie**, čo je časové obdobie kontroly a čistenia, kedy bioštatistik môže odhaliť problémy napr. v podobe odľahlých pozorovaní počas **predbežnej štatistickej analýzy**. Po skončení tohto obdobia sa vytvorí **finálna kópia databázy určená na štatistické účely**, do ktorej môže zasahovať len bioštatistik. Po ukončení analýz sa vykoná tzv. **trvalé zmrazenie** databázy, po ktorom nie je možná žiadna ďalšia zmena ani v rámci autorizovaného prístupu.

1.3.10 Rozsah náhodného výberu

Veľmi dôležitou súčasťou plánovania výskumu je určenie **minimálneho rozsahu náhodného výberu**, ktorý môže ovplyvniť nielen trvanie výskumu a jeho finančnú náročnosť, ale v neposlednom rade aj vlastné štatistické analýzy. Výpočet rozsahu náhodného výberu by mal prebiehať v spolupráci so štatistikom, keďže treba vopred zvážiť, aký štatistický model sa použije v následných analýzach. Rovnako treba zohľadniť oprávnenosť použitia tohto modelu, ako aj dostatočnú reliabilitu získaných odpovedí na sledované otázky. Na výpočet sú potrebné tieto informácie (Hackshaw, 2009):

1. Správny **štatistický model** konzultovaný so štatistikom. Spolu s typom premenných a ich jednotkami musí byť jasne a zrozumiteľne opísaný.

2. **Veľkosť efektu** je číslo získané z istého štatistického modelu, ktoré charakterizuje sledovanú premennú v súvislosti s testovanou hypotézou, kde efektom je napr. rozdiel priemerov, podiel rozptylov, rozdiel pravdepodobností, pomer šancí, relatívne riziko, Pearsonov korelačný koeficient, regresný parameter a pod. Približnú (predpokladanú) veľkosť efektu zistíme z predchádzajúcich štúdií v rámci tvorby systematického prehľadu z literatúry, pilotnej štúdie alebo zo skúsenosti lekára/antropológa, pričom predstavuje *minimálny (realistický) klinicky/antropologicky dôležitý efekt.* Čím je efekt väčší, tým menší rozsah výberu potrebujeme. Rozsah výberu však v súvislosti so štatistickým modelom nemožno podceňovať, pretože menší rozsah nielenže nemusí dostatočne reprezentovať sledovanú populáciu, ale ani získaný výsledok testovania určitej hypotézy nemusí byť štatisticky signifikantný.

3. **Hladina významnosti** je číslo, ktoré zjednodušene predstavuje maximálnu pravdepodobnosť nájdenia štatisticky signifikantného efektu, ktorý ale reálne neexistuje. V zjednodušenom chápaní vyjadruje maximálnu pravdepodobnosť chybného úsudku.

4. **Sila testu** je číslo, ktoré predstavuje pravdepodobnosť nájdenia štatisticky signifikantnej veľkosti efektu (rovnej alebo väčšej ako nejaké dopredu zvolené číslo), ak signifikantný efekt v realite skutočne existuje.

Príklad 39 (štúdia prípadov a kontrol) *Cieľom štúdie je hľadanie sily asociácie medzi fajčením matiek a vrodenými poruchami novorodencov, kde **prípady** sú matky s dieťaťom s vrodenou poruchou a **kontroly** matky s dieťaťom bez vrodenej poruchy. Očakávame približne* 25 % *fajčiarok medzi kontrolami, kde predpokladáme nájdenie pomeru šancí (príslušný efekt) približne* 1.75. *Pri* 80% *sile a* 5% *hladine významnosti pri dvojvýberovom Studentovom t-teste s obojstrannou alternatívnou hypotézou za prepokladu alokačného pomeru* 3 : 1 *(t.j. troch kontrol na jeden prípad), bude minimálny rozsah náhodného výberu* 170 *prípadov a* 510 *kontrol.*

Niekedy sa môže stať, že už vypočítaný **základný minimálny rozsah náhodného výberu** treba z nejakého dôvodu zvýšiť. K tejto situácii dochádza napr. v týchto prípadoch:

- máme niekoľko závislých premenných alebo mnohonásobné porovnávania, čím sa automaticky mení nielen hladina významnosti, ale aj štatistický model (napr. zo Studentovho t-testu na ANOVA model s fixnými efektmi, z jednoduchého lineárneho regresného modelu na mnohorozmerný lineárny regresný model a pod.);
- nemôžeme predpokladať, že pomer počtu zahrnutých do štúdie k počtu vyšetrených bude číslo rovné jednej (t.j. že všetci vyšetrení budú zahrnutí do štúdie), že nejaké subjekty budú pokračovať v štúdii a stratíme o nich ďalšie informácie (napr. v kohortovej štúdii a pod.); preto musíme plánovať štúdiu na základe rozsahu, ktorý hovorí o dostupnej populácii v rozsahu väčšom ako je minimálny a bezprostredne potrebný.

1.3.11 Kontroly a ich hodnotenie

Úspešný nábor či získavanie subjektov je vo fyzickej a klinickej antropológii kľúčom k úspechu každej vedeckej štúdie. Efektívny nábor vyžaduje zabezpečenie potrebnej infraštruktúry, chodu a dokumentácie s cieľom optimalizovať kontroly a ich hodnotenie. Pod **infraštruktúrou** rozumieme miesto, kde sa bude nábor/získavanie konať, a to buď v laboratóriu, alebo v pracovni výskumníka, alebo mimo týchto priestorov. Je potrebné odhadnúť veľkosť výskumného priestoru, identifikovať ho a zariadiť podľa potrieb štúdie. Vybrané miesto musí byť bezpečné z hľadiska uchovávania a ochrany dôverných informácií či už v písomnej, alebo v elekronickej podobe v enkryptovaných dátových súboroch (napr. informovaný súhlas subjektov, dáta, dotazníky, 2D alebo 3D obraz a pod.). Rovnako musí byť pripravnený aj priestor na uchovávanie biologického materiálu. Miestnosť musí obsahovať priestor, v ktorom prebehne interview a vyšetrenia subjektov, ale aj priestor pre výskumníkov, ktorí vyšetrenia vykonávajú.

Propagácia výskumu sa môže konať rôznym spôsobom, napr. pomocou letákov posielaných písomnou alebo eletronickou formou, pomocou internetových stránok či periodickej tlače. Akákoľvek použitá metóda musí byť vopred schválená inštitucionálnou etickou komisiou. Spôsob a miesto propagácie ovplyvňuje výber vzorky a tým aj výsledky štúdie, preto by mali byť zahrnuté v pláne vedeckej štúdie. Najefektívnešie je hromadné posielanie písomnej alebo elektronickej pošty, po čom zvyčajne nasleduje telefonický rozhovor. V mnohých prípadoch je nevyhnutná úzka spolupráca so zdravotníckymi zariadeniami alebo príslušnými organizáciami.

Alokačný čas potrebný na vykonanie interview, vypísanie informovaného súhlasu, zistenie zdravotného stavu (v minulosti a súčasnosti), vyplnenie dotazníka, vykonanie vyšetrení, meraní alebo odobratia krvi na krvné testy a pod. je potrebné dopredu odhadnúť ešte pred návštevou prvého subjektu a pred zistením jeho spôsobilosti zúčastniť sa štúdie. Tiež je potrebné odhadnúť **počet návštev** na skompletizovanie vyšetrení (celkovo, ale aj denne alebo týždenne). Pokiaľ chceme subjekt zaradiť do štúdie, treba ho najskôr vyšetriť alebo zhodnotiť, a preto je potrebné odhadnúť **pomer vyšetrených ku zahrnutým** do štúdie, čo je efektívna miera úspešnosti náboru. Najefektívnejšie je ponúknuť subjektom flexibilný čas návštevy/návštev na zlepšenie zápisu, oboznámiť ich, najlepšie písomnou formou, s dôležitými informáciami (ako napr. poloha, čas, čo očakávame počas vyšetrenia a hodnotenia, dĺžka návštevy a pod.). Vyžaduje sa tiež dodržanie HIPPA, ako aj princípov etickej komisie inštitúcie.

Cieľom vyšetrení je začať proces zistenia spôsobilosti subjektov zúčastniť sa štúdie, zoznámenie sa s potencionálnymi subjektmi, získanie informovaného súhlasu a zhromaždenie dát potrebných na vlastný zápis (zaradenie) subjektov do štúdie. Vyšetrenie a hodnotenie subjektov môžeme rozdeliť do desiatich krokov (Hulley a kol., 2007):

1. **Získanie informovaného súhlasu** – umožnenie realizácie ďalších krokov; potencionálny subjekt dostane kópiu informovaného súhlasu a má dostatočný čas na to, aby si dokument pred podpisom prečítal a pochopil obsah; subjekt tiež dostane po podpise kópiu dokumentu potvrdzujúcu schválenie štúdie inštitucionálnou etickou komisiou.
2. **Školenie subjektov** – zapojenie potencionálnych subjektov a ich informovanie o dôležitosti a cieľoch štúdie, ako aj ich účasti na nej; na tomto mieste je potrebné zdôrazniť, že subjekt je (v istom zmysle) súčasťou výskumného tímu.
3. **Odpovedanie na otázky** – odpovedanie na otázky vybuduje dôveryhodnosť a možnosť vytvorenia pozitívnej väzby medzi subjektom a výskumníkom.
4. **Zhromaždenie dát vzhľadom na spôsobilosť subjektov zúčastniť sa štúdie** – zhrnutie získaných údajov napr. vyplnenie dotazníka, vyšetrenia, laboratórne testy.

5. **Poskytnutie inštrukcií v písomnej a ústnej forme** – o ďalšej návšteve, vyšetreniach alebo o zápise do štúdie.

6. **Dokumentácia dát súvisiacich so štúdiou v podobe formulárov** – podpísaný originál informovaného súhlasu a všetky ostatné dokumenty, do ktorých možno počas štúdie v prípade potreby nahliadnuť.

7. **Kontrola spôsobilosti vykonaná hlavným riešiteľom** – zabezpečenie spôsobilosti subjektu zúčastniť sa štúdie.

8. **Kontrolný zoznam prijímacích a vylučovacích kritérií** – zabezpečenie celistvosti výberu a kontrola dodržiavania protokolu, pričom subjekty musia splniť všetky kritériá špecifikované v protokole (zoznam musí byť vytvorený už pred vyšetrením prvého subjektu a následne striktne dodržiavaný).

9. **Dokumentácia výsledkov vyšetrení, hodnotenia, prijímacích a vylučovacích kritérií** – zaistenie správnosti a celistvosti dát.

10. **Zápis subjektov prijatých do štúdie** – umožnenie pokračovania štúdie v ďalších krokoch.

Informovaný súhlas obsahuje tieto informácie: obsah, ciele a význam štúdie, účel a zdôvodnenie náboru subjektov, náplň a dĺžka účasti subjektov, náplň základnej a experimentálnej časti štúdie, rola a zodpovednosť subjektov a členov vedeckého tímu, zoznam vyšetrení a prípadné riziká a pod. Subjekty musia mať dostatočný čas na prečítanie súhlasu, ako aj zodpovedanie prípadných otázkok pred vlastným podpisom. Pri tvorbe informovaného súhlasu treba rozlišovať medzi štandardnou klinickou praxou a klinickým výskumom; pri ich zámene sa hovorí o *therapeutical misconception.* Z etického hľadiska je potrebné rozlišovať medzi

1. **subjektom klinickej štúdie** – podpisom informovaného súhlasu sa subjekt vzdáva práva na liečbu, ktorá by inak plne zodpovedala jeho potrebám a preferenciám; subjekt musí byť informovaný o kontrolných podmienkach, podávaní neúčinných liečiv (placebo), randomizácii a zaslepení;

2. **subjektom observačnej štúdie** – podpisom subjekt súhlasí s procedúrami a vyšetreniami špecifikovanými v pláne štúdie.

Príklad 40 (vedecká etika) *Schopnosť porozumieť obsahu informovaného súhlasu môže byť limitovaná pre napr.* ***deti, duševne chorých*** *alebo* ***mentálne postihnutých*** *pacientov. Takéto prípady musia byť legislatívne ošetrené a mali by zahŕňať* ***hodnotenie schopnosti subjektu dať súhlas, splnomocnenie k rozhodovaniu*** *alebo špeciálne odporučenie inštitucionálnej etickej komisie.* ***Väzni*** *a* ***podmienečne prepustení*** *tiež vyžadujú pozornosť vzhľadom na obmedzenie slobodnej voľby. Subjekty* ***v trestnom konaní****, stíhané napr.* ***za zneužívanie návykových látok****, tiež vyžadujú pozornosť vzhľadom na etiku, ochranu súkromia alebo pocit bezpečnosti pri účasti v štúdii. Na všetky vyššie uvedené príklady vyžadujúce špeciálnu legislatívnu pozornosť je potrebné myslieť už pri plánovaní vedeckej štúdie.*

Porušenie prijímacích kritérií. Subjekt, ktorý nespĺňa kritériá štúdie, nazývame **neúspešný kandidát**. Vzhľadom na to, že evidenciu vhodnosti subjektov je potrebné neskôr sumarizovať, je potrebné evidovať všetky dôvody vylúčenia. Niektoré štúdie umožňujú opakované vyšetrenia aj neúspešných subjektov po skončení časového intervalu vyšetrení. Ak sa dôvody nesplnenia stanovených kritérií medzi potenciálnymi subjektmi opakujú, na základe počtu neúspešných kandidátov možno protokol modifikovať. Protokol musí byť dostatočne flexibilný a musí umožniť dodatočné doplnenie ďalších kritérií alebo opakované vyšetrenia, ako sa uvádza v nasledujúcich príkladoch.

Príklad 41 (dodatok k protokolu) *Výskumník potrebuje pridať jeden subjekt do databázy, aby mohol ukončiť nábor a potencionálny subjekt spĺňa všetky kritéria až na jedno, BMI* $= 46\ kg/m^2$*. Vylučovacie kritérium je však BMI* $\geq 45\ kg/m^2$*. Na tomto mieste si môžeme položiť niekoľko otázok.*

Máme právo odporučiť zníženie hmotnosti? Môže výskumník urobiť ***výnimku*** *a pridať tento subjekt do štúdie a zdokumentovať dôvod prijatia? Odpoveď na obe otázky znie* **nie**. *Čo urobiť môže? Subjekty, ktoré nesplnili kritériá možno znovu vyšetriť, pokiaľ je opakované vyšetrenie špecifikované v protokole a schválené inštitucionálnou etickou komisiou. Alternatívnou je modifikácia existujúceho protokolu pomocou* ***dodatku***, *ktorá však vyžaduje schválenie.*

Príklad 42 (druhé (dobrovoľné) vyšetrenie) *Výskumník vyšetrí subjektu systolický krvný tlak (SKT) a zmeraná hodnota* $SKT = 128$ *mm Hg. Vylučovacie kritérium predstavuje* $SKT \leq 130$ *mm Hg a bolo dopredu stanovené a schválené inštitucionálnou etickou komisiou s možnosťou* ***druhého (dobrovoľného) vyšetrenia*** *ako súčasť protokolu. Presná formulácia v protokole môže byť nasledovná: „Ak je hodnota SKT subjektu na antihypertenznej liečbe počas prvého vyšetrenia menšia ako* 130 *mm Hg, je možné uskutočniť druhé vyšetrenie merania SKT. Potencionálnym subjektom bude odporučené, aby na toto vyšetrenie prišli do jedného až dvoch týždňov od prvého vyšetrenia. Ak bude hodnota SKT menšia ako* 130 *mm Hg, bude subjekt z ďalšieho výskumu vylúčený.“*

1.3.12 Nábor (získavanie) subjektov vo fyzickej a klinickej antropológii

Alternatívou ku klasickým metódam náboru (letáky, periodická tlač, e-mail a telefonovanie) sú sofistikované multimediálne reklamné kampane, globálne telefonické centrá pre multicentrické štúdie, internetové stránky štúdie. Pomocou nich možno nielen identifikovať potencionálny subjekt, ale aj sa s ním spojiť. Štúdia môže byť dobre naplánovaná, ale bez dostatočného počtu subjektov sa nedá realizovať. Výskumníci musia byť preto kreatívni nielen pri tvorbe stratégie **náboru**, ale aj pri **udržaní si subjektov** a pri **zabezpečení vernosti subjektov**. Tieto stratégie sú často jedinečné a prispôsobené konkrétnej štúdii. Plán náboru sa vytvára už počas plánovania štúdie, kedy si treba strategicky premyslieť, akým spôsobom možno získať prístup k cieľovej populácii, ako prekonať prípadné bariéry zistením charakteristík cieľovej populácie a ako minimalizovať náklady. Nábor je potom možné začať po získaní súhlasu inštitucionálnej etickej komisie.

WWH stratégia. Táto stratégia predstavuje trojicu základných princípov, t.j. kto (*who*), čo (aký alebo aké; *what*) a ako (*how*) v podobe nasledovnej implementácie (McPaul a Toto, 2011):

1. **sledovaná choroba alebo iné podmienky** – definujú, *kto* bude tvoriť cieľovú populáciu; cieľová populácia definuje, *aké* miesta treba použiť na nábor (berúc do úvahy vek, pohlavie, sociálny a ekonomický status, kultúru, etnicitu, gramotnosť a prístupnosť (deti, starí ľudia, väzni a pod.), ako aj prevalenciu choroby); miesta definujú, *ako* vyvinúť realistické stratégie náboru a udržania si subjektov;

2. **prijímacie a vylučovacie kritériá** – vytvárajú sa s cieľom vybrať do štúdie vhodné subjekty, pričom sa vychádza zo systematického prehľadu, z ktorého možno zistiť, *kto* tvoril cieľovú populáciu, *aký* typ dizajnu sa použil, a následne *ako* stanoviť typ najvhodnejšej štúdie na zodpovedanie cieľov a testovanie hypotéz;

3. **spôsob práce** – je potrebné idenfikovať, ktoré procedúry a aký spôsob práce (vrátane súvisiacich komplikácií) je potrebný na testovanie hypotéz (môže sa stať, že subjekty nebudú ochotné zúčastniť sa štúdie, ak budú použité **invazívne metódy** či bolestivé, nepohodlné alebo časovo náročné procedúry); pozadie štúdie a jej význam by mali nasmerovať výskumníkov na správnu cestu, t.j. určiť *aké* procedúry sú potrebné, *kto* a *ako* bude vykonávať tieto procedúry potrebné na zodpovedanie cieľov a testovanie hypotéz;

4. **trvanie štúdie** – tento faktor môže ovplyvniť rozhodnutie potencionálnych subjektov zúčastniť sa štúdie; ak je štúdia priveľmi dlhá, je vysoká pravdepodobnosť straty subjektu; jej dôvodom môže byť presťahovanie sa, odcestovanie alebo rodinné problémy, ako aj choroba, strata záujmu alebo smrť subjektu; týmto problémom sa možno vyhnúť buď kratším trvaním štúdie, alebo získaním väčšieho množstva subjektov, a to po diskusii s bioštatistikom, čo zaručí dostatočnú istotu a zakomponovanie tejto úpravy do plánu štúdie; takýmto plánovaním v predstihu možno

zobrať do úvahy aj nasledovné skutočnosti: *kto* sú potenciálne subjekty, *aká* frekvencia vyšetrení je potrebná, kde budú vyšetrenia vykonávané a *ako* sa získajú dostatočné ľudské zdroje a grantové prostriedky;

5. **zázemie a ľudské zdroje** – vedecký tím (koordinátor, spoluriešitelia vrátane bioštatistika, databázový programátor, laboratórni technici, administratívna podpora a pod.), priestor, školenia a výučba bezprostredne súvisiaca s vedeckou štúdiou (bez školenia alebo nedostatočným školením a výučbou môže byť narušená integrita štúdie a vedeckého tímu a zároveň aj poškodené dobré meno hlavného riešiteľa, ktorý získal na štúdiu grantové prostriedky), tovary a služby, prístup k diagnózam a pod.; pokiaľ sa zdroje hľadajú oneskorene, môže to spôsobiť oneskorený začiatok štúdie, chýbanie potenciálnych subjektov a nedosiahnutie cieľov náboru a neudržanie si subjektov a nakoniec aj neschopnosť ukončiť štúdiu;

6. **grantové prostriedky**.

Príklad 43 (over-recruitment) *Ak bol vypočítaný minimálny rozsah súboru n a očakávaná strata subjektov je* 10 %, *potom upravený rozsah súboru* $n' = n/(1-0.1)$. *Teda ak* $n = 100$, *potom* $n' = 100/(1-0.1) = 111$.

Príklad 44 (prostriedky na nábor) *Prezentácie štúdie medzi kolegami, zabezpečenie všeobecných informácií o štúdii na webovej stránke, na nástenke, letáku a v brožúre, zoznam kontaktných informácií riešiteľského kolektívu, vizitky s kontaktným číslom na koordinátora, informácia o tom, čo sa očakáva na každej návšteve, najčastejšie kladené otázky (FAQs), zabezpečenie zmysluplného transportu, použitie elektronických lekárskych záznamov a pod.*

Príklad 45 (prostriedky na udržanie si subjektov) *Zabezpečenie dopravy, občerstvenie, flexibilný plán náboru realizovaný aj večer a cez víkend, motivačné platby na ocenenie nadčasov, motivačné darčeky, zabezpečenie kalendára štúdie s pripomienkami termínov ďalších návštev, pohľadnice k narodeninám a pod.*

1.4 Úloha bioštatistika pri tvorbe vedeckých štúdií

Vedeckú štúdiu charakterizujeme ako validnú nielen na základe jej výsledkov, ale aj na základe jej plánu. Často sa hovorí „Rob to správne alebo to radšej nerob.“ **(Aplikovaný) výskumník** (biológ, antropológ alebo lekár), ktorý vo svojom výskume aplikuje (používa) metódy štatistickej analýzy (takýto výskum štatistici často nazývajú **aplikovaný výskum**) si často kladú otázku, či potrebujú bioštatistika vo svojom vedeckom tíme. Odpoveď znie jednoznačne áno, pretože bioštatistik môže nielen optimalizovať plán štúdie, analyzovať dáta a interpretovať výsledky, ale môže aj načrtnúť závery. Bioštatistik by sa mal stať členom riešiteľského kolektívu čo najskôr – už na začiatku plánovania štúdie, lebo zle navrhnutú štúdiu, ktorá sa už začala, nemôže zachrániť, rovnako ako nemôže analyzovať zle zozbierané dáta alebo interpretovať výsledky z nevhodne použitých štatistických analýz. Hoci úloha bioštatistika pri štatistickej analýze je celkom očividná, výskumníci si dobre neuvedomujú jeho dôležitosť už pri plánovaní štúdie a vývoji výskumného protokolu. Je dôležité mať na zreteli veľký význam dopredu naplánovaného zberu dát, správne naformulovaných a realistických hypotéz a správne naplánovanej, vykonanej a interpretovanej štatistickej analýzy. Na ich základe musí byť výskumník schopný správne zhodnotiť a interpretovať vedeckú hypotézu/hypotézy. Bioštatistik teda dokáže poradiť aj v dobrej previazanosti hypotéz, ktoré priamo korešpondujú s plánom štúdie.

Interakcia s bioštatistikom. Interakcia (aplikovaného) výskumníka s bioštatistikom je dlhodobý proces, ktorý predstavuje sériu kontaktov, e-mailov alebo telefonátov, počas ktorých výskumník musí koncepčne pochopiť navrhovanú štatistickú metodológiu a terminológiu. Následky zle vyvinutého štatistického prístupu môžu nielen znemožniť získanie grantových prostriedkov na výskum, ale môžu najmä poškodiť vedeckú štúdiu pre neschopnosť adekvátne testovať navrhované hypotézy. Krátka a oneskorená konzultácia so štatistikom nestačí na vyriešenie všetkých problémov.

Bioštatistik. Bioštatistik musí počas stretnutí s (aplikovaným) výskumníkom koncepčne pochopiť navrhovanú biologickú, antropologickú alebo lekársku metodológiu a terminológiu. (Aplikovaný) výskumník nesmie chápať bioštatistika a jeho podiel na výskume ako servis, ale ako rovnocennú spoluprácu dvoch (alebo viacerých) vzájomne sa doplňujúcich odborov. Tak ako antropológia nie je servisom archeológie, ani štatistika nie je servisom antropológie. Rovnaký princíp platí aj pre spoluprácu bioštatistika s biológom či lekárom.

Vedecká etika. Kvalitný plán štúdie definuje vedecké výhody v porovnaní s rizikami, ktoré na seba berú zaangažované subjekty. Ak je plán štúdie chybný, alebo sa dáta nesprávne analyzovali, potom štúdia nemôže byť odobrená po etickej stránke ani schválená inštitucionálnou vedeckou komisiou.

Medzi základné úlohy (otázky) bioštatistika pri vývoji štatistického plánu patrí:

- Ujasnenie vedeckých otázok. Sú primárne hypotézy jednoznačne formulované, primerané a realistické?
- Identifikácia závislých premenných priamo súvisiacich s vedeckou otázkou/otázkami. Sú základné alebo vedľajšie závislé premenné jednoznačne definované?
- Podáva plán vedeckej štúdie primeraný obraz o navrhovaných hypotézach?
- Sú otázky ohľadom výchylky, zaslepenia alebo stratifikácie primerane ošetrené a jednoznačne opísané?
- Obsahuje plán vedeckej štúdie jasne špecifikovaný a vhodný plán štatistickej analýzy?
- Sú dáta zozbierané na základe opísaných kritérií, obsahuje štúdia bezpečný plán sledovania alebo pravidlá ukončenia štúdie?
- Je zaistená interná a externá validita štúdie?

Otázky, ktoré zaujímajú bioštatistika:

- Aká je vedecká hypotéza?
- O aký typ plánu štúdie ide?
- Ktorá závislá premenná je najdôležitejšia?
- O aký typ premennej ide a v akých jednotkách sa merala?
- Aký je klinicky/biologicky/antropologicky dôležitý efekt pre primárnu závislú premennú?
- Koľko subjektov možno získať alebo pozorovať v rámci výskumu?
- Koľko skupín alebo typov vyšetrení a pod. bude zahrnutých v pláne štúdie?
- Koľko je sledovaných prediktorov? V akých jednotkách sa merajú?
- Bude každá skupina obsahovať rovnaký počet pozorovaní, t.j. aký je alokačný pomer?
- Ak ide o opakované merania, koľko ich je a aký je časový interval medzi meraniami?

2 Model rozdelenia pravdepodobnosti a štatistický model

Pri riešení biomedicínskych problémov pomocou matematickej štatistiky sa vyskytuje stochastický (náhodný) element, ktorým sa nie je možné zaoberať len pomocou základných zákonov aritmetiky. Preto štatistické metódy potrebujú stochastické modely. Vývoj takýchto modelov tvorí *deduktívny*[1] (matematický) aspekt štatistiky. Štatistické problémy sú však *induktívne*[2], pretože vznikajú ako dôsledok pozorovania istých javov v reálnych biomedicínskych situáciách. Tieto javy sú výsledkom nejakého experimentu alebo pozorovania. Otázky, ktoré experiment alebo pozorovanie rieši, sú všeobecnejšieho typu. Pýtajú sa na niečo, čo nie je priamo pozorovanteľné, ale je logicky obsiahnuté v dátach. Hovoríme, že usudzujeme („inferujeme“) niečo na základe dát. Na riešenie deduktívnych problémov matematiky často postačuje čiastočne dostupná informácia, aby sme boli schopní vytvoriť novú matematickú vetu. Na riešenie induktívnych problémov štatistiky potrebujeme všetky dostupné dáta. Len tak môžeme vyvodiť závery. Ignorovanie nejakej časti dát nie je akceptovateľné. Pri deduktívnych problémoch je kvalita novej matematickej vety rovnaká ako kvalita jej predchádzajúcich axióm, definícií alebo viet. Pri induktívnych problémoch je stupeň istoty v záveroch väčší ako v samotných dátach. Čím máme viac dát, tým sa kvalita výstupov zväčšuje. Avšak jedno nové pozorovanie môže naše závery zmeniť.

Dáta môžeme jednoducho popísať aj pomocou **charakteristík (parametrov) polohy a variability** a zobraziť ich pomocou **štatistickej grafiky** (pozri kap. Charakteristiky polohy a variability a štatistická grafika), čo je súčastou tzv. **exploratórnej analýzy dát** (EDA). Avšak použitím nejakého modelu sa o dátach dozvieme viac v zjednodušenej podobe, tento model nám navyše umožní interpretáciu výsledkov a uľahčí nám aj komunikáciu medzi dátami a biomedicínskou praxou.

Koncepcia modelu rozdelenia pravdepodobnosti a štatistického modelu predstavuje jeden zo základných piliérov bioštatistiky. Tieto modely sú charakterizované ich parametrami v prípade parametrického modelu, ktorým sa budeme podrobnejšie zaoberať. Jeho parametre slúžia na jednoduchú charakterizáciu dát a zjednodušujú interpretáciu výsledkov. Modely rozlišujeme podľa toho, či sú to modely na diskrétne alebo spojité dáta. **Model rozdelenia pravdepodobnosti** je charakterizovaný funkciou hustoty a distribučnou funkciou na základe presne špecifikovaných parametrov, ktoré je potrebné odhadnúť z dát pomocou funkcie vierohodnosti – ide teda o „model na dáta“[3]. Dáta sú realizáciami nahodnej premennej, o ktorej predpokladáme, že má asymptoticky (pre veľké n) nejaké rozdelenie, napr. normálne, binomické, multinomické, súčinové multinomické alebo Poissonove. Tento asymptotický predpoklad je fundamentálnym základom štatistickej inferencie, ktorá je procesom vyvodenia záverov (na základe dát a z nich vypočítaných odhadov parametrov) prostredníctvom postupu testovania hypotéz, ktorý používa tzv. štatistiky, testovacie štatistiky a ich asymptotické rozdelenia pravdepodobnosti (pozri kap. Testovanie hypotéz). Hypotézy testujeme aj v **štatistických modeloch**, ktoré často predstavujú modely kauzálnej závislosti zavislých premenných na prediktoroch. V týchto modeloch pomocou funkcie vierohodnosti odhadujeme parametre, ktoré zjednodušujú interpretácie výsledkov, či už štatisticky nesignifikantných alebo signifikantných, ale aj biologicky (medicínsky) nevýznamných alebo významných.

Všetky vyššie uvedené postupy sú súčasťou širšieho pojmu **(bio)štatistická analýza**. Bez správnej formulácie a aplikácie modelu rozdelenia pravdepodobnosti a štatistického modelu na dáta by štatistická analýza nebola možná a závery z nej by boli problematické.

Základom každej empirickej štúdie, či už experimentálnej alebo observačnej, je zabezpečiť **dáta** (**dátový súbor**, **realizácie**), ktoré označíme $\mathbf{x}$, procesom, ktorý nazývame **experiment** (**pokus**) alebo **meranie**. V najjednoduchšej podobe $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ (vektor pozorovaní, vektor výberových hodnôt, vektor realizácií), kde n je **rozsah (náhodného) výberu** a x_i sú realizácie vždy označované malými písmenami (všeobecne ozn. x). (Jednorozmerná) **náhodná premenná** X je funkcia z výberového priestoru $\mathcal{Y}$ (označovaná aj ako priestor elementárnych udalostí alebo množina

[1] Dedukcia – na základe všeobecne platných záverov hľadáme riešenia nejakého konkrétneho prípadu.
[2] Indukcia – na základe konkrétnych prípadov vyvodzujeme všeobecnejšie platné závery.
[3] Model rozdelenia pravdepodobnosti je možné použiť aj v súvislosti so štatistikou – tu ide o „model pre štatistiku“ a s testovacou štatistikou – tu ide o „model pre testovaciu štatistiku“, ale aj v súvislosti s chybami štatistického modelu – tu ide o „model pre chyby (reziduály)“.

výsledkov náhodného pokusu) do množiny reálnych čísel $\mathbb{R}$. Pozorovanie x je realizáciou náhodnej premennej X. Analogicky môžeme definovať **k-rozmerný náhodný vektor** $(X_1, X_2, \ldots, X_k)^T$. Príkladom takéhoto vektora je dvojrozmerný náhodný vektor $(X_{1i}, X_{2i})^T$, $i = 1, 2, \ldots, n$, s realizáciami $(x_{1i}, x_{2i})^T$ usporiadanými po riadkoch do matice $n \times 2$, ktorá je v tomto prípade dátovým súborom[4].

Príklad 46 (príklady náhodných premenných) *(1) Chirurg vykoná 100 transplantácií srdca, kde náhodná premenná X je napr. počet úspešne vykonaných transplantácií. (2) 50 bežcov beží maratón, kde náhodná premenná X je napr. čas odbehnutia maratónu v hodinách. (3) Hodíme 30-krát kockou, kde náhodná premenná X je napr. počet hodených šestiek. (4) Na 75 deťoch vo veku 10 rokov zmeriame výšku (v metroch) a hmotnosť (v kilogramoch), kde náhodná premenná X je napr. Rohrerov index* $(RI = \frac{\textit{hmotnosť v kg}}{(\textit{výška v m})^3})$.

Príklad 47 (porovnanie dvoch typov modelov) *Model rozdelenia pravdepodobnosti je modelom náhodnej premennej X, napr. model rozdelenia pravdepodobnosti náhodnej premennej X šírka dolnej čeľuste alebo (2) model rozdelenia pravdepodobnosti náhodnej premennej X hrúbka kožných rias u dospelých zdravých žien. Štatistický model je modelom náhodnej premennej* $Y|X$ *(Y kauzálne zavisí na X), napr. (1) model závislosti náhodnej premennej Y šírka dolnej čeľuste na premennej X pohlavie alebo (2) model závislosti náhodnej premennej Y hrúbka kožných rias u dospelých zdravých žien na premennej X BMI. Všimnime si, že náhodné premenné označujeme X alebo Y podľa toho, aký model ich charakterizuje.*

Základný predpoklad je, že nejaké pozorovanie (výberová hodnota, realizácia) x je hodnota (realizácia) **náhodnej premennej** X a naším cieľom je použiť x na vyvodenie záverov o neznámom modeli (rozdelenia pravdepodobnosti alebo štatistickom) $F_*(\cdot)$ premennej X. Naše závery o $F_*(\cdot)$ sú zaťažené neistotou kvôli náhodnosti X, z ktorej pochádzajú x. Cieľom je zabezpečiť,

1. aby stupeň neistoty bol čo možno najmenší, berúc do úvahy náhodnosť X,
2. a aby bol tento stupeň neistoty vyjadrený v našich záveroch.

Ekvivalentne $(x_1, x_2, \ldots, x_k)^T$ je realizácia náhodného vektora $(X_1, X_2, \ldots, X_k)^T$ a naším cieľom je použiť túto realizáciu na vyvodenie záverov o neznámom k-rozmernom modeli $F_*^{(k)}(\cdot)$ náhodného vektora $(X_1, X_2, \ldots, X_k)^T$.

Definícia 14 (štatistická inferencia) *Štatistická inferencia (zriedkavo nazývaná aj štatistická indukcia) je proces vyvodenia záverov na základe dát prostredníctvom testovania hypotéz, modelu rozdelenia pravdepodobnosti alebo štatistického modelu (Cox, 2006). Tento proces je ovplyvnený náhodnými chybami, náhodným výberom, voľbou testovacieho kritéria, štatistického modelu alebo modelu rozdelenia pravdepodobnosti. Výsledkom tohto procesu sú zmysluplné závery aplikované na dobre definované dostatočne všeobecné situácie.*

Podstata vytvárania realizácií x limituje možnosti voľby modelu F_*. Inferencia bude o to presnejšia, o čo lepšie bude vybraná čo najmenšia množina $\mathcal{F}$ tak, aby $F_* \in \mathcal{F}$, kde $\mathcal{F}$ nazývame **množinou modelov** (štatistických modelov a modelov rozdelenia pravdepodobnosti). V mnohých prípadoch predpokladáme, že $X_i, i = 1, 2, \ldots, n$, sú **rovnako rozdelené náhodné premenné** (ozn. iid; *independently identically distributed*). V tomto prípade hovoríme o **jednoduchom náhodnom výbere** (srs; *simple random sample*) s rozsahom n, kde X je charakterizované rozdelením $F_*(\cdot)$.

Príklad 48 (jednoduchý náhodný výber) *V jednoduchom náhodnom výbere s rozsahom n z populácie s konečným rozsahom N má každý prvok rovnakú pravdepodobnosť vybratia. Ak vyberáme bez vrátenia, hovoríme o* ***jednoduchom náhodnom výbere bez vrátenia*** *(Dalgaard, 2008). Ak vyberáme s vrátením, hovoríme o* ***jednoduchom náhodnom výbere s vrátením****. Majme množinu* $\mathcal{M}$ *s* $N = 10$

[4]Pre dvojrozmerný náhodný vektor sa často používa ozn. $(X, Y)^T$ a pre jeho realizáciu ozn. $(x, y)^T$.

prvkami a chceme z nej vybrať $n = 3$ prvkov (a) bez vrátenia a (b) s vrátením. Koľko máme možností? Ako vyzerá jedna takáto možnosť, ak ide o množinu $\mathcal{M} = \{1, 2, \ldots, 10\}$. Zopakujte to isté pre $N = 100$, $n = 30$ a množinu $\mathcal{M} = \{1, 2, \ldots, 100\}$.

Riešenie aj v R[5]

(a) Spolu máme $\binom{N}{n}$ možných náhodných výberov (kombinácie bez opakovania n-tej triedy z N prvkov množiny $\mathcal{M}$). Ak $N = 10$ a $n = 3$, potom kombinačné číslo $\binom{N}{n} = \frac{N!}{(N-n)!n!} = \binom{10}{3} = 120$ možností. Ak $N = 100$ a $n = 30$, potom $\binom{N}{n} = \binom{100}{30} = 2.937234 \times 10^{25}$ možností.

```
choose(10,3) # pocet vsetkych moznych vyberov bez vratenia
choose(100,30)
library(utils)
combn(10,3) # pocet vsetkych moznych vyberov bez vratenia
combn(100,30)
sample(x=1:10,size=3,replace=FALSE) # jednoduchy nahodny vyber bez vratenia
sample(x=1:100,size=30,replace=FALSE)
```

(b) Spolu máme $\binom{N+n-1}{n}$ možných náhodných výberov (kombinácie s opakovaním n-tej triedy z N prvkov množiny $\mathcal{M}$). Ak $N = 10$ a $n = 3$, potom $\binom{N+n-1}{n} = \frac{(N+n-1)!}{(N-1)!n!} = \binom{10+3-1}{3} = 220$ možností. Ak $N = 100$ a $n = 30$, potom $\binom{N+n-1}{n} = \binom{100+30-1}{30} = 2.009491 \times 10^{29}$ možností.

```
choose(10+3-1,3) # pocet vsetkych moznych vyberov s vratenim
choose(100+30-1,30)
library(utils)
combn(10+3-1,3) # pocet vsetkych moznych vyberov s vratenim
combn(100+30-1,30)
sample(x=1:10,size=3,replace=TRUE) # jednoduchy nahodny vyber s vratenim
sample(x=1:100,size=30,replace=TRUE)
```

Príklad 49 (jednoduchý náhodný výber) *Nech je skupina ľudí označená identifikačnými číslami (ID) od 1 do 30. Vyberte (a) náhodne 5 ľudí z 30 bez návratu, (b) náhodne 5 ľudí z 30 s návratom a nakoniec (c) náhodne 5 ľudí z 30 bez návratu, kde ľudia s ID od 28 do 30 majú pravdepodobnosť vybratia 4× väčšiu ako ľudia s ID od 1 do 27.*

Riešenie v R

```
sample(x=1:30,size=5,replace=FALSE)
sample(x=1:30,size=5,replace=TRUE)
sample(x=1:30,size=5,prob=c(rep(1/39,27),rep(4/39,3)),replace=FALSE)
```

Distribučná funkcia F náhodnej premennej X je definovaná ako $F_X(x) = \Pr(X < x)$, kde zápis znamená pravdepodobnosť, že náhodná premenná X nadobúda hodnoty menšie alebo rovné ako nejaké číslo x, ktoré nazývame *kvantil*. Dolný index v F_X sa spravidla vynecháva a píšeme F.

Náhodná premenná X sa nazýva **diskrétna**, ak jej distribučná fukcia F je schodovitá funkcia. V tomto prípade hovoríme, že ide o **diskrétne rozdelenie** a množina $\mathcal{Y}$, z ktorej pochádzajú X je konečná alebo nanajvýš spočítateľná (má spočítateľne veľa prvkov).

Distribučná funkcia diskrétnej náhodnej premennej X je definovaná ako (Azzalini, 1996)

$$F_X(x) = \Pr(X \leq x) = \sum_{i:x_i \leq x} \Pr(X = x_i),$$

kde $\sum_{i=1}^{k(\infty)} p_i = 1$, $\Pr(X = x_i) = p_i = f_X(x_i) = f(x_i), \forall x_i$, sa nazýva **pravdepodobnostná funkcia**. Často zapisujeme $\{x_i, p_i\}_{i=1}^{k(\infty)}$, kde x_i sú realizácie, p_i sú pravdepodobnosti výskytu x_i, $k \in \mathbb{N}^+$, kde $\mathbb{N}^+$ znamená množinu kladných prirodzených čísel[6].

[5] Detaily o jazyku R pozri napr. v Chambers (2008), Becker a kol. (1988) alebo Matloff (2011).

[6] Písmeno $\mathbb{N}$ sa v angličtine nazýva *blackboard bold* N (*double-struck capital* N).

Príklad 50 (diskrétna premenná) *(1) počet úspešne vykonaných transplantácií v SR, (2) počet hodených šestiek, (3) štyri vekové kategórie, do ktorých zaradíme subjekty, (4) pohlavie, (5) počet starších súrodencov (o.sib.N; dve kategórie; dáta: two-samples-means-birth.txt), (6) vzdelanie matky (edu.M, štyri kategórie; dáta: anova-newborns.txt).*

Náhodná premenná X sa nazýva **spojitá**, ak jej distribučná fukcia F je absolútne spojitá funkcia. V tomto prípade hovoríme, že ide o **spojité rozdelenie** a množina $\mathcal{Y}$ je nekonečná (má nekonečne veľa prvkov).

Distribučná funkcia spojitej náhodnej premennej X je definovaná ako

$$F_X(x) = \int_{-\infty}^{x} f(t)\, dt, f(x) \geq 0,$$

kde $\int_{-\infty}^{\infty} f(x)\, dx = 1$, $f_X(x) = f(x) = \frac{\partial}{\partial x} F_X(x)$ sa nazýva **hustota**.

Príklad 51 (spojitá premenná) *(1) čas odbehnutia maratónu v hodinách; (2) hmotnosť v kilogramoch; (3) výška v centimetroch; (4) Rohrerov index; (5) dĺžkošírkový index lebky vypočítaný ako podiel náhodných premenných najväčšia šírka mozgovne a najväčšia dĺžka mozgovne (skull.B a skull.L; v mm; dáta: one-sample-mean-skull-mf.txt); (6) stranový rozdiel vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej a ľavej strane tela (length.R a length.L; v mm; dáta: paired-means-clavicle2.txt); (7) najväčšia výška mozgovne a morfologická výška tváre (skull.pH a face.H; v mm; dáta: one-sample-correlation-skull-mf.txt).*

Rozmery na živom objekte nemožno merať s absolútnou presnosťou a počet desatinných miest závisí na presnosti merania; ako posledné desatinné miesto by sa malo uvádzať to, na ktorom sa pri opakovanom meraní rovnakého objektu na viac desatinných miest zhodujú všetky merania.

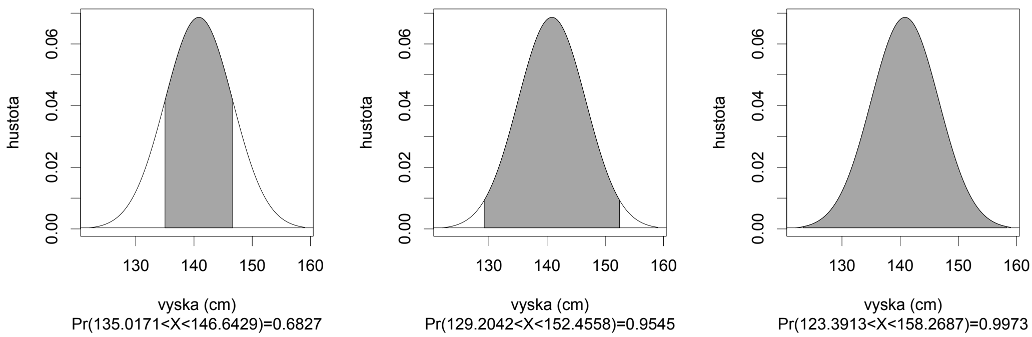
Príklad 52 (normálne rozdelenie) *Majme náhodnú premennú X (môže to byť napr. výška postavy 10-ročných dievčat) a predpokladáme, že má normálne rozdelenie s parametrami μ (stredná hodnota) a σ^2 (rozptyl), čo zapisujeme ako $X \sim N(\mu, \sigma^2)$, $\mu = 140.83, \sigma^2 = 33.79$. Normálne rozdelenie predstavuje model rozdelenia pravdepodobnosti pre túto náhodnú premennú. Vypočítajte pravdepodobnosť $\Pr(a \leq X < b) = \Pr(X < b) - \Pr(X < a) = F_X(b) - F_X(a)$, kde $a = \mu - k\sigma$, $b = \mu + k\sigma$, $k = 1, 2, 3$.*

Riešenie (aj v R); (pozri obrázok 2.1)
$a = \mu - \sigma = 135.0171$, $b = \mu + \sigma = 146.6429$,
$\Pr(|X - \mu| > \sigma) = 0.3173, \Pr(|X - \mu| < \sigma) = 1 - 0.3173 = 0.6827$,
$a = \mu - 2\sigma = 129.2042$, $b = \mu + 2\sigma = 152.4558$,
$\Pr(|X - \mu| > 2\sigma) = 0.0455, \Pr(|X - \mu| < 2\sigma) = 1 - 0.0455 = 0.9545$,
$a = \mu - 3\sigma = 123.3913$, $b = \mu + 3\sigma = 158.2687$,
$\Pr(|X - \mu| > 3\sigma) = 0.0027, \Pr(|X - \mu| < 3\sigma) = 1 - 0.0027 = 0.9973$.
Pozn.: Pravdepodobnosť $\Pr(a < X < b) = \Pr(a \leq X \leq b)$, pretože pravdepobnosť v bode (tu a a b) je rovná nule pre spojité premenné, t.j. $\Pr(a) = \Pr(b) = 0$. Pre diskrétne premenné to neplatí.

Alternatívny výpočet cez štandardizované normálne rozdelenie (syn. normálne normované rozdelenie) je nasledovný:

```
mu <- 0
sig <- 1
bin <- seq(mu-3*sig,mu+3*sig,by=sig)
pnorm(bin[7])-pnorm(bin[1]) # 0.9973002
pnorm(bin[6])-pnorm(bin[2]) # 0.9544997
pnorm(bin[5])-pnorm(bin[3]) # 0.6826895
```

Dostaneme pravidlo 68.27 – 95.45 – 99.73 (tzv. „**miery normálneho rozdelenia**“).



Obr. 2.1: Miery normálneho rozdelenia; krivka hustoty s vyfarbeným obsahom pod touto krivkou medzi príslušnými kvantilmi na osi x; obsah je rovný pravdepodobnosti výskytu subjektov s danou výškou v rozpätí týchto kvantilov

Príklad 53 (normálne rozdelenie) *Majme* $X \sim N(\mu, \sigma^2)$, *kde* $\mu = 150, \sigma^2 = 6.25$. *Vypočítajte* $a = \mu - x_{1-\alpha/2}\sigma$ *a* $b = \mu + x_{1-\alpha/2}\sigma$ *tak, aby* $\Pr(a \le X \le b) = 1 - \alpha$, *bola rovná* 0.90, 0.95 *a* 0.99. *Číslo* $x_{1-\alpha}$ *je kvantil normálneho normovaného rozdelenia, t.j.* $\Pr(Z = \frac{X-\mu}{\sigma} < x_{1-\alpha}) = 1 - \alpha, Z \sim N(0,1)$.

Riešenie (aj v R); (pozri obrázok 2.2)
$\Pr\left(\mu - x_{1-\alpha/2}\sigma < X < \mu + x_{1-\alpha/2}\sigma\right) = \Pr\left(X < \mu + x_{1-\alpha}\sigma\right) - \Pr\left(X < \mu - x_{1-\alpha}\sigma\right) = 1 - \alpha = 0.9$. Z-transformáciou[7] na normálne normované rozdelenie dostaneme $\Pr(-x_{1-\alpha/2} < Z < x_{1-\alpha/2}) = 0.9$, kde $\frac{\mu - x_{1-\alpha/2}\sigma - \mu}{\sigma} = -x_{1-\alpha/2}$, $\frac{\mu + x_{1-\alpha/2}\sigma - \mu}{\sigma} = x_{1-\alpha/2}$, $x_{1-\alpha} = x_{0.95} = 1.64$, t.j. 90.00 % dát leží v intervale $\mu \pm 1.64\sigma$.
$\Pr(a < X < b) = 0.95$. Potom $x_{0.975} = 1.96$, t.j. 95.00 % dát leží v intervale $\mu \pm 1.96\sigma$.
$\Pr(a < X < b) = 0.99$. Potom $x_{0.995} = 2.58$, t.j. 99.00 % dát leží v intervale $\mu \pm 2.58\sigma$.

```
Q95 <- qnorm(0.95,0,1) # 1.644854
Q05 <- qnorm(0.05,0,1) # -1.644854
Q975 <- qnorm(0.975,0,1) # 1.959964
Q025 <- qnorm(0.025,0,1) # -1.959964
Q995 <- qnorm(0.995,0,1) # 2.575829
Q005 <- qnorm(0.005,0,1) # -2.575829
```

Dostaneme pravidlo 90 − 95 − 99 (tzv. „**upravené miery normálneho rozdelenia**“). Použili sme nerovnosť $\Pr\left(x_{\alpha/2} < Z < x_{1-\alpha/2}\right) = \Phi\left(x_{1-\alpha/2}\right) - \Phi\left(x_{\alpha/2}\right) = 1 - \alpha$, kde Φ je distribučná funkcia normálneho normovaného rozdelenia a všeobecne $\alpha \in (0, 1/2)$; v príklade $\alpha = 0.1$, 0.05 a 0.01.

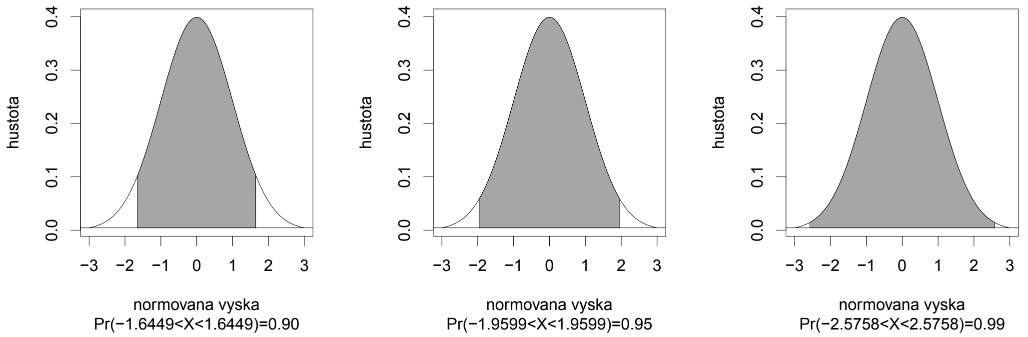
Príklad 54 (normálne rozdelenie) *Predpokladajme model normálneho rozdelenia* $N\left(132, 13^2\right)$ *pre systolický krvný tlak. Aká časť populácie (v %) bude mať hodnoty väčšie ako* 160 *mm Hg?*

Riešenie (aj v R)
Pomocou Z-transformácie dostaneme
$\Pr(X > 160) = \Pr\left(\frac{X-132}{13} > \frac{160-132}{13}\right) = \Pr\left(\frac{X-132}{13} > 2.154\right) = 0.016$.

```
(1-pnorm(160,mean=132,sd=13))*100 # 1.562612 %
z.transf <- (160-132)/13
(1-pnorm(z.transf))*100 # 1.562612 %
```

Teda asi 1.6 % populácie z $N\left(132, 13^2\right)$ bude mať systolický krvný tlak väčší ako 160 mm Hg.

[7] Z-transformácia je spôsob transformácie náhodnej premennej $X \sim N(\mu, \sigma^2)$ pomocou centrovania strednou hodnotou μ a normovania smerodajnou odchýlkou σ, kde $Z = \frac{X-\mu}{\sigma}$; $Z \sim N(0,1)$.



Obr. 2.2: Upravené miery normálneho rozdelenia; krivka hustoty s vyfarbeným obsahom pod touto krivkou medzi príslušnými kvantilmi na osi x; obsah je rovný pravdepodobnosti výskytu subjektov s danou normovanou výškou v rozpätí týchto kvantilov

Príklad 55 (binomické rozdelenie) *Predpokladajme, že počet ľudí uprednostňujúcich liečbu A pred liečbou B sa správa podľa modelu binomického rozdelenia s parametrami p (pravdepodobnosť výskytu udalosti) a N (rozsah náhodného výberu), ozn. Bin* (N, p), *kde* $N = 20, p = 0.5$, *t.j. ľudia preferujú oba typy liečby rovnako. (a) Aká je pravdepodobnosť, že bude* 16 *a viac pacientov uprednostňovať liečbu A pred liečbou B? (b) Aká je pravdepodonosť, že bude* 16 *a viac a zároveň* 4 *alebo menej pacientov uprednostňovať liečbu A pred liečbou B?*

Riešenie (aj v R)
(a) $\Pr(X \geq 16) = 1 - \sum_{i:x_i \leq 15} \Pr(X = x_i) = 1 - \sum_{i:x_i \leq 15} \binom{N}{x_i} p^{x_i} (1-p)^{N-x_i} = 1 - \sum_{i:x_i \leq 15} \binom{20}{x_i} 0.5^{x_i} (1-0.5)^{20-x_i} = 0.006$.

```
pbinom(16,size=20,prob=0.5) # 0.9987116
1-pbinom(16,size=20,prob=0.5) # 0.001288414
```

Z vyššie uvedeného R-kódu vyplýva, že ide o pravdepodobnosť $\Pr(X \leq 16)$ a $\Pr(X > 16)$, ale my potrebujeme $\Pr(X \geq 16)$. Preto R-kód upravíme nasledovne

```
1-pbinom(15,size=20,prob=0.5) # 0.005908966
sum(choose(20,16:20)*0.5^(16:20)*0.5^(20-16:20)) # 0.005908966
```

(b) $\Pr(X \leq 4, X \geq 16) = 1 - \sum_{i:x_i \leq 15} \Pr(X = x_i) + \sum_{i:x_i \leq 4} \Pr(X = x_i) = 0.012$. Táto pravdepodobnosť je dvojnásobkom predchádzajúcej pravdepodobnosťi, lebo $Bin(N, 0.5)$ je symetrické okolo 0.5, t.j.

```
1-pbinom(15,size=20,prob=0.5)+pbinom(4,size=20,prob=0.5) # 0.01181793
```

Príklad 56 (binomické rozdelenie) *Predpokladajme, že pohlavie novorodencov (mužské alebo ženské) sa správa podľa modelu binomického rozdelenia s parametrami p (pravdepodobnosť výskytu chlapcov) a N (rozsah náhodného výberu), ozn.* $Bin(N, p)$, *kde* $N = 1113$, $p = 0.52$, *t.j. rodí sa o niečo viac chlapcov než dievčat (dáta:* two-samples-probabilities-sexratio.txt*). Aká je pravdepodobnosť, že sa narodí* 700 *a viac dievčat?*

Príklad 57 (binomické rozdelenie) *Predpokladajme, že* $\Pr(vír) = 0.533 = p_1$ *je pravdepodobnosť výskytu dermatoglyfického vzoru vír na palci pravej ruky mužov českej populácie a* $\Pr(ostatné) = 0.467 = p_2$ *je pravdepodobnosť výskytu ostatných vzorov na palci pravej ruky mužov českej populácie,* *pričom X je počet vírov a Y je počet ostatných vzorov, kde $X \sim Bin(N, p_1)$ a $Y \sim Bin(N, p_2)$. Vypočítajte (1) $\Pr(X \leq 120)$, keď $N = 300$ a (2) $\Pr(Y \leq 120)$, keď $N = 300$.*

$\mathcal{F}$ môže byť nejaká množina **distribučných funkcií** identifikovateľných pomocou parametra $\boldsymbol{\theta}$ z parametrického priestoru $\boldsymbol{\Theta} \in \mathbb{R}^k$ ($\mathbb{R}$ znamená množinu reálnych čísel[8]), čo môžeme formálne zapísať ako (Azzalini, 1996)

$$\mathcal{F} = \left\{ F(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^k \right\},$$

kde pre každé fixné $\boldsymbol{\theta}$ je $F(\cdot; \boldsymbol{\theta})$ distribučnou funkciou, ktorej nosič je $\mathcal{Y}_{\boldsymbol{\theta}} \subseteq \mathbb{R}^n$. **Nosič** $\mathcal{Y}_{\boldsymbol{\theta}}$ je najmenšia množina, na ktorej je hustota definovaná. Výberový priestor je množina $\mathcal{Y}$ všetkých možných hodnôt x, ktoré charakterizujeme modelom. Formálne $\mathcal{Y} = \cup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \mathcal{Y}_{\boldsymbol{\theta}}$. Často však $\mathcal{Y}_{\boldsymbol{\theta}}$ je rovnaký pre všetky $\boldsymbol{\theta}$, a preto koinciduje s $\mathcal{Y}$.

Príklad 58 (parametre) *Príklady parametrov θ – stredná hodnota μ, rozptyl σ^2, korelačný koeficient ρ, pravdepodobnosť p výskytu nejakej udalosti, rozdiel dvoch stredných hodnôt $\mu_1 - \mu_2$, podiel dvoch rozptylov σ_1^2/σ_2^2, rozdiel dvoch korelačných koeficientov $\rho_1 - \rho_2$, rozdiel dvoch pravdepodobností $p_1 - p_2$ a pod.*

Príklad 59 (parametre) *Príkladmi parametrov môžu byť: (1) stredná hodnota μ a (2) rozptyl σ^2 dĺžky lebky (skull-L, v mm) egyptskej stredovekej mužskej populácie; (3) rozdiel medzi strednými hodnotami $\mu_1 - \mu_2$ dĺžky lebky mužskej a ženskej; (4) podiel rozptylov σ_1^2/σ_2^2 hodnôt dĺžky lebky u mužov a u žien (dáta: two-samples-means-skull.txt); (5) korelačný koeficient ρ medzi dĺžkou dolnej končatiny (lowex.L, v mm) a dĺžkou trupu (tru.L, v mm); (6) rozdiel korelačných koeficientov $\rho_1 - \rho_2$ medzi dĺžkou dolnej končatiny a dĺžkou trupu u mužov a u žien (dáta: two-samples-correlations-trunk.txt); (7) pravdepodobnosť p výskytu mužov (sex; m – muž, f – žena); (8) pravdepodobnosť výskytu popôrodných zmien p na ženských panvových kostiach u Afričaniek, ako aj rozdiel pravdepodobností $p_1 - p_2$ výskytu výrazných popôrodných zmien na panvových kostiach Afričaniek a Inuitiek (dáta: more-samples-probabilities-pubis.txt); (9) rozdiel pravdepodobností $p_1 - p_2$ sexuálnej orientácie na opačné pohlavie (sexor – sexuálna orientácia; op – výlučne na opačné pohlavie, sa – minimálne/občas na rovnaké pohlavie) u mužov a žien (dáta: anova-head.txt).*

Čítanie označení. Pojem „model rozdelenia pravdepodobnosti" sa často skracuje na „rozdelenie". Potom hovoríme, že „X má rozdelenie $F_X(x)$", „X je charakterizované rozdelením $F_X(x)$" alebo „X pochádza z rozdelenia $F_X(x)$", čo označujeme ako $X \sim F_X(x)$, kde symbol „$\sim$" čítame ako „je rozdelená ako" alebo „pochádza z rozdelenia" (často sa uvádza aj pojem „**asymptoticky**", čo znamená „pre veľké n"). Mohli by sme písať aj $X \sim f_X(x)$, to sa však používa len zriedkavo. Ak porovnávame rozdelenia dvoch náhodných premenných X a Y, hovoríme „X a Y majú rovnaké rozdelenie" alebo „X a Y sú rovnako rozdelené", ozn. $X \sim Y$ alebo $F_X(x) \sim F_Y(y)$. Pojem „štatistický model" sa často skracuje na „model".

Tri typy priestorov. Výberový priestor $\mathcal{Y}$ súvisí s náhodnou premennou a jej realizáciou, nosič $\mathcal{Y}_{\boldsymbol{\theta}}$ súvisí s hodnotami, na ktorých je definovaná hustota rozdelenia pravdepodobnosti a parametrický priestor $\boldsymbol{\Theta}$ súvisí s parametrom $\boldsymbol{\theta}$.

Príklad 60 (binomické rozdelenie) *Ak $X \sim Bin(N, \theta), \theta = p \in \langle 0, 1 \rangle$, potom $\mathcal{Y}_\theta$ je rovnaký pre všetky θ a koinciduje s výberovým priestorom $\mathcal{Y} = \{0, 1, \ldots, N\}$.*

[8] $\mathbb{R}^k$ znamená k-rozmernú množinu reálnych čísel (k je dĺžka vektora parametrov $\boldsymbol{\theta} \in \boldsymbol{\Theta}$), kde ak $k = 1$, potom $\theta \in \Theta$ je číslo (skalár). $\mathbb{R}^n$ znamená n-rozmernú množinu reálnych čísel (n je rozsah náhodného výberu). $\mathbb{R}^+$ znamená množinu kladných reálnych čísel.

Aproximácia binomického rozdelenia normálnym. Ak $X \sim Bin(N, p)$, $Np > 5$ a $Nq > 5$, kde $q = 1 - p$, potom rozdelenie náhodnej premennej X môžeme aproximovať normálnym rozdelením, kde $X \sim N(Np, Npq)$; príklady pozri v tabuľke 2.1.

Tabuľka 2.1: Príklady minimálnych N pre fixované p potrebných na aproximáciu

p	0.1	0.2	0.3	0.4	0.5
q	0.9	0.8	0.7	0.6	0.5
N	51	26	17	13	11

Príklad 61 (aproximácia binomického rozdelenia normálnym) *Nech* $\Pr(\textit{muž}) = 0.515$ *znamená pravdepodobnosť výskytu mužov v populácii a* $\Pr(\textit{žena}) = 0.485$ *pravdepodobnosť výskytu žien. Nech X je počet mužov a Y počet žien. Za predpokladu modelu $Bin(N, p)$ vypočítajte (a)* $\Pr(X \leq 3)$*, ak $N = 5$, (b)* $\Pr(X \leq 5)$*, ak $N = 10$ a (c)* $\Pr(X \leq 25)$*, ak $N = 50$. Porovnajte vypočítané pravdepodobnosti s pravdepodobnosťami aproximovanými normálnym rozdelením $N(Np, Npq)$.*

Riešenie (aj v R) (pozri obrázok 2.3 a 2.4)
Aproximácia znamená „približné vyjadrenie“, t.j. buď nejaké rozdelenie aproximujeme iným (majúcim isté výhody oproti tomu, ktoré aproximujeme), alebo aproximujeme dáta nejakým rozdelením (ktoré popisuje dáta pomocou ľahko interpretovateľných parametrov).
(a) $E[X] = Np = 5 \times 0.515 = 2.575, E[Y] = 5 \times 0.485 = 2.425,$
$\Pr(X \leq 3) = \sum_{k \leq 3} \binom{5}{k} 0.515^k 0.485^{5-k} = 0.793,$
$\Pr(X \leq 3) = 0.648, N(5 \times 0.515, 5 \times 0.515 \times 0.485).$
(b) $E[X] = 10 \times 0.515 = 5.15, E[Y] = 10 \times 0.485 = 4.85,$
$\Pr(X \leq 5) = \sum_{k \leq 5} \binom{10}{k} 0.515^k 0.485^{10-k} = 0.586,$
$\Pr(X \leq 5) = 0.462, N(10 \times 0.515, 10 \times 0.515 \times 0.485).$
(c) $E[X] = 50 \times 0.515 = 25.75, E[Y] = 50 \times 0.485 = 24.25,$
$\Pr(X \leq 25) = \sum_{k \leq 25} \binom{50}{k} 0.515^k 0.485^{50-k} = 0.471,$
$\Pr(X \leq 25) = 0.416, N(50 \times 0.515, 50 \times 0.515 \times 0.485).$

```
pbinom(3,size=5,prob=0.515) # 0.7931878
pnorm(3,mean=5*0.515,sd=sqrt(5*0.515*0.485)) # 0.6481396
pbinom(5,size=10,prob=0.515) # 0.5856244
pnorm(5,mean=10*0.515,sd=sqrt(10*0.515*0.485)) # 0.4621927
pbinom(25,size=50,prob=0.515) # 0.4712842
pnorm(25,mean=50*0.515,sd=sqrt(50*0.515*0.485)) # 0.4159648
```

Z príkladu 61 vyplýva, že pre pravdepodobnosť $p = 0.515$ a $N = 50$ aproximácia stále nie je postačujúca (ani na jedno desatinné miesto) a pre $N = 10$ a $N = 5$ ju nie je možné použiť. Pre pravdepodobnosti p blížiace sa jednotke alebo nule sú potrebné väčšie početnosti ako pre pravdepodobnosti p blízke hodnote 0.5.

Distribučné funkcie patriace do množiny $\mathcal{F}$ sú distribučnými funkciami diskrétnych alebo spojitých náhodných premenných. Potom $\mathcal{F}$ môže byť definovaná ako množina **pravdepodobnostných funkcií** alebo **funkcií hustoty** a model môžeme formálne zapísať ako (Casella a Berger, 2002)

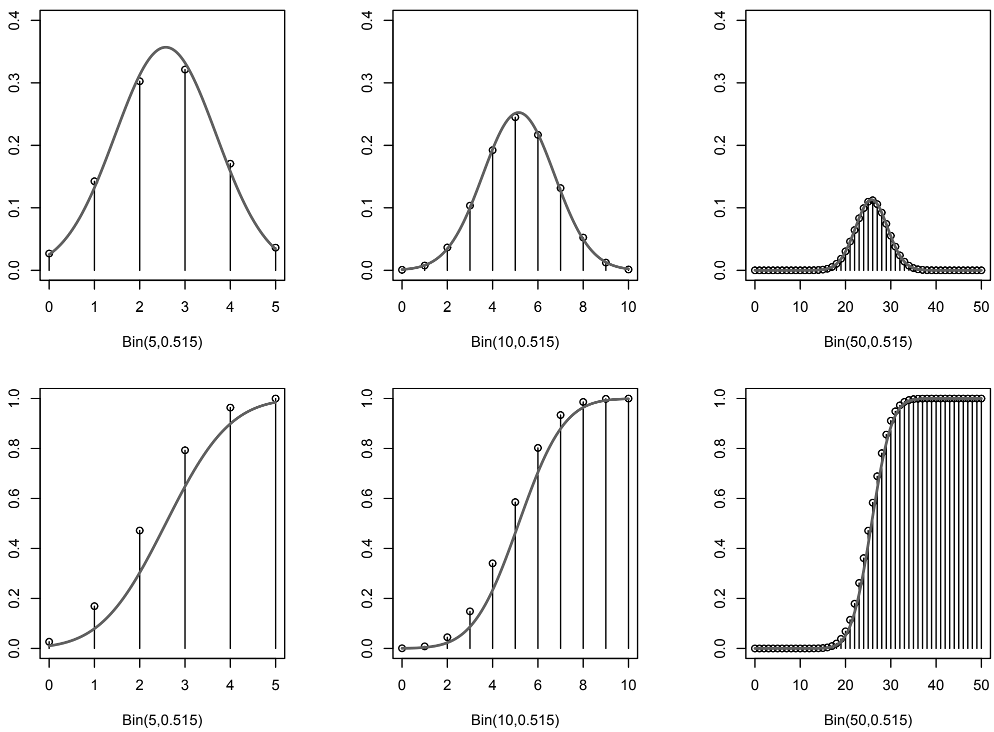
$$\mathcal{F} = \left\{ f(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^k \right\}$$

pre nejakú funkciu hustoty f. Vektor $\boldsymbol{\theta}$ sa nazýva **parameter**, množina $\boldsymbol{\Theta}$ **parametrický priestor** a $\mathcal{F}$ **parametrický model**. Keďže prvky $\mathcal{F}$ sú asociované s prvkami $\boldsymbol{\Theta}$, existuje $\boldsymbol{\theta}_* \in \boldsymbol{\Theta}$ asociovaná s F_* a nazýva sa **skutočná hodnota parametra**. A štatistická inferencia je práve o $\boldsymbol{\theta}_*$. Paramertický štatistický model môže byť charakterizovaný aj pomocou k-rozmerného vektora parametrov $\boldsymbol{\theta}$, preto sa v označení používa $\mathbb{R}^k$, kde pre skalár θ bude $k = 1$.

Ak však

$$\mathcal{F} = \{\text{množina všetkých hustôt funkcií jednej premennej}\},$$

ide o **neparametrický model** (Wasserman, 2006).



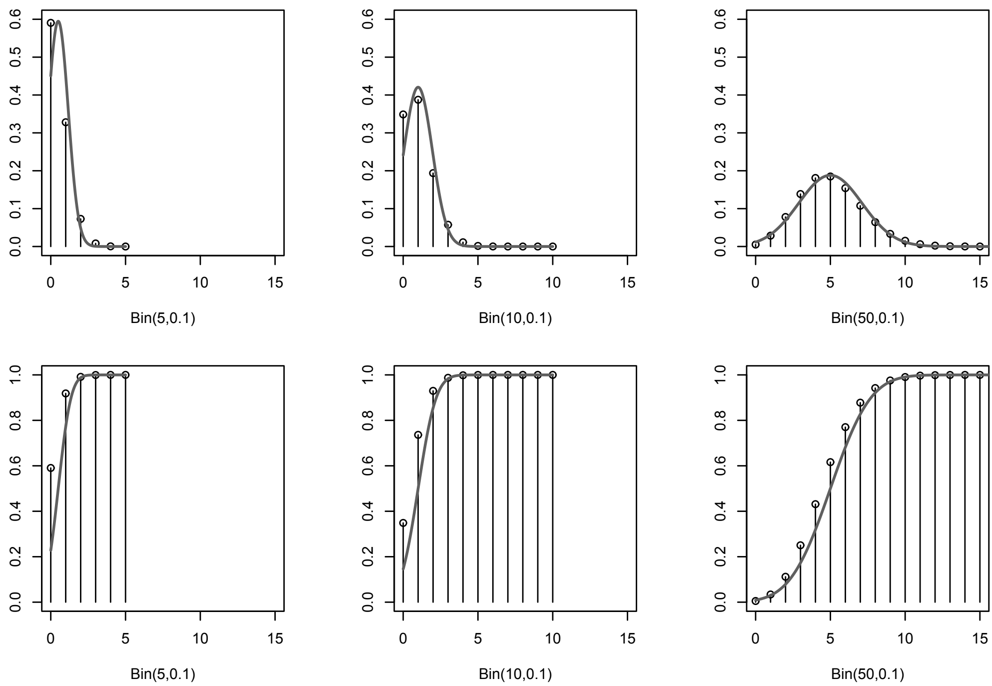
Obr. 2.3: Aproximácia binomického rozdelenia normálnym pre $p = 0.515$ a $N = 5, 10$ a 50; spojnicový graf superponovaný hustotou (prvý riadok) a distribučnou funkciou (druhý riadok)

Normálne rozdelenie. Model pre náhodný výber $X_1, X_2, \ldots, X_n$ je $N(\mu, \sigma^2)$ a hovoríme, že $X_1, X_2, \ldots, X_n$ pochádza z normálneho rozdelenia, t.j. $X \sim N(\mu, \sigma^2)$. Parameter modelu $N(\mu, \sigma^2)$ je vektor $\boldsymbol{\theta} = (\mu, \sigma^2)^T$. Hustota tohto rozdelenia má tvar $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$.

Príklad 62 (normálne rozdelenie) *Predpokladajme, že náhodná premenná X má asymptoticky (pre veľké n) normálne rozdelenie so strednou hodnotou $E[X] = \mu$ a rozptylom $Var[X] = \sigma^2$, čo zapisujeme ako $X \sim N(\mu, \sigma^2)$. Príkladmi takýchto náhodných premenných sú: (1) dĺžka pravej kľúčnej kosti u mužov (length.R; dáta: paired-means-clavicle2.txt); (2) šírka lebky u žien (skull.B; dáta: one-sample-mean-skull-mf.txt).*

Štandardizované normálne rozdelenie. Model pre náhodný výber $X_1, X_2, \ldots, X_n$ je $N(0, 1)$ a hovoríme, že $X_1, X_2, \ldots, X_n$ pochádza zo štandardizovaného normálneho rozdelenia, t.j. $X \sim N(\mu, \sigma^2)$, kde $\mu = 0$ a $\sigma^2 = 1$. Parameter modelu $N(\mu, \sigma^2)$ je vektor $\boldsymbol{\theta} = (0, 1)^T$. Hustota tohto rozdelenia má tvar $\phi(x) = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$.

Príklad 63 (štandardizované normálne rozdelenie) *Predpokladáme, že náhodná premenná X šírka lebky u mužov (skull.B; dáta: one-sample-mean-skull-mf.txt) má asymptoticky normálne rozdelenie so strednou hodnotou μ a rozptylom σ^2, čo zapisujeme ako $X \sim N(\mu, \sigma^2)$. Keď od X odpočítame jej strednú hodnotu μ a tento rozdiel vydelíme odmocninou z rozptylu $\sigma = \sqrt{\sigma^2}$, dostaneme náhodnú*



Obr. 2.4: Aproximácia binomického rozdelenia normálnym pre $p = 0.1$ a $N = 5, 10$ a 50; spojnicový graf superponovaný hustotou (prvý riadok) a distribučnou funkciou (druhý riadok)

premennú Z, ktorá má asymptoticky normálne rozdelenie so strednou hodnotou $\mu = 0$ a rozptylom $\sigma^2 = 1$, čo zapisujeme ako $Z \sim N(0, 1)$.

Dvojrozmerné normálne rozdelenie. Náhodný vektor $(X, Y)^T$ má dvojrozmerné normálne rozdelenie

$$N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \text{ kde } \boldsymbol{\mu} = (\mu_1, \mu_2)^T \text{ a } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

s hustotou (Casella a Berger, 2002)

$$f(x, y) = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left\{ \frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right\}\right\},$$

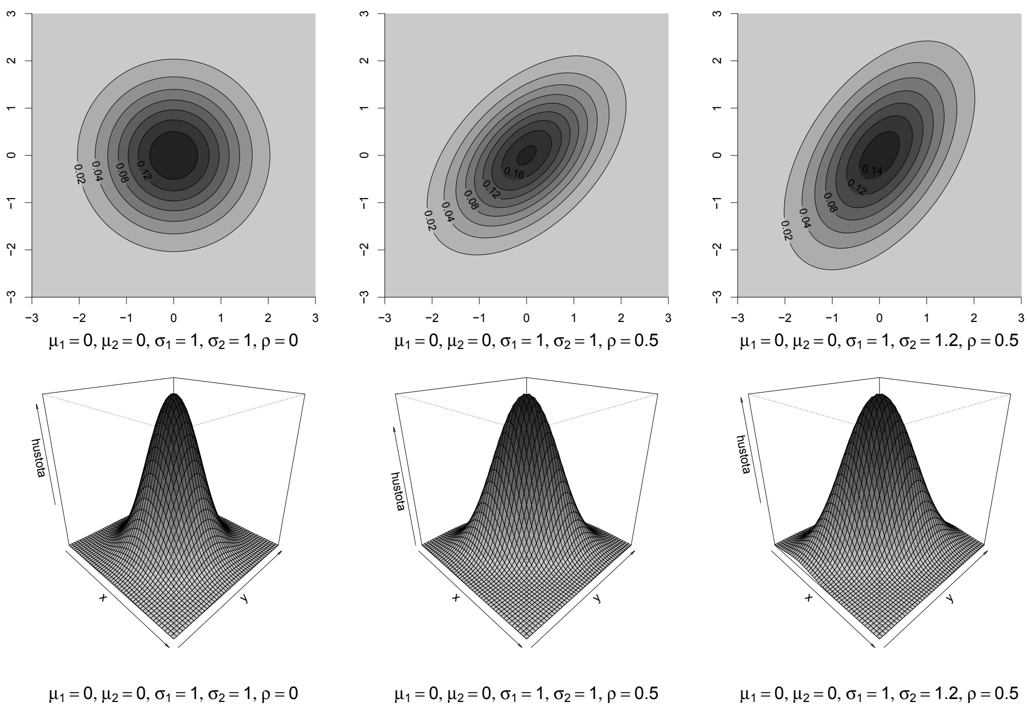
kde $(x, y)^T \in \mathbb{R}^2$, $\mu_i \in \mathbb{R}^1$, $\sigma_i^2 > 0$, $i = 1, 2$, $\rho \in \langle -1, 1\rangle$ sú parametre, potom $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)^T$. Výraz v exponente môžeme písať ako

$$-\frac{1}{2}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}^T \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix},$$

marginálne rozdelenia sú $X \sim N(\mu_1, \sigma_1^2)$ a $Y \sim N(\mu_2, \sigma_2^2)$, ρ je koeficient korelácie (detaily maticovej algebry pozri v Gentle (2007). Margináne rozdelenie je rozdelenie marginálnej náhodnej premennej, tu X nezávisle na Y a naopak Y nezávisle na X.

Z vyššie uvedeného textu je zrejmé, že na dostatočný popis dvojrozmerného normálneho rozdelenia potrebujeme päť parametrov, t.j. strednú hodnotu a rozptyl pre marginálne rozdelenie náhodných premenných X a Y a korelačný koeficient $\rho = \rho(X, Y)$ popisujúci silu lineárneho vzťahu X a Y.

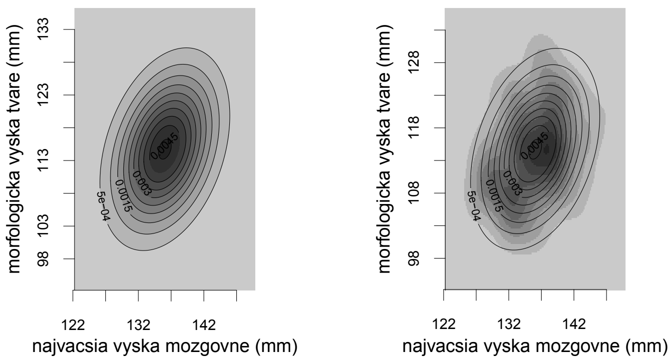
Príklad 64 (dvojrozmerné normálne rozdelenie) *(1) Nakreslite hustotu dvojrozmerného normálneho rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ pomocou funkcie* image() *a superponujte ho s kontúrovým grafom hustoty toho istého rozdelenia pomocou funkcie* contour(). *(2) Nakreslite hustotu dvojrozmerného normálneho rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ pomocou funkcie* persp(). *Hustotu rozsekajte na 12 intervalov, kde hodnoty v týchto intervaloch budú zodpovedať farbám* terrain.colors(12). *Použite nasledovné parametre*
(a) $\mu_1 = 0, \mu_2 = 0,\ \sigma_1 = 1, \sigma_2 = 1,\ \rho = 0$;
(b) $\mu_1 = 0, \mu_2 = 0,\ \sigma_1 = 1, \sigma_2 = 1,\ \rho = 0.5$;
(c) $\mu_1 = 0, \mu_2 = 0,\ \sigma_1 = 1, \sigma_2 = 1.2,\ \rho = 0.5$.
Vzorové riešenie pozri na obrázku 2.5.



Obr. 2.5: Hustoty dvojrozmerného normálneho rozdelenia pri rôznych parametroch (prvý riadok – kontúrový graf, druhý riadok – perspektívny trojrozmerný graf v podobe plochy); čím je ρ odlišnejšie od nuly, tým viac sa kontúry líšia od kruhov (menia sa na elipsy); so zväčšujúcim sa rozdielom medzi σ_1 a σ_2 sa zväčšuje rozdiel rozptýlenia koncetrických kruhov v smere jednotlivých osí (hovoríme, že rozdiel variability premenných X_1 a X_2 sa zväčšuje)

Príklad 65 (dvojrozmerné normálne rozdelenie) *Nech náhodnou premennou X je najväčšia výška mozgovne u mužov (*skull.pH*; v mm) a náhodnou premennou Y je morfologická výška tváre u mužov (*face.H*; v mm); dáta:* one-sample-correlation-skull-mf.txt. *Nech $E[X] = \mu_1$ je stredná hodnota najväčšej výšky mozgovne a $Var[X] = \sigma_1^2$ je rozptyl najväčšej výšky mozgovne, $E[Y] = \mu_2$ je stredná hodnota morfologickej výšky tváre a $Var[Y] = \sigma_2^2$ je rozptyl morfologickej výšky tváre. Predpokladajme, že najväčšia výška mozgovne X má normálne rozdelenie $N(\mu_1, \sigma_1^2)$ a morfologická výška tváre Y má normálne rozdelenie $N(\mu_2, \sigma_2^2)$. Potom $(X, Y)^T$ má dvojrozmerné normálne rozdelenie $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ s parametrami $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, čo je vektor stredných hodnôt a σ_1^2, σ_2^2 a ρ, čo sú parametre kovariančnej matice $\boldsymbol{\Sigma}$, kde sila lineárneho vzťahu týchto dvoch premenných je daná veľkosťou a znamienkom ρ. Potom $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)^T$. (1) Nakreslite hustotu dvojrozmerného normálneho rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*

pomocou funkcie image() *a superponujte ho s kontúrovým grafom hustoty toho istého rozdelenia pomocou funkcie* contour(). *(2) Nakreslite dvojrozmerný jadrový odhad hustoty pomocou funkcií* kde2d() *a* image() *a superponujte ho s kontúrovým grafom hustoty dvojrozmerného normálneho rozdelenia* $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *pomocou funkcie* contour(). *Hustotu rozsekajte na 12 intervalov, kde hodnoty v týchto intervaloch budú zodpovedať farbám* terrain.colors(12). *Namiesto* $\boldsymbol{\theta}$ *použite vektor* $\widehat{\boldsymbol{\theta}} = (\overline{x}_1, \overline{x}_2, s_1^2, s_2^2, r)^T$ *odhadnutý z dát, kde* r *je Pearsonov korelačný koeficient. Riešenie pozri na obrázku 2.6.*



Obr. 2.6: Hustota dvojrozmerného normálneho rozdelenia s parametrom $\widehat{\boldsymbol{\theta}}$, ktorý je odhadnutý z dát (vľavo) a superimpozícia kontúr hustoty dvojrozmerného normálneho rozdelenia s parametrom $\widehat{\boldsymbol{\theta}}$, ktorý je odhadnutý z dát a dvojrozmerného jadrového odhadu hustoty (vpravo)

Štandardizované dvojrozmerné normálne rozdelenie. Náhodný vektor $(X, Y)^T$ má dvojrozmerné normálne rozdelenie

$$N_2(\mathbf{0}, \boldsymbol{\Sigma}), \text{ kde } \mathbf{0} = (0,0)^T \text{ a } \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

s hustotou (Bickel a Doksum, 2006)

$$\phi(x,y) = f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\},$$

kde $(x,y)^T \in \mathbb{R}^2$, $\rho \in \langle -1, 1\rangle$ sú parametre, potom $\boldsymbol{\theta} = (0,0,1,1,\rho)^T$. Výraz v exponente môžeme písať ako

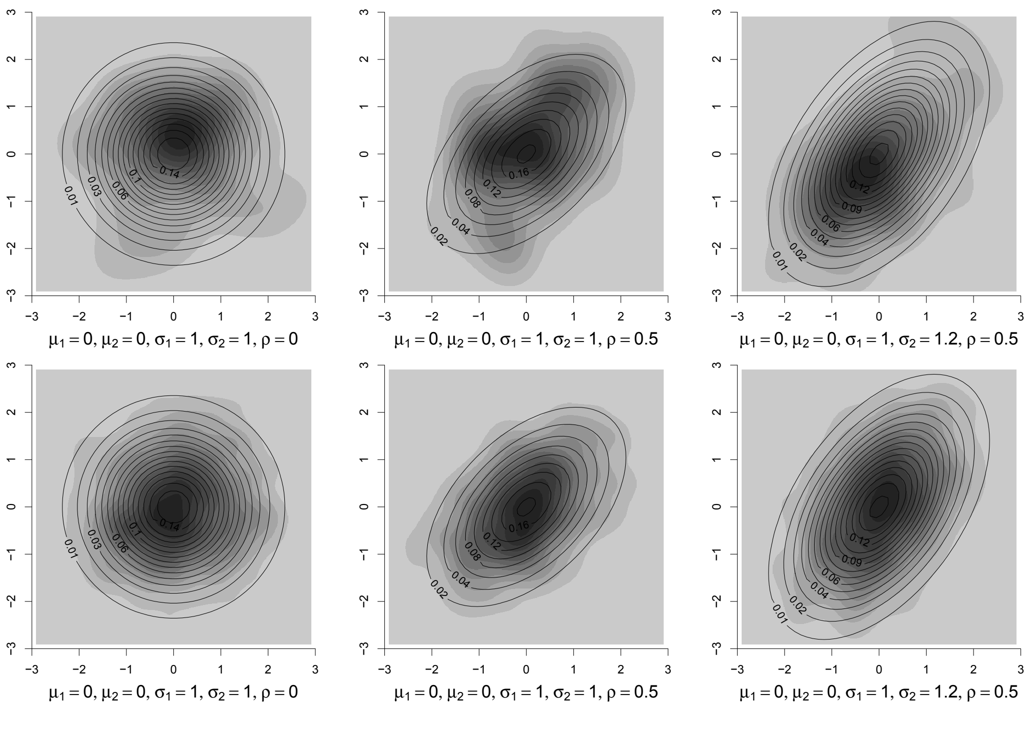
$$-\frac{1}{2}\begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix},$$

marginálne rozdelenia sú obe $N(0,1)$ a ρ je koeficient korelácie.

Príklad 66 (štandardizované dvojrozmerné normálne rozdelenie) *Nech náhodnou premennou* $X \sim N(\mu_1, \sigma_1^2)$ *je najväčšia výška mozgovne u mužov (*skull.pH*; v mm) a náhodnou premennou* $Y \sim N(\mu_2, \sigma_2^2)$ *je morfologická výška tváre u mužov (*face.H*; v mm). Nech* X *a* Y *majú dvojrozmerné normálne rozdelenie s parametrami* $(\mu_1, \mu_2)^T$ *a* σ_1^2, σ_2^2 *a* ρ *sú parametre kovariančnej matice* $\boldsymbol{\Sigma}$. *Keď od* X *odpočítame jej strednú hodnotu* μ_1 *a tento rozdiel vydelíme odmocninou z rozptylu* σ_1, *dostaneme náhodnú premennú* Z_X, *ktorá má asymptoticky normálne rozdelenie so strednou hodnotou* $\mu_1 = 0$ *a rozptylom* $\sigma_1^2 = 1$, *čo zapisujeme ako* $Z_X \sim N(0,1)$. *Keď od* Y *odpočítame jej strednú hodnotu* μ_2 *a tento rozdiel vydelíme odmocninou z rozptylu* σ_2, *dostaneme náhodnú premennú* Z_Y, *ktorá má asymptoticky normálne rozdelenie so strednou hodnotou* $\mu_2 = 0$ *a rozptylom* $\sigma_2^2 = 1$, *čo zapisujeme ako*

$Z_Y \sim N(0,1)$. Potom $(Z_X, Z_Y)^T$ má štandardizované dvojrozmerné normálne rozdelenie $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ s parametrami $\boldsymbol{\mu} = (0,0)^T$ a $\sigma_1^2 = 1$, $\sigma_2^2 = 1$ a ρ sú parametre kovariančnej matice $\boldsymbol{\Sigma}$ (dáta: one-sample-correlation-skull-mf.txt).

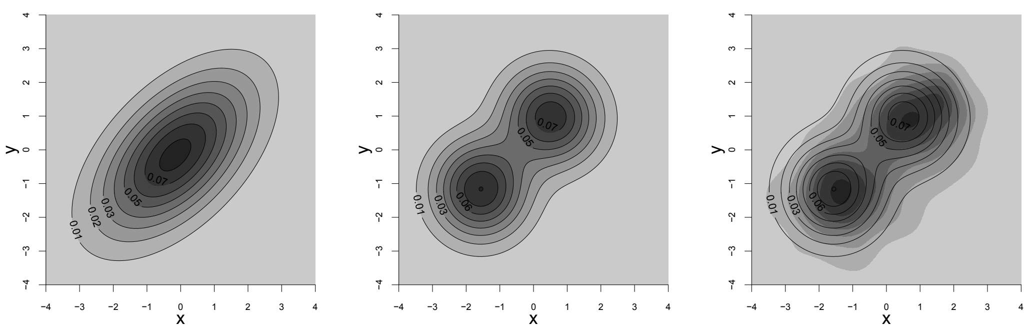
Príklad 67 (dvojrozmerné normálne rozdelenie) *Simuláciu pseudonáhodných čísel z $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ môžeme v R urobiť použitím nasledovných alternatívnych funkcií:*
1) knižnice library(MASS) a funkcie mvrnorm();
2) knižnice library(mvtnorm) a funkcie rmvnorm();
3) funkcie rnorm() a nasledovného algoritmu – nech $X_1 \sim N(0,1)$ a $X_2 \sim N(0,1)$; potom $(Y_1, Y_2) \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, čo je vektor stredných hodnôt a σ_1^2, σ_2^2 a ρ, čo sú parametre kovariančnej matice $\boldsymbol{\Sigma}$, kde sila lineárneho vzťahu Y_1 a Y_2 je daná veľkosťou a znamienkom ρ; $Y_1 = \sigma_1 X_1 + \mu_1$ a $Y_2 = \sigma_2(\rho X_1 + \sqrt{1-\rho^2} X_2) + \mu_2$. Nasimulujte pseudonáhodné čísla Y_1 a Y_2 z $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Vypočítajte dvojrozmerný jadrový odhad hustoty $(Y_1, Y_2)^T$ pomocou funkcie kde2d(). Nakreslite ho pomocou funkcie image() a superponujte ho s kontúrovým grafom hustoty dvojrozmerného normálneho rozdelenia $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ pomocou funkcie contour(). Hustotu rozsekajte na 12 intervalov, kde hodnoty v týchto intervaloch budú zodpovedať farbám terrain.colors(12). Pri simulácii použite nasledovné parametre
(a) $\mu_1 = 0, \mu_2 = 0$, $\sigma_1 = 1, \sigma_2 = 1$, $\rho = 0$; (1) $n = 50$ a (2) $n = 1000$;
(b) $\mu_1 = 0, \mu_2 = 0$, $\sigma_1 = 1, \sigma_2 = 1$, $\rho = 0.5$; (1) $n = 50$ a (2) $n = 1000$;
(c) $\mu_1 = 0, \mu_2 = 0$, $\sigma_1 = 1, \sigma_2 = 1.2$, $\rho = 0.5$; (1) $n = 50$ a (2) $n = 1000$.
Vzorové riešenie pozri na obrázku 2.7.



Obr. 2.7: Hustoty dvojrozmerného normálneho rozdelenia (prvý riadok $n = 50$; druhý riadok $n = 1000$)

Príklad 68 (zmes dvoch dvojrozmerných normálnych rozdelení) *Simuláciu pseudonáhodných čísel zo zmesi dvoch normálnych rozdelení $pN_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + (1-p)N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ môžeme v R urobiť použitím jedného z alternatínych postupov z príkladu 67. Nasimulujte pseudonáhodné čísla X a Y (1)*

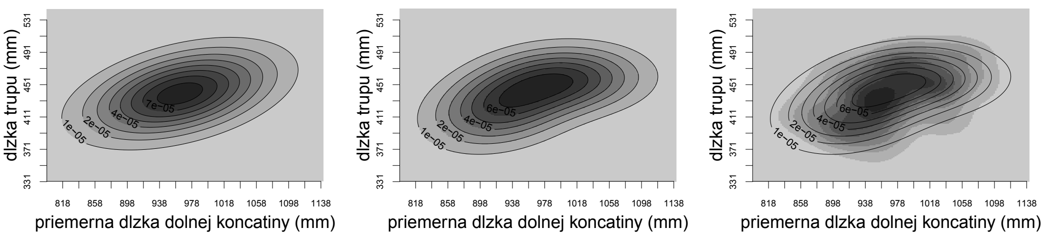
zo zmesi $pN_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + (1-p)N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$*, kde* $\boldsymbol{\theta} = (\mu_{11}, \mu_{12}, \sigma_{11}^2, \sigma_{12}^2, \rho_1, \mu_{21}, \mu_{22}, \sigma_{21}^2, \sigma_{22}^2, \rho_2)^T$ *a (2) z dvojrozmerného rozdelenia* $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, kde parametre predstavujú spoločný vektor stredných hodnôt a spoločnú kovariančnú maticu. t.j.* $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)^T$*. Pre (1) vypočítajte dvojrozmerný jadrový odhad hustoty* $(X, Y)^T$ *pomocou funkcie* kde2d()*.*
(a) Nakreslite teoretickú hustotu (2) pomocou funkcie image() *a superponujte ju s kontúrovým grafom teoretickej hustoty (2) pomocou funkcie* contour()*. Teoretickým rozdelením v tomto prípade bude* $N_2\left(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}\right)$*.*
(b) Nakreslite teoretickú hustotu (1) pomocou funkcie image() *a superponujte ju s kontúrovým grafom teoretickej hustoty (1) pomocou funkcie* contour()*.*
(c) Nakreslite dvojrozmerný jadrový odhad hustoty realizácí (1) pomocou funkcie image() *a superponujte ju s kontúrovým grafom teoretickej hustoty (1) pomocou funkcie* contour()*.*
Hustotu rozsekajte na 12 intervalov, kde hodnoty v týchto intervaloch budú zodpovedať farbám terrain.colors(12)*. Pri simulácii použite* $\boldsymbol{\theta} = (-1.2, -1.2, 1, 1, 0, 1, 1, 1, 1, 0)^T$*,*
(1) $\widehat{\boldsymbol{\theta}} = (\overline{x}_{11}, \overline{x}_{12}, s_{11}^2, s_{12}^2, r_1, \overline{x}_{21}, \overline{x}_{22}, s_{21}^2, s_{22}^2, r_2)^T$, $n_1 = n_2 = 50$ *a* $p = 0.5$ *(odhady pochádzajú z nasimulovaných dát).*
(2) $\widehat{\boldsymbol{\theta}} = (\overline{x}_1, \overline{x}_2, s_1^2, s_2^2, r)^T$ *a* $n_1 = n_2 = 50$ *(odhady pochádzajú zo spoločného výberu nasimulovaných dát).*
Vzorové riešenie pozri na obrázku 2.8.



Obr. 2.8: Spoločná hustota dvojrozmerného normálneho rozdelenia (vľavo), hustota zmesi dvoch dvojrozmerných normálnych rozdelení (uprostred) a dvojrozmerný jadrový odhad superponovaný hustotou zmesi dvoch dvojrozmerných normálnych rozdelení (vpravo) – simulačná štúdia

Príklad 69 (zmes dvoch dvojrozmerných normálnych rozdelení) *Nech* $(X_1, Y_1)^T$ *pochádza z rozdelenia* $N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$*, kde* X_1 *je priemerná dĺžka dolnej končatiny* lowex.L *v milimetroch a* Y_1 *dĺžka trupu* tru.L *v milimetroch (u mužov). Nech* $(X_2, Y_2)^T$ *pochádza z rozdelenia* $N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$*, kde* X_2 *je priemerná dĺžka dolnej končatiny* lowex.L *v milimetroch a* Y_2 *dĺžka trupu* tru.L *v milimetroch (u žien). Predpokladajme, že* X *je priemerná dĺžka dolnej končatiny a* Y *dĺžka trupu pochádzajú (1) zo zmesi* $pN_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + (1-p)N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$*, kde* $\boldsymbol{\theta} = (\mu_{11}, \mu_{12}, \sigma_{11}^2, \sigma_{12}^2, \rho_1, \mu_{21}, \mu_{22}, \sigma_{21}^2, \sigma_{22}^2, \rho_2)^T$ *a (2) z dvojrozmerného rozdelenia* $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, kde parametre predstavujú spoločný vektor stredných hodnôt a spoločnú kovariančnú maticu. t.j.* $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)^T$*. Pre (1) vypočítajte dvojrozmerný jadrový odhad hustoty* $(X, Y)^T$ *pomocou funkcie* kde2d()*.*
(a) Nakreslite teoretickú hustotu (2) pomocou funkcie image() *a superponujte ju s kontúrovým grafom teoretickej hustoty (2) pomocou funkcie* contour()*.*
(b) Nakreslite teoretickú hustotu (1) pomocou funkcie image() *a superponujte ju s kontúrovým grafom teoretickej hustoty (1) pomocou funkcie* contour()*.*
(c) Nakreslite dvojrozmerný jadrový odhad hustoty realizácí (1) pomocou funkcie image() *a superponujte ho kontúrovým grafom teoretickej hustoty (1) pomocou funkcie* contour()*.*
Hustotu rozsekajte na 12 intervalov, kde hodnoty v týchto intervaloch budú zodpovedať farbám terra-

in.colors(12).
(1) $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_{11}, \widehat{\mu}_{12}, \widehat{\sigma}^2_{11}, \widehat{\sigma}^2_{12}, \widehat{\rho}_1, \widehat{\mu}_{21}, \widehat{\mu}_{22}, \widehat{\sigma}^2_{21}, \widehat{\sigma}^2_{22}, \widehat{\rho}_2)^T$ *a* $p = n_1/(n_1+n_2)$*; parametre sú odhadnuté z dát.*
(2) $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2_1, \widehat{\sigma}^2_2, \widehat{\rho})^T$*; parametre sú odhadnuté zo spoločného výberu.*
Vzorové riešenie pozri na obrázku 2.9 (dáta two-samples-correlations-trunk.txt).



Obr. 2.9: Spoločná hustota dvojrozmerného normálneho rozdelenia (vľavo), hustota zmesi dvoch dvojrozmerných normálnych rozdelení (uprostred) a dvojrozmerný jadrový odhad superponovaný hustotou zmesi dvoch dvojrozmerných normálnych rozdelení (vpravo) – reálne dáta

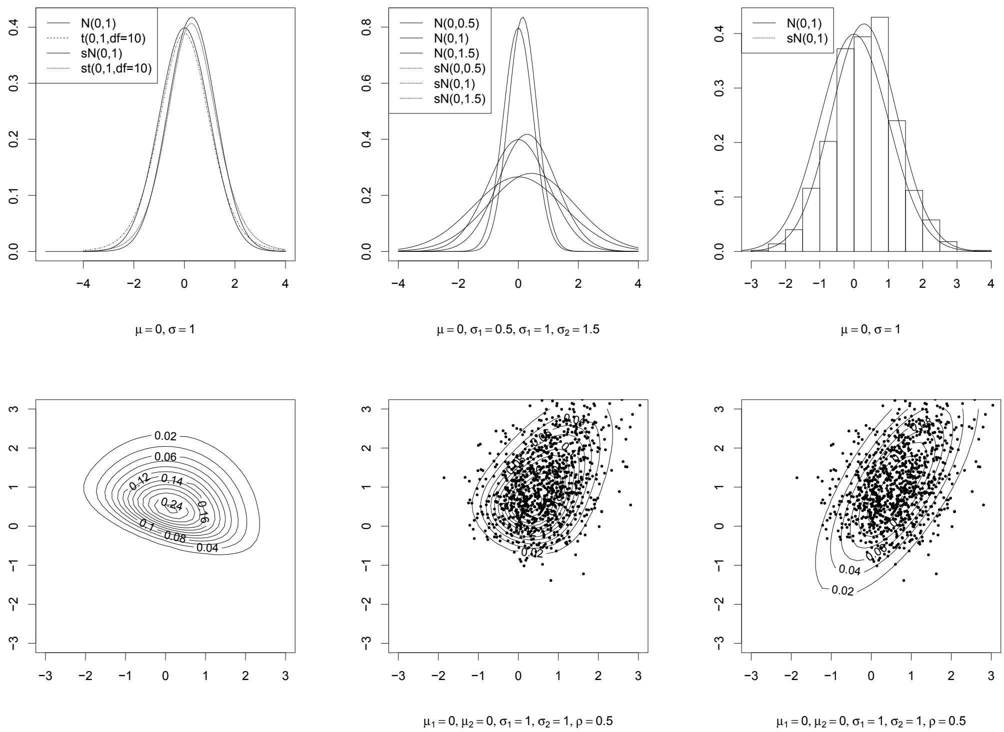
Odlišnosti od teoretického rozdelenia. Odlišnosti empirického rozdelenia (rozdelenia realizácií) od teoretického (napr. normálneho) rozdelenia, môžeme charakterizovať napr. ako pravostranne alebo ľavostranne zošikmené rozdelenie (obrázok 2.10, prvý riadok vľavo a vpravo), ploché alebo špicaté rozdelenie (obrázok 2.10, prvý riadok uprostred). Pri viacrozmerných rozdeleniach je situácia komplikovanejšia. Pri dvojrozmernom normálnom rozdelení može byť napr. zošikmená jedna alebo obe premenné (príklad zošikmenia oboch premenných zľava pozri na obrázku 2.10, dolný riadok).

Binomické rozdelenie. Majme nezávislé identické *Bernoulliho* pokusy s odpoveďami $X_i = 1$ (udalosť nastala) alebo $X_i = 0$ (udalosť nenastala) pre $i = 1, 2, ..., N$, kde N je počet nezávislých pokusov. Pravdepodobnosť nastatia udalosti pre každý pokus $\Pr(X_i = 1) = p$, pravdepodobnosť neúspechu pre každý pokus $\Pr(X_i = 0) = 1 - p$. Počet nastatí udalosti $X = \sum_{i=1}^{N} X_i$, pravdepodobnosť nastatia udalosti je p. Náhodná premenná X má binomické rozdelenie s parametrami N a p, t.j. $X \sim Bin\,(N, p)$, kde $\theta = p$. Pravdepodobnosť, že X je rovné nejakému číslu $x = n$ (x je realizácia X) zapisujeme ako $\Pr(X = x) = \binom{N}{x} p^x (1-p)^{N-x}$, pre $x = 0, 1, 2, \ldots, N$ (Christensen, 1997). Stredná hodnota náhodnej premennej X je definovaná ako $E[X] = \sum_{x=0}^{N} x \binom{N}{x} p^x (1-p)^{N-x} = Np$ a rozptyl $Var[X] = \sum_{x=0}^{N} (x - Np)^2 \binom{N}{x} p^x (1-p)^{N-x} = Np\,(1-p)$.

Strieška v označení $\widehat{p}$ sa všeobecne používa na označenie odhadov parametrov rozdelení pravdepodobnosti. Tieto odhady sa počítajú z dát. Pre spojité náhodné premenné označujeme rozsah náhodného výberu n, ale pri binomickom (a iných diskrétnych) rozdeleniach máme početnosti dve – počet úspechov a rozsah náhodného výberu – preto pre počet úspechov rezervujeme ozn. n a rozsah náhodného výberu N. Počet úspechov ozn. aj $x = n$, ak ide o realizáciu náhodnej premennej X.

Ekvivalentne môžeme rozdelenie náhodnej premennej pochádzajúcej z binomického rozdelenia zapísať ako $\mathbf{X} \sim Bin\,(N, p, 1-p)$, kde $\mathbf{X} = (X_1, X_2)^T$, $\boldsymbol{\theta} = (p, 1-p)^T$, X_1 je počet úspechov, $X_2 = N - X_1$ je počet neúspechov, $X_1 \sim Bin\,(N, p)$, $X_2 \sim Bin\,(N, 1-p)$. Potom $E[X_1] = Np$, $E[X_2] = N\,(1-p)$, $Var[X_2] = Np\,(1-p) = Var[X_1]$ nezávisle na p, $Cov\,[X_1, X_2] = -Np\,(1-p)$ a $Cor\,[X_1, X_2] = -1$. (Vysvetlenia označení pozri v poznámke pod čiarou[9]). Realizácie náhodných premenných X_1 a X_2 budeme označovať ako n_1 a n_2. Tento typ označenia je vhodnejší z dôvodu zovšeobecnenia binomického

[9] $E[X]$ je označenie pre strednú hodnotu náhodnej premennej X, $Var[X]$ pre rozptyl, $Cov[X, Y]$ pre kovarianciu dvoch premenných X a Y a $Cor[X, Y]$ pre korelačný koeficient.



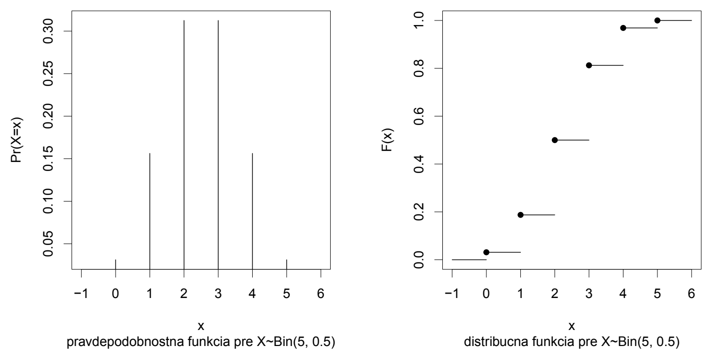
Obr. 2.10: Hustoty normálneho rozdelenia a zošikmeného normálneho rozdelenia pri rôznych parametroch (prvý riadok); hustoty dvojrozmerného zošikmeného normálneho rozdelenia (druhý riadok vľavo a uprostred) a dvojrozmerného normálneho rozdelenia (druhý riadok vpravo) pri rôznych parametroch

rozdelenia na viacrozmerné rozdelenia, kde $\mathbf{n} = (n_1, n_2)^T$ a $\mathbf{p} = (p_1, p_2)^T$, $p_1 = p$ a $p_2 = 1 - p$ (pre porovnanie pozri ozn. v práci Verzani, 2005). Potom $\boldsymbol{\theta} = \mathbf{p}$.

Príklad 70 (binomické rozdelenie, binomický experiment) *Experiment pozostávajúci z fixného počtu Bernouliho experimentov (ozn. N) sa nazýva binomický experiment. Pravdepodobnosť úspechu ozn. p, pravdepodobnosť neúspechu $q = 1 - p$. Náhodná premenná X je počet pozorovaných úspechov počas experimentu. Pravdepodobnosť $X = x$ za podmienky, že X pochádza z binomického rozdelenia $Bin(N, p)$ píšeme ako $\Pr(X = x) = \binom{N}{x} p^x (1-p)^{N-x}, x = 0, 1, \ldots, N$ (Ugarte a kol., 2008). Stredná hodnota $E[X] = Np$ a rozptyl $Var[X] = Np(1-p)$. Naprogramujte a zobrazte v R pravdepodobnostnú funkciu a (kumulatívnu) distribučnú funkciu pre Bin(5,0.5).*

Riešenie v R (pozri obrázok 2.11)

```
par(mfrow=c(1,2),mar=c(6,5,1,1),pty="s")
plot(0:5,dbinom(0:5,5,0.5),type="h",xlab="x",ylab="Pr(X=x)",
xlim=c(-1,6))
title(sub="hustota pre X~Bin(5,0.5)")
plot(0:5,pbinom(0:5,5,0.5),type="n",xlab="x",ylab="F(x)",
xlim=c(-1,6),ylim=c(0,1))
segments(-1,0,0,0)
segments(0:5,pbinom(0:5,5,.5),1:6,pbinom(0:5,5,.5))
lines(0:5,pbinom(0:5,5,.5),type="p",pch=16)
segments(-1,1,9,1,lty=2)
title(sub="distribucna funkcia pre X~Bin(5,0.5)")
```



Obr. 2.11: Pravdepodobnostná funkcia a distribučná funkcia $Bin(5, 0.5)$

Multinomické rozdelenie. Nech N je počet nezávislých identických pokusov a v každom z nich môže nastať $J \geq 2$ navzájom disjunktných udalostí s možnými odpoveďami $X_{ij} = 1$ (udalosť nastala) alebo $X_{ij} = 0$ (udalosť nenastala), kde $i = 1, 2, \ldots, N$ a $j = 1, 2, \ldots, J$. Potom $X_j = \sum_{i=1}^{N} X_{ij}$. Pravdepodobnosť nastatia j-tej udalosti v i-tom pokuse $\Pr(X_{ij} = 1) = p_j$, $\sum_{j=1}^{J} p_j = 1$. Náhodná premenná $\mathbf{X} = (X_1, X_2, \ldots, X_J)^T$ má (J-rozmerné) multinomické rozdelenie s parametrami N a $\mathbf{p}$, t.j. $\mathbf{X} \sim Mult_J\,(N, \mathbf{p})$, kde $\boldsymbol{\theta} = \mathbf{p}$ a $\mathbf{p} = (p_1, p_2, \ldots, p_J)^T$. Pravdepodobnosť, že X_j je rovné nejakému číslu n_j zapisujeme ako (Casella a Berger, 2002)

$$\Pr(X_1 = x_1, X_2 = x_2, \ldots, X_J = x_J) = \frac{N!}{x_1!x_2! \ldots x_J!} p_1^{x_1} p_2^{x_2} \ldots p_J^{x_J} = \frac{N!}{\prod_j x_j!} \prod_{j=1}^{J} p_j^{x_j},$$

kde $N = \sum_{j=1}^{J} X_j$, $X_j \geq 0$ a $x_j = n_j$ sú realizácie X_j. Potom $\mathbf{n} = (n_1, n_2, \ldots, n_J)^T$. Pre marginálne rozdelenie píšeme $X_j \sim Bin\,(N, p_j)$, kde stredná hodnota $E[X_j] = Np_j$, rozptyl $Var[X_j] = Np_j\,(1 - p_j)$, kovariancia $Cov\,[X_i, X_j] = -Np_i p_j$, korelačný koeficient $Cor\,[X_i, X_j] = (-p_i p_j) / \sqrt{p_i\,(1 - p_i)\, p_j\,(1 - p_j)}$. Stredná hodnota $E[\mathbf{X}] = N\mathbf{p}$ a kovariančná matica $Var\,[\mathbf{X}] = N\left(\mathbf{D_p} - \mathbf{pp}^T\right)$, kde $\mathbf{D_p} = \operatorname{diag}(\mathbf{p})$ a

$$\left(\mathbf{D_p} - \mathbf{pp}^T\right)_{ij} = \begin{cases} p_i\,(1 - p_i) & \text{ak } i = j \\ -p_i p_j & \text{ak } i \neq j \end{cases}.$$

Disjunktnosť znamená, že v i-tom pokuse mohla udalosť nastať len raz, t.j. výsledkom takéhoto pokusu môže byť vektor napr. $(1, 0, \ldots, 0)^T$ alebo $(0, 1, \ldots, 0)^T$, kde okrem jednej jednotky na j-tom mieste máme vždy $J-1$ núl na ostatných miestach. Kedže sumácia $\mathbf{p}$ nám dáva jednotku, $Var\,[\mathbf{X}]$ je singulárna matica. Matica diag($\mathbf{p}$) má diagonálu v podobe $\mathbf{p}$ a mimodiagonálne prvky rovné 0. Ak $J = 2$, potom $Bin\,(N, p) \approx Mult_2\,(N, \mathbf{p})$, t.j. multinomické rozdelenie je zovšeobecnením binomickeho rozdelenia.

Príklad 71 (multinomické rozdelenie; príklady) *Príklady premenných, o ktorých predpokladáme, že majú multinomické rozdelenie:*
(1) farba dúhovky – hodnotená podľa škály R. Martina (Martin, 1914/1928) a kategorizovaná do štyroch kategórií: hnedá, hnedozelená, melírovaná a modrá (dáta: multinom-iris-color.txt*);*
(2) zakončenie troch hlavných dlaňových línií – kategorizované do troch kategórií: vysoké, stredné a nízke (dáta: multinom-palmar-lines.txt*);*
(3) priľahlosť ušného laloka – podľa priľahlosti k hlave kategorizovaná do troch kategórií: priľahlý, stredne priľahlý, odstávajúci (dáta: multinom-earlobe.txt*);*
(4) krvná skupina – kategorizovaná v AB0 systéme do štyroch kategórií: skupina 0, A, B a AB (dáta: multinom-blood-groups.txt*).*

Príklad 72 (multinomické rozdelenie) *Majme náhodné premenné (1) socioekonomický status (vysoký – H, nízky – Lo), (2) politická prislušnosť (demokrat – D, republikán – R) a (3) politická filozofia (liberál – Li, konzervatívec – C). Označme ich interakcie nasledovne* X_1 *(H-D-Li),* X_2 *(H-D-C),* X_3 *(H-R-Li),* X_4 *(H-R-C),* X_5 *(Lo-D-Li),* X_6 *(Lo-D-C),* X_7 *(Lo-R-Li) a* X_8 *(Lo-R-C). Predpokladajme, že máme náhodný výber s rozsahom* $N = 50$. *Pravdepodobnosti* p_j *pozri v tabuľke 2.2. Vypočítajte* $Var[X_1]$, $Var[X_3]$, $Cov\,[X_1, X_3]$, $Cor\,[X_1, X_3]$ *a očakávané početnosti* $Np_j, j = 1, 2, \ldots, 8$.

Tabuľka 2.2: Kontingenčná tabuľka 2×3 pravdepodobností p_j pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (multinomické rozdelenie)

	D-Li	D-C	R-Li	R-C	spolu
H	0.12	0.12	0.04	0.12	0.4
Lo	0.18	0.18	0.06	0.18	0.6
spolu	0.30	0.30	0.10	0.30	1.0

Riešenie

$X = (X_1, X_2, \ldots, X_8) \sim Mult_J(N, \mathbf{p})$, kde $N = 50$, $\mathbf{p} = (p_1, p_2, \ldots, p_8)^T$, vieme, že $X_j \sim Bin(N, p_j)$, p_j sú v tabuľke 2.2 a $j = 1, 2, \ldots, 8$. Potom
$Var[X_1] = 50 \times 0.12 \times (1 - 0.12) = 5.28$,
$Var[X_3] = 50 \times 0.04 \times (1 - 0.04) = 1.92$.
Vybraná kovariancia a korelácia (medzi počtami prislušných skupín) je rovná
$Cov\,[X_1, X_3] = -50 \times 0.12 \times 0.04 = -0.24$, $Cor\,[X_1, X_3] = -0.24/\sqrt{5.28 \times 1.92} = -0.075$.
Očakávané početnosti pre každú bunku tabuľky sú (všeobecne nemusia byť) celé čísla, pozri tabuľku 2.3.

Tabuľka 2.3: Kontingenčná tabuľka 2×3 očakávaných početností Np_j pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (multinomické rozdelenie)

	D-Li	D-C	R-Li	R-C
H	6	6	2	6
Lo	9	9	3	9

Súčinové multinomické rozdelenie. Nech N_k je počet nezávislých identických pokusov a v každom z nich môže nastať $J \geq 2$ navzájom disjunktných udalostí s možnými odpoveďami $X_{kji} = 1$ (udalosť nastala) alebo $X_{kji} = 0$ (udalosť nenastala), kde $i = 1, 2, \ldots, N_k$, $k = 1, 2, \ldots, K$ a $j = 1, 2, \ldots, J$. Nech $X_{kj} = \sum_{i=1}^{N_k} X_{kji}$ a $\sum_{k=1}^{K} N_k = N$. Pravdepodobnosť nastatia (kj)-tej udalosti v i-tom pokuse $\Pr(X_{kji} = 1) = p_{kj}$, $\sum_{k=1}^{K}\sum_{j=1}^{J} p_{kj} = 1$. Náhodná premenná $\mathbf{X}_k = (X_{k1}, X_{k2}, \ldots, X_{kJ})^T$ má (J-rozmerné) multinomické rozdelenie s parametrami N_k a $\mathbf{p}_k$, t.j. $Mult_J\,(N_k, \mathbf{p}_k)$, kde $\boldsymbol{\theta}_k = \mathbf{p}_k$ a $\mathbf{p}_k = (p_{k1}, p_{k2}, \ldots, p_{kJ})^T$. Realizácie náhodnej premennej $\mathbf{X}_k$ označujeme ako $\mathbf{x}_k$. Potom $x_{kj} = n_{kj}$ a naviac $\mathbf{n}_k = (n_{k1}, n_{k2}, \ldots, n_{kJ})^T$. Nech $\mathbf{X}_k$ sú nezávislé, potom platí (Christensen, 1997)

$$\Pr(X_{kj} = x_{kj}, \forall k, j; j = 1, 2, \ldots, J; k = 1, 2, \ldots, K) = \prod_{k=1}^{K} \Pr(X_{kj} = x_{kj}, \forall j; j = 1, 2, \ldots, J)$$

a $\sum_{j=1}^{J} X_{kj} = N_k$ pre $\forall k$. Ďalej platí

$$\Pr(X_{kj} = x_{kj}, \forall j) = \left(N_k!/\prod_{j=1}^{J} x_{kj}!\right)\prod_{j=1}^{J} p_{kj}^{x_{kj}}.$$

Z toho vyplýva, že

$$\Pr(X_{kj} = x_{kj}, \forall k, j; j = 1, 2, \ldots, J; k = 1, 2, \ldots, K) = \prod_{k=1}^{K}\left(\left(N_k!/\prod_{j=1}^{J} x_{kj}!\right)\prod_{j=1}^{J} p_{kj}^{x_{kj}}\right).$$

Očakávané hodnoty $N_k p_{kj}$, rozptyly $Var[X_{kj}]$ a kovariancie $Cov[X_{kj}]$ a korelácie $Cor[X_{kj}]$ vnútri nejakého $\mathbf{X}_k$ vieme vypočítať. Kovariancie medzi rôznymi $\mathbf{X}_k$, napr. $Cov[\mathbf{X}_1, \mathbf{X}_2]$, sú nuly kvôli nezávislosti jednotlivých $\mathbf{X}_k$. Potom $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_K)^T$ má súčinové multinomické rozdelenie s parametrami $\boldsymbol{\theta}_k = \mathbf{p}_k, k = 1, 2, \ldots, K$.

Príklad 73 (súčinové multinomické rozdelenie; príklady) *Príklady premenných, o ktorých predpokladáme, že majú súčinové multinomické rozdelenie:*
(1) farba dúhovky – hodnotená podľa škály R. Martina (Martin, 1914/1928) a kategorizovaná do štyroch kategórií (hnedá, hnedozelená, melírovaná a modrá) zisťovaná súčasne s výskytom (prítomnosť, neprítomnosť) radiálnych útvarov v štruktúre dúhovky (dáta: multinom-iris-color.txt);
(2) zakončenie hlavných dlaňových línií – kategorizované do troch kategórií (vysoké, stredné a nízke) zisťované súčasne s odtieňmi farby vlasov (svetlá, stredná a tmavá) na základe štandardnej Fischer-Sallerovej stupnice odtieňov (dáta: multinom-palmar-lines.txt);
(3) priľahlosť ušného laloka – podľa priľahlosti k hlave kategorizovaná do troch kategórií (priľahlý, stredne priľahlý, odstávajúci) zisťovaná u mužov a u žien (dáta: multinom-earlobe.txt);
(4) krvná skupina – kategorizovaná v AB0 systéme do štyroch kategórií (0, A, B a AB) zisťovaná v Košiciach a v Prahe (dáta: multinom-blood-groups.txt).

Príklad 74 (súčinové multinomické rozdelenie) *Majme dáta z príkladu 72 a náhodný výber s rozsahom $N_1 = 30$ zo skupiny H, ďalší náhodný výber s rozsahom $N_2 = 20$ zo skupiny Lo. Označme interakcie premenných nasledovne $X_{11} = X_{1|1}$ (H-D-Li), $X_{12} = X_{2|1}$ (H-D-C), $X_{13} = X_{3|1}$ (H-R-Li), $X_{14} = X_{4|1}$ (H-R-C), $X_{21} = X_{1|2}$ (Lo-D-Li), $X_{22} = X_{2|2}$ (Lo-D-C), $X_{23} = X_{3|2}$ (Lo-R-Li) a $X_{24} = X_{4|2}$ (Lo-R-C), kde $\mathbf{X}_1 = (X_{11}, X_{12}, X_{13}, X_{14})^T$ a $\mathbf{X}_2 = (X_{21}, X_{22}, X_{23}, X_{24})^T$. Potom $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ má súčinové multinomické rozdelenie s $K = 2$, $N_1 = 30, J_1 = 4, N_2 = 20, J_2 = 4$. Zápis s $X_{j|k}$, kde $j = 1, 2, 3, 4$ a $k = 1, 2$ zvýrazňuje fakt, že rozdelenie je podmienené socioekonomickým statusom (vysoký – H, nízky – Lo), t.j. rozdelenie v stĺpcoch tabuľky je podmienené jej riadkom. Realizácie $X_{j|k}$ označujeme ako $n_{j|k} = n_{kj}$, pravdepodobnosti ekvivalentné $X_{j|k} = X_{kj}$ ako $p_{j|k} = p_{kj}$. Vypočítajte podmienené pravdepodobnosti $p_{j|k}$, očakávané početnosti $N_k p_{kj}$, $Var[X_{13}]$, $Cov\,[X_{21}, X_{23}]$ a $Cor\,[X_{11}, X_{23}]$.*

Riešenie

Pravdepodobnosti štyroch kategórií asociovaných s H statusom sú podmienené pravdepodobnosti dané H statusom. Napr. $\Pr(X_{3|1}) = 0.04/0.4 = 0.1$. $\Pr(X_{1|1}) = 0.12/0.4 = 0.3$, $\Pr(X_{3|2}) = 0.06/0.6 = 0.1$. Musíme ale tabuľku 2.2 prepísať na súčinovo-multinomický model, teda podmienené pravdepodobnosti $p_{j|i}$ budú dané socioekonomickým statusom i (pozri tabuľku 2.4).

Tabuľka 2.4: Kontingenčná tabuľka 2×3 pravdepodobností $p_{j|i}$ pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (súčinové multinomické rozdelenie)

	D-Li	D-C	R-Li	R-C	spolu
H	0.3	0.3	0.1	0.3	1.0
Lo	0.3	0.3	0.1	0.3	1.0

Pre $N_1 = 30$ a $N_2 = 20$ pozri očakávané početnosti v tabuľke 2.5.

Tabuľka 2.5: Kontingenčná tabuľka 2×3 očakáných početností $N_i p_{j|i}$ pre dva socioekonomické statusy, dve politické príslušnosti a dve politické filozofie (súčinové multinomické rozdelenie)

	D-Li	D-C	R-Li	R-C	spolu
H	9	9	3	9	30
Lo	6	6	2	6	20

$Var(X_{3|1}) = 30 \times 0.1 \times (1 - 0.1) = 2.7$.
Vybrané kovariancie (medzi počtami prislušných skupín) sú rovné

$Cov\left[X_{1|2}, X_{3|2}\right] = -20 \times 0.3 \times 0.1 = -0.6\,,$
$Cov\left[X_{1|1}, X_{3|2}\right] = 0$, lebo $\mathbf{X}_1$ a $\mathbf{X}_2$ sú nezávislé.

Príklad 75 (farba očí a vlasov) *Majme premenné farba vlasov (blond BlH, hnedá BrH, ryšavá RH) a farba očí (modrá BlE, hnedá BrE, zelená GE). Ich interakcie sú usporiadané v tabuľke ako* X_1 *(BlH-BlE),* X_2 *(BlH-BrE),* X_3 *(BlH-GE),* X_4 *(BrH-BlE),* X_5 *(BrH-BrE),* X_6 *(BrH-GE),* X_7 *(RH-BlE),* X_8 *(RH-BrE),* X_9 *(RH-GE). Nim zodpovedajúce pravdepodobnosti* $p_j, j = 1, 2, \ldots, 9$*, pozri v tabuľke 2.6.* $\mathbf{X} = (X_1, X_2, \ldots, X_9)^T \sim Mult_9(N, \mathbf{p})$*. Transformujte multinomický model na súčinový multinomický model nasledovne – vypočítajte (a) riadkové marginálne pravdepodobnosti* $p_{1\cdot} = \sum_{j=1}^{3} p_j$*,* $p_{2\cdot} = \sum_{j=4}^{6} p_j$*,* $p_{3\cdot} = \sum_{j=7}^{9} p_j$*, (b) stĺpcové marginálne pravdepodobnosti* $p_{\cdot 1} = p_1 + p_4 + p_7$*,* $p_{\cdot 2} = p_2 + p_5 + p_8$*,* $p_{\cdot 3} = p_3 + p_6 + p_9$*, (c) podmienené pravdepodobnosti* $p_{j|k} = p_{kj}$*; (d) podmienené pravdepodobnosti* $p_{k|j} = p_{jk}$*; (d) akému číslu sú rovné sumy* $\sum_{j=1}^{3} p_{j|k}$ *pre každé k a* $\sum_{k=1}^{3} p_{k|j}$ *pre každé j?*

Tabuľka 2.6: Kontingenčná tabuľka 3×3 pravdepodobností p_j pre tri farby vlasov a tri farby očí (multinomické rozdelenie)

farba vlasov/farba očí	modrá (BlE)	hnedá (BrE)	zelená (GE)
blond (BlH)	0.12	0.15	0.03
hnedá (BrH)	0.22	0.34	0.04
ryšavá (RH)	0.06	0.01	0.03

Riešenie (čiastkové)
Marginálne pravdepodobnosti sú
$\Pr(BlH) = 0.3$, $\Pr(BrH) = 0.6$, $\Pr(RH) = 0.1$,
$\Pr(BlE) = 0.4$, $\Pr(BrE) = 0.5$, $\Pr(GE) = 0.1$.
Podmienené pravdepodobnosti $p_{k|j}$ sú
$\Pr(BlH|BlE) = \Pr(BlH \cap BlE) / \Pr(BlE) = 0.12/0.4 = 0.3$,
$\Pr(BlH|BlE) = \Pr(BlH)$,
$\Pr(BrH|BlE) = 0.22/0.4 = 0.55$,
$\Pr(BrH) = 0.6$.
Ak vieme, že niekto má modré oči, potom bude menej pravdepodobné, že má hnedé vlasy v porovnaní s tým, keď nevieme, akej farby má oči. Teda
$\Pr(BlE|BlH) = 0.12/0.3 = 0.4$,
$\Pr(BlE|BlH) = \Pr(BlE)$,
$\Pr(BrE|BlH) = \Pr(BrE)$,
$\Pr(GE|BlH) = \Pr(GE)$.
Informácia, že má niekto blond vlasy, nám nedáva ďalšiu informáciu o farbe jeho očí.

Binomické, multinomické a súčinové multinomické rozdelenie sú vhodné v prípadoch, keď máme počet pokusov N nie príliš veľké a pravdepodobnosti výskytu udalostí p nie príliš malé. V opačnom prípade je vhodné Poissonovo rozdelenie (Agresti, 2002).

Poissonovo rozdelenie. Poissonovo rozdelenie je limitným pripadom Binomického rozdelenie $Bin(N, p)$, kde $N \to \infty, p \to 0$, teda $Np \to \lambda$. Ak je X náhodná premenná s Poissonovým rozdelením a parametrom $\theta = \lambda$, $X \sim Poiss(\lambda)$, potom (Casella a Berger, 2002)

$$\Pr(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \ldots.$$

Realizáciu náhodnej premennej X označujeme ako $x = n$. $E[X] = \lambda$ a $Var[X] = \lambda$. Tomu korešponduje

$$\binom{N}{x} p^x (1-p)^{N-x} = \left[(Np)^x (1-p)^N / x!\right] (1-p)^{-x} \frac{N!}{(N-x)! N^x},$$

ak $N \to \infty, p \to 0$ a $Np \to \lambda \Rightarrow$ $(Np)^x \to \lambda^x$, $(1-p)^N \to e^{-\lambda}$, $(1-p)^{-x} \to 1$, $\frac{N!}{(N-x)!N^x} \to 1$.

Znak „→“ čítame ako „konverguje“ (do nekonečna, k nule, k λ a pod.). V prípade, že za „→“ je znak nekonečna, napr. $N \to \infty$, potom čítame celý výraz ako „pre (dostatočne) veľké N“ alebo „N konverguje do nekonečna“.

Príklad 76 (Poissonovo rozdelenie; počet havárií za týždeň) *Ak každý z 50 miliónov ľudí šoféruje auto v Taliansku budúci týždeň nezávisle, potom pravdepodobnosť smrti pri autonehode bude* 0.000002*, kde počet úmrtí má binomické rozdelenie* $Bin(50mil, 0.000002)$ *alebo limitne Poissonovo rozdelenie s parametrom* $50mil \times 0.000002 = 100$.

Príklad 77 (Poissonovo rozdelenie; príklady) *Príklady premenných, o ktorých predpokladáme, že majú Poissonove rozdelenie:*
(1) početnosť absencie palmárneho trirádia d (Okajima a Iwayanagi, 1986);
(2) incidencia os japonicum (os zygomaticum bipartitum/tripartitum), variety lícnej kosti, kedy bežne celistvá kosť je aj v dospelosti rozdelená švami na dve až tri časti, pozri Hauser, De Stefano a kol. (1989, s. 222) a Dodo (1974);
(3) celková perioperačná mortalita, t.j. mortalita v súvislosti s výkonom chirurgických operačných zákrokov (Bainbridge a kol., 2012);
(4) výskyt ďalšieho infarktu myokardu do 30 dní po operácii u pacientov, ktorí boli liečení na ochorenie koronárnych tepien chirurgickým zásahom pomocou by-passu, kde operácia prebehla klasickým spôsobom so zastavením srdca a umelou cirkuláciou (Cheng a kol., 2005);
(5) výskyt ďalšieho infarktu myokardu do 30 dní po operácii u pacientov, ktorí boli liečení na ochorenie koronárnych tepien chirurgickým zásahom pomocou by-passu, kde operácia prebehla moderným postupom na tlčúcom srdci (off-pump coronary artery bypass; Cheng a kol. (2005);
(6) výskyt bilaterálnej agenézy horných laterálnych rezákov 12 *a* 22 *(Alt a kol., 2013).*

Príklad 78 (Poissonovo rozdelenie; pruské armádne jednotky) *Nech početnosti Pruských armádnych jednotiek, v ktorých nastalo n úmrtí zapríčinených kopnutím koňom za rok (von Bortkiewicz, 1898), má Poissonovo rozdelenie s parametrom* λ*, t.j.* $X \sim Poiss(\lambda)$*. Pravdepodobnosť, že niekto bude smrteľne zranený v danom dni je extrémne malá. Majme* 10 *vojenských jednotiek za* 20*-ročnú periódu s rozsahom* $M = 200$ *(*$200 = 10 \times 20$*), kde popri početnostiach úmrtí* $n = 1, 2, 3, 4, 5+$*, v danej jednotke a v danom roku, zaznamenávame aj početnosti vojenských jednotiek* m_n *pri danom n, kde* $M = \sum m_n$ *(pozri tabuľku 2.7). Vypočítajte očakávané početnosti, za predpokladu* $X \sim Poiss(\lambda)$*, kde* $\lambda = \frac{\sum_n n m_n}{\sum_n m_n}$.

Tabuľka 2.7: Pozorované a očakávané početnosti m_n (zaokrúhlené na nula desatinných miest) Pruských armádnych jednotiek, v ktorých nastalo n úmrtí zapríčinených kopnutím koňom

n	0	1	2	3	4	5+
pozorované m_n	109	65	22	3	1	0
očakávané m_n	109	66	20	4	1	0

Príklad 79 (Poissonove rozdelenie; tri typy havárií) *Nech* n_1 *je počet ľudí, ktorí zahynú pri automobilovej nehode,* n_2 *je počet ľudí, ktorí zahynú pri havárii lietadla,* n_3 *je počet ľudí, ktorí zahynú pri havárii vlaku v Taliansku budúci týždeň. Potom Poissonov model pre* X_1*,* X_2 *a* X_3 *pre nezávislé poissonovské náhodné premenné s parametrami* λ_1*,* λ_2 *a* λ_3 *je definovaný ako* $X_1 + X_2 + X_3 \sim Poiss(\lambda_1 + \lambda_2 + \lambda_3)$.

Zovšeobecnením príkladu 79 dostaneme

$$X_1 + X_2 + \ldots + X_J \sim Poiss\left(\lambda_1 + \lambda_2 + \ldots + \lambda_J\right)$$

Poissonovo vs. multinomické rozdelenie. Dá sa ukázať, že

$$(X_1 + X_2 + \ldots + X_J)\,|N \sim Mult_J\,(N, p_1, p_2, \ldots, p_J)\,,$$

kde $N = \sum_j X_j$ a $p_j = \lambda_j / \sum_j \lambda_j, j = 1, 2, \ldots, J$. Ak $X_j, j = 1, 2, \ldots, J$ sú nezávislé, $X_j \sim Poiss\,(\lambda_j)$, kde $E(X_j) = \lambda_j$, potom podmienená pravdepodobnosť, že všetky $X_j = x_j$ za podmienky $N = \sum_j X_j$ sa rovná

$$\begin{aligned}
\Pr\left[(X_1 = x_1, X_2 = x_2, \ldots, X_J = x_J)|\sum_j X_j = N\right] &= \frac{\Pr(X_1 = x_1, X_2 = x_2, \ldots, X_J = x_J)}{\Pr(\sum_j X_j = N)} \\
&= \frac{\prod_j \lambda_j^{x_j} e^{-\lambda_j}/x_j!}{\lambda^N e^{-\lambda}/N!} = \frac{N! e^{-\lambda} \prod_j \lambda_j^{x_j}}{\prod_j \lambda^x e^{-\lambda} x_j!} \\
&= \frac{N!}{\prod_j x_j!} \prod_j \left(\frac{\lambda_j}{\lambda}\right)^{x_j}, \text{ kde } p_j = \frac{\lambda_j}{\lambda}.
\end{aligned}$$

Toto podmienené rozdelenie sa často používa pri log-lineárnych modeloch. Ak máme početnosti X_j pochádzajúce z Poissonovho rozdelenia, potrebujeme ich celkovú sumu N (*grand total*). Teda potrebujeme podmiené rozdelenie pri danom N, čo je v podstate multinomické rozdelenie.

Overdispersion a underdispersion. V praxi často variabilita presahuje variabilitu danú binomickým či Poissonovým modelom alebo je variabilita menšia ako variabilita daná binomickým či Poissonovým modelom. Prepokladáme, že každý človek má rovnakú pravdepodobnosť úmrtia pri dopravnej nehode budúci týždeň. Realistickejšie však tieto pravdepodobnosti varírujú napr. podľa času stráveného šoférovaním, závisia od toho, či osoba má zapnuté pásy, závisia od geografickej polohy a pod.

V prípade *overdispersion*, teda v prípade, keď rozptyl presahuje strednú hodnotu, je realistickejšie nahradiť Poissonovo rozdelenie **negatívne binomickým rozdelením** (Agresti, 2002). Ak máme binomické (prip. multinomické) rozdelenie, tiež môže nastať prípad *overdispersion*, pretože skutočné rozdelenie je zmes rôznych binomických rozdelení s parametrami varírujúcimi kvôli nenameraným (mätúcim) premenným.

Negatívne binomické rozdelenie. Majme nezávislé identické *Bernoulliho* pokusy s odpoveďami $X_i = 1$ (udalosť nastala) alebo $X_i = 0$ (udalosť nenastala) pre $i = 1, 2, \ldots$. Pravdepodobnosť nastatia udalosti pre každý pokus $\Pr(X_i = 1) = p$, pravdepodobnosť neúspechu pre každý pokus $\Pr(X_i = 0) = 1 - p$. Nech X je počet úspechov pred k-tym neúspechom. Potom $\Pr(X = x) = \binom{x+k-1}{x} p^x (1-p)^k$ má negatívne binomické rozdelenie s $E[X] = k\frac{p}{1-p}$ a $Var[X] = k\frac{p}{(1-p)^2}$, ozn. $X \sim Negbin(k, p)$. Poissonovo rozdelenie je limitným prípadom negatívne binomického rozdelenia, kde $k \to \infty$, $p \to 0$ a fixovaným $kp = \lambda$.

Príklad 80 (podiel chlapcov a dievčat v rodinách) *Nech X predstavuje početnosť chlapcov medzi deťmi v rodinách. Tu môžeme predpokladať, že $X \nsim Bin(N, p)$, t.j. rodina môže mať vychýlený pomer pohlaví detí v smere ku chlapcom alebo dievčatám. V realite teda môžeme mať príliš veľa rodín len s chlapcami alebo len s dievčatami a nemáme dostatok rodín s pomerom pohlaví blízkym* 51 : 49 *(pomer chlapcov ku dievčatám). Z toho nám vyplýva, že rozptyl početnosti chlapcov bude v skutočnosti väčší ako rozptyl predpokladaný binomickým modelom $Bin(N, p)$.*

Overdispersion v binomickom modeli. Nech $X_i \sim Bin(N, p_i)$ a nech $X = X_I$ je náhodne zvolené z X_i, kde náhodný index $I = i$ má pravdepodobnosť $1/m$. Budeme teda mať zmes binomických rozdelení s marginálnou pravdepodobnosťou

$$\Pr(X = x) = E[\Pr(X_I = x|I)] = \frac{1}{m}\sum_{i=1}^{m} \Pr(X_i = x) = \frac{1}{m}\sum_{i=1}^{m} \binom{N}{x} p_i^x (1 - p_i)^x,$$

$$E[X] = E[E[X_I|I]] = \frac{1}{m}\sum_{i=1}^{m} E[X_i] = \frac{N}{m}\sum_{i=1}^{m} p_i = N\pi,$$

kde $\pi = \sum_{i=1}^{m} p_i/m$ a

$$\begin{aligned} Var[X] &= E[Var[X_I|I]] + Var[E[X_I|I]] = \frac{1}{m}\sum_{i=1}^{m} Var[X_i] + \frac{1}{m}\sum_{i=1}^{m} E^2[X_i] - \left(\sum_{i=1}^{m} E[X_i]/m\right)^2 \\ &= \frac{1}{m}\sum_{i=1}^{m} Np_i(1-p_i) + \frac{1}{m}\sum_{i=1}^{m}(Np_i)^2 - (N\pi)^2 = N\pi(1-\pi) + N(N-1)\sigma_p^2, \end{aligned}$$

kde σ_p^2 je definovaná ako pri *underdispersion*. Z toho vyplýva, že pri náhodnej voľbe z rôznych individuálnych pravdepodobností p_i je rozptyl väčší ako rozptyl za platnosti binomického modelu.

Príklad 81 (overdispersion v binomickom modeli) *V klasickej štúdii pomeru pohlaví u ľudí z roku 1889 na základe záznamov z nemocníc v Sasku (bližšie pozri Lindsey a Altham, 1998; Geissler, 1889) zaznamenal rozdelenie počtu chlapcov v rodinách. Medzi $M = 6115$ rodinami s $N = 12$ deťmi pozoroval početnosti chlapcov (pozri tabuľku 2.8). Vypočítajte m_n za predpokladu, že početnosti chlapcov X v rodinách majú binomické rozdelenie s parametrami $\pi = \frac{\sum_{n=0}^{N} nm_n}{NM} = 0.5192$ a $N = 12$, ozn. $X \sim Bin(N, \pi)$.*

Tabuľka 2.8: Pozorované početnosti rodín m_n s n chlapcami

n	0	1	2	3	4	5	6	7	8	9	10	11	12
pozorované m_n	3	24	104	286	670	1033	1343	1112	829	478	181	45	7

Riešenie

Tabuľka 2.9: Očakávané početnosti rodín m_n (zaokrúhlené na nula desatinných miest) s n chlapcami (binomické rozdelenie)

n	0	1	2	3	4	5	6	7	8	9	10	11	12
očakávané m_n	1	12	72	258	628	1085	1367	1266	854	410	133	26	2

Keď porovnáme pozorované m_n (pozri tabuľku 2.8) a vypočítané (teoretické) m_n (pozri tabuľku 2.9) zistíme, že pozorované poukazujú na *overdispersion*, t.j. máme väčšie početnosti rodín s malým a veľkým množstvom chlapcov v porovnaní s teroretickými početnosťami.

Underdispersion v binomickom modeli. Nech $X_1, X_2, \ldots, X_N$ sú nezávislé binomické pokusy s pravdepodobnosťami $p_1, p_2, \ldots, p_N$. Nech $X = \sum_{i=1}^{N} X_i$, potom $E[X] = \sum_{i=1}^{N} p_i = N\pi$, kde $\pi = \frac{1}{N}\sum_{i=1}^{N} p_i$, ale

$$\begin{aligned} Var[X] &= \sum_{i=1}^{N} Var[X_i] = \sum_{i=1}^{N} p_i(1-p_i) = \sum_{i=1}^{N} p_i - \sum_{i=1}^{N} p_i^2 \\ &= \sum_{i=1}^{N} p_i - \frac{\left(\sum_{i=1}^{N} p_i\right)^2}{N} - \left(\sum_{i=1}^{N} p_i^2 - \frac{\left(\sum_{i=1}^{N} p_i\right)^2}{N}\right) \\ &= N\pi - N\pi^2 - N\sigma_p^2 = N\pi(1-\pi) - N\sigma_p^2, \end{aligned}$$

kde $\sigma_p^2 = \frac{1}{N}\left(\sum_{i=1}^{N} p_i^2 - \frac{\left(\sum_{i=1}^{N} p_i\right)^2}{N}\right)$ je rozptyl medzi p_i. Z toho vyplýva, že pri rôznych individuálnych pravdepodobnostiach p_i je rozptyl menší ako rozptyl za platnosti binomického modelu.

Overdispersion v Poissonovom modeli. Predpokladajme, že náhodná premenná X má rozptyl $Var[X]$ podmienený strednou hodnotou $E[X] = \mu$, kde μ je náhodná premenná so strednou hodnotou $E[\mu]$ a rozptylom $Var[\mu] = \sigma_\mu^2$. Teda pre jednotlivé subjekty μ, charakterizujúca napr. nehodu, varíruje. Hoci počet nehôd na subjekt má rozdelenie $Poiss(\mu)$, marginálne rozdelenie bude charakterizované *overdispersion*, a teda

$$E[X_\mu] = E[E[X_\mu|\mu]] = E[\mu] \text{ a } Var[X_\mu] = E[Var[X_\mu|\mu]] + Var[E[X_\mu|\mu]] = E[\mu] + \sigma_\mu^2,$$

čo poukazuje na väčšiu variabilitu v porovnaní s Poissonovým modelom. Za predpokladu, že μ má **gama rozdelenie**, môžeme ľahko spočítať marginálne pravdepodobnosti, t.j. ak X_μ má Poissonovo rozdelenie so strednou hodnotou μ, μ má hustotu $f(\mu) = \frac{1}{\Gamma(\alpha)}\mu^{\alpha-1}\lambda^\alpha e^{-\lambda\mu}$. Náhodná premenná X predstavujúca zmes X_μ má strednú hodnotu $E[X] = E[\mu] = \frac{\alpha}{\lambda}$ a rozptyl $Var[X] = E[\mu] + Var[\mu] = \frac{\alpha}{\lambda} + \frac{\alpha}{\lambda^2}$. Marginálna pravdepodobnosť pre $x = 0, 1, \ldots,$ je potom rovná

$$\begin{aligned}
\Pr(X = x) &= E[\Pr(X_\mu = x|\mu)] = E[e^{-\mu}\frac{\mu^x}{x!}] = \frac{\lambda^\alpha}{\Gamma(\alpha)x!}\int e^{-\mu}\mu^x\mu^{\alpha-1}e^{-\lambda\mu}d\mu \\
&= \frac{\lambda^\alpha\Gamma(x+\alpha)}{(\lambda+1)^{x+\alpha}\Gamma(\alpha)x!} = \binom{x+\alpha-1}{\alpha-1}\left(\frac{\lambda}{\lambda+1}\right)^\alpha\left(1-\frac{\lambda}{\lambda+1}\right)^x,
\end{aligned}$$

kde $(x + \alpha - 1)! = \Gamma(x + \alpha)$. Ide teda o negatívne binomické rozdelenie, kde X je počet neúspechov (úrazov, zlyhaní) zaznamenaných po α úspechoch, pravdepodobnosť úspechu $\pi = \frac{\lambda}{\lambda+1}$ a pomer zlyhaní $\mu = \frac{1-\pi}{\pi}\alpha$.

Príklad 82 (overdispersion v Poissonovom modeli) *Majme početnosti robotníkov m_n s n úrazmi v továrni (pozri v tabuľku 2.10; Greenwood a Yule (1920)). Vypočítajte očakávané početnosti robotníkov za predpokladu, že početnosti úrazov na robotníka X majú Poissonove rozdelenie s parametrom $\lambda = \frac{\sum_n nm_n}{\sum_n m_n} = 0.47$, ozn. $X \sim Poiss(\lambda)$.*

Tabuľka 2.10: Pozorované početnosti robotníkov m_n s n úrazmi v továrni

n	0	1	2	3	4	≥ 5
pozorované m_n	447	132	42	21	3	2

Riešenie
Keď porovnáme pozorované m_n (pozri tabuľku 2.10) a vypočítané (teoretické, očakávané) m_n (pozri tabuľku 2.11) zistíme, že pozorované poukazujú na *overdispersion*, t.j. máme viac robotníkov bez úrazu ako aj viac robotníkov s väčším množstvom úrazov v porovnaní s teoretickými početnosťami.

Tabuľka 2.11: Očakávané početnosti robotníkov m_n (zaokrúhlené na nula desatinných miest) s n úrazmi v továrni (Poissonovo rozdelenie)

n	0	1	2	3	≥ 4
očakávané m_n	406	189	44	7	1

2.1 *Simulačný experiment ako nástroj štúdia teoretických vlastností modelov

Monte Carlo (MC) experiment. Pojem MC metóda je známy od 40. rokov dvadsiateho storočia a zaviedli ho fyzici pracujúci na projekte o jadrových zbraniach v Los Alamos National Laboratory, menovite Stanislav Ulam, Enrico Fermi, John von Neuman a Nicholas Metropolis. MC je odvodené od kasína v Monaku, kde Ulamov strýko hrával hazardné hry. Použitie tzv. **náhodnosti a opakovateľnosti MC procesu** je analogické aktivitám v kasíne. Prvýkrát metódu použil Enrico Fermi v roku 1930 na výpočet vlastností novoobjaveného neutrónu. MC experiment bol použitý tiež v 50. rokoch 20. storočia počas vývoja vodíkovej bomby. U.S. Air Force bola v tom čase hlavnou organizáciou zodpovednou za financovanie a rozšírenie informácie o MC metódach, ktoré si začali hľadať cestu k mnohým ďalším aplikáciám, najprv vo fyzike, neskôr v chémii a nakoniec aj v matematike a štatistike. V štatistike sa MC metódy používajú na študovanie asymptotických vlastností odhadov a testovacích štatistík (príp. štatistických modelov) a zisťovanie ich správania sa za kontrolovaných podmienok (pozri napr. Rizzo, 2007; Gentle, 2009). Jediný predpoklad **dobrej simulácie pseudonáhodných čísel** používaný v MC metódach je **dostatočná náhodnosť v širšom slova zmysle** (Suess a Trumbo, 2010).

Simulačný experiment (mnohonásobne opakované náhodné výbery) musí spĺňať nasledovné tri kritériá (Robert a Casella, 2010)

1. **relevantnosť** – vygenerované (simulované) dáta musia byť generované na základe relevantných pravidiel, napr. kombinácia minulých skúseností a súčasných dát, *hypotetického pravdepodobnostného* alebo *štatistického modelu*, ktorý chceme študovať a pod.;

2. **stabilita** – *centrálna limitná veta* (CLV; použiteľná aj pre n menšie ako 100) a *dva zákony veľkých čísel* (použiteľné pre n väčšie ako 100 alebo 1000) garantujú, že ak je plán simulačnej štúdie správny a simulácia má dostatočne veľa opakovaní, dostaneme stabilný výsledok namiesto náhodného šumu;

3. **diagnostika** – pomocou rôznych numerických a grafických metód môžeme rozlíšiť signál od šumu, napr. porovnanie numerického a analytického riešenia, porovnanie viacerých podobných modelov, použitie dostatočného množstva opakovaní a pod.

Veta 1 (CLV) *Nech X_i, $i = 1, 2, \ldots, n$, sú iid náhodné premenné s rovnakou strednou hodnotou $E[X_i] = \mu \in \mathbb{R}$ a konečným rozptylom $Var[X_i] = \sigma^2 < \infty$. Potom $\frac{1}{\sigma\sqrt{n}}(\sum_{i=1}^{n} X_i - n\mu)$ má limitne (pre dostatočne veľké n) štandardizované normálne rozdelenie (t.j. konverguje v distribúcii k štandardizovanému normálnemu rozdeleniu $N(0,1)$; Wasserman (2006).*

Veta 2 (Slabý zákon veľkých čísel) *Nech X_i, $i = 1, 2, \ldots, n$, sú iid náhodné premenné s rovnakou strednou hodnotou $E[X_i] = \mu \in \mathbb{R}$ a konečným rozptylom $Var[X_i] = \sigma^2 < \infty$. Nech $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Potom pre každé $\epsilon > 0$ platí $\lim_{n\to\infty} \Pr(|\overline{X}_n - \mu| < \epsilon) = 1$, t.j. pre dostatočne veľké n platí, že $\overline{X}_n$ konverguje v pravdepodobnosti k μ (Lehmann, 1999).*

Veta 3 (Silný zákon veľkých čísel) *Nech X_i, $i = 1, 2, \ldots, n$, sú iid náhodné premenné s rovnakou strednou hodnotou $E[X_i] = \mu \in \mathbb{R}$ a konečným rozptylom $Var[X_i] = \sigma^2 < \infty$. Nech $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Potom pre každé $\epsilon > 0$ platí $\Pr(\lim_{n\to\infty} |\overline{X}_n - \mu| < \epsilon) = 1$, t.j. pre dostatočne veľké n platí, že $\overline{X}_n$ konverguje skoro iste k μ (Lehmann, 1999).*

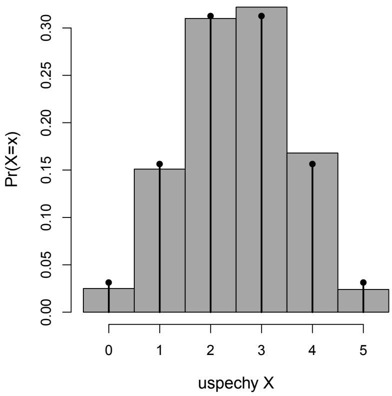
CLV a zákony veľkých čísel nám v praxi zabezpečia použitie nejakého modelu rozdelenia pravdepodobnosti na reálne dáta za predpokladu, že máme dostatočne veľký rozsah náhodného výberu (DasGupta, 2008). Dôsledky použitia modelu rozdelenia pravdepodobnosti na reálne dáta s malým rozsahom siahajú od nesprávneho použitia štatistického modelu po nesprávne použitie štatistického testu, čoho dôsledkom je nedôveryhodná interpretácia výsledkov štatistickej analýzy.

Príklad 83 (binomický rozdelenie, simulačná štúdia) *Vygenerujte pseudonáhodné čísla X (početnosti úspechov) opakovane M-krát ($M = 1000$) z $Bin(N, p)$, kde $N = 5$ a $p = 0.5$. Vytvorte tabuľku vygenerovaných (simulovaných) ako aj teoretických relatívnych početností (pre $n = 0, 1, \ldots, 5$). Superponujte histogram vygenerovaných pseudonáhodných čísel s teoretickou pravdepodobnostnou funkciou.*

Riešenie (pozri tabuľku 2.12 a obrázok 2.12)

Tabuľka 2.12: Simulované a teoretické relatívne početnosti úspechov

relatívne početnosti/n	0	1	2	3	4	5
simulované	0.025	0.151	0.310	0.322	0.168	0.024
teoretické	0.031	0.156	0.312	0.312	0.156	0.031



Obr. 2.12: Histogram vygenerovaných pseudonáhodných čísel superponovaný spojnicovým grafom teoretickej pravdepodobnostnej funkcie X

Príklad 84 (binomické vs. normálne rozdelenie) *Nech $X_N \sim Bin(N, p)$, potom môžeme aproximovať binomické rozdelenie normálnym nasledovne – $X_N \sim N(Np, Np(1-p))$, kde tiež platí $Z_N = \frac{X_N - Np}{\sqrt{Np(1-p)}} \sim N(0, 1)$. Ukážte, že CLV platí pre $N = 100$ a $p = 1/2$ na tri desatinné miesta.*

Riešenie (aj v R)
Príklad hovorí o tom, ako dobre normálne rozdelenie aproximuje binomické pri rozsahu $N = 100$, čo je dôležité pri testovaní hypotéz.
$E[X_N] = Np = 50$, $\sqrt{Var[X_N]} = \sqrt{Np(1-p)} = 5$.
Ak $Y_N = X_N/N$, potom $\Pr(|Y_N - 1/2| < \epsilon) = 0.236$, kde $\epsilon = 0.02$. $\Pr(0.48 < Y_{100} < 0.52) = \Pr(48 < X_{100} < 52) = \Pr(48.5 < X_{100} < 51.5) = \Pr(\frac{48.5-50}{5} < Z_{100} < \frac{51.5-50}{5})$, kde $X_{100} \sim N(50, 5)$ s použitím úpravy na spojitosť.

```
pbinom(51,100,.5)-pbinom(48,100,.5) # 0.2356466
pnorm(51.5,50,5)-pnorm(48.5,50,5) # 0.2358228
```

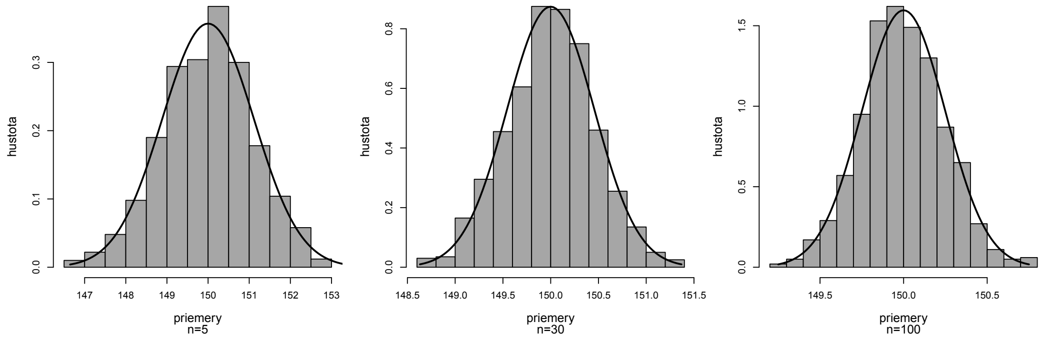
Výsledky sa zhodujú na tri desatinné miesta. Všeobecne platí $X_M \sim N(M/2, M/4)$ a $Y_M = X_M/M \sim N(1/2, 1/(4M))$.

Príklad 85 (normálne rozdelenie, simulačná štúdia) *Na základe simulačnej štúdie preverte, že ak $X \sim N(150, 6.25)$, potom $\overline{X}_n \sim N(150, 6.25/n)$. Použite (1) $n = 5$, (2) $n = 30$ a (3) $n = 100$. Pre každú simuláciu X vypočítajte aritmetické priemery $\overline{x}_m$, $m = 1, 2, \ldots, M$, kde $M = 1000$. Superponujte ich histogram v relatívnej škále s teoretickou krivkou hustoty pre $\overline{X}_n$. Vypočítajte $\Pr(\overline{X}_n > 151)$ pre $n = 30$ zo simulovaných dát a porovnajte tento výsledok s teoretickou (očakávanou) pravdepodobnosťou.*

Riešenie (aj v R) (pozri obrázok 2.13)
Príklad hovorí o tom, že ak má náhodná premenná X_n normálne rozdelenie, bude mať normálne rozdelenie aj aritmetický priemer $\overline{X}_n$, čo je dôležité pri testovaní hypotéz.
$\Pr(\overline{X}_n > 151) = \Pr(\frac{\overline{X}_n - 150}{\sqrt{6.25}/\sqrt{n}} > \frac{151-150}{\sqrt{6.25}/\sqrt{n}}) \approx \Phi(2.190890) = 0.01422987$, kde $n = 30$.

```
M <- 1000; n30 <- 30
X30 <- matrix(0,M,n30)
for (i in 1:M) X30[i,] <- rnorm(n30,150,sqrt(6.25))
x.bar30 <- rowMeans(X30)
# pravdepodobnosti
mean(x.bar30>151) # 0.014238
1-pnorm((151-150)/sqrt(6.25/n30)) # 0.01422987
# obrazok pre n=30
windows(5,5)
par(mar=c(5,4.5,1,1))
hist(x.bar30, probability=TRUE, col="gray", main="", ylab="hustota", xlab="priemery", sub="n=30", cex.lab=1.2,
     cex.sub=1.2)
xmin <- 150-3*sqrt(6.25/n30)
xmax <- 150+3*sqrt(6.25/n30)
curve(dnorm(x,150,sqrt(6.25/n30)),from=xmin, to=xmax, lwd=2, add=TRUE)
```

Pri dostatočne veľkom počte opakovaní vidíme zhodu medzi teoretickým a simulovaným rozdelením $\overline{X}_n$ na tri desatiné miesta (pri výpočte zadanej pravdepodobnosti).



Obr. 2.13: Histogram vygenerovaných priemerov superponovaný teoretickou krivkou hustoty $\overline{X}_n$

Príklad 86 (normálne rozdelenie, simulačná štúdia) *Nech (a)* $X \sim N(\mu, \sigma^2)$, $\mu = 0$, $\sigma^2 = 1$, *(b)* $X \sim Exp(\lambda)$, $\lambda = 1/3$, *(c)* $X \sim Unif(\min, \max), \min = 0, \max = 1$, *(d)* $X \sim [pN(0,1) + (1-p)N(0,10)]$, *kde* $p = 0.9$. *Použite* R *na simuláciu pseudonáhodných čísel z daných rozdelení – rozsahy náhodných výberov* $n = 2, 5, 20, 50, 100$ *a* 500. *Pre každú simuláciu X vypočítajte aritmetické priemery* $\overline{x}_m$, $m = 1, 2, \ldots, M$, *kde* $M = 1000$. *Zobrazte ich do histogramu v relatívnej škále a superponujte ho s teoretickou krivkou hustoty* $N(\mu, \sigma^2/n)$ *prislúchajúcej danej simulácii.*

Príklad 86 slúži na zistenie vlastností rozdelenia výberového priemeru pri rôznych situáciách. $Exp(\lambda)$ je exponenciálne rozdelenie s parametrom λ, $Unif(\min, \max)$ je rovnomerné rozdelenie s parametrami min a max. Zmes dvoch normálnych rozdelení predstavuje 10% prímes normálneho rozdelenia s väčším rozptylom rovným $\sigma^2 = 10$ v normálnom rozdelení s menším rozptylom rovným $\sigma^2 = 1$, čím sme docielili výskyt 10 % odľahlých pozorovaní.

Príklad 87 (normálne rozdelenie, simulačná štúdia) *Nech* $X \sim N(\mu_1, \sigma_1^2)$ *a* $Y \sim N(\mu_2, \sigma_2^2)$. *Potom* $\overline{X}_{n_1} - \overline{Y}_{n_2} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$. *Generujte pseudonáhodné čísla X a Y rozdelení* $N(\mu_j, \sigma_j^2), j = 1, 2$, *kde* $\mu_1 = 100, \sigma_1 = 10, \mu_2 = 50, \sigma_2 = 9$ *pri (a)* $n_1 = 4, n_2 = 5$, *(b)* $n_1 = 100, n_2 = 81$. *Pre každú simuláciu X a Y vypočítajte rozdiel* $\overline{x}_m - \overline{y}_m, m = 1, 2, \ldots, M$, *kde* $M = 1000$. *Superponujte histogram týchto rozdielov v relatívnej škále s teoretickou krivkou hustoty rozdielu* $\overline{X}_{n_1} - \overline{Y}_{n_2}$. *Pre prípad (a) aj (b) vypočítajte* $\Pr(\overline{X}_{n_1} - \overline{Y}_{n_2}) < 52$ *na základe empirického (vygenerovaného) a teoretického rozdelenia* $\overline{X}_{n_1} - \overline{Y}_{n_2}$.

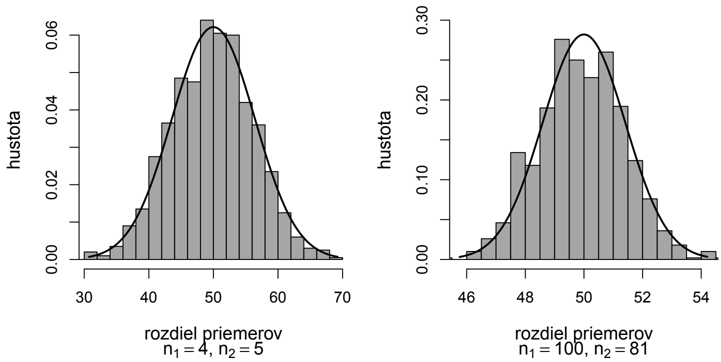
Príklad 87 slúži na zistenie vlastností rozdelenia rozdielu dvoch výberových priemerov pri rôznych situáciách. Pri dostatočne veľkom počte opakovaní vidíme zhodu medzi teoretickým a simulovaným rozdelením $\overline{X}_{n_1} - \overline{Y}_{n_2}$ na dve desatiné miesta (pri výpočte zadanej pravdepodobnosti; pozri obrázok 2.14).

2.2 *Štatistika

Definícia 15 (štatistika) *Ľubovoľná funkcia* $T(\cdot)\colon \mathcal{Y} \to \mathbb{R}^r$, *pre nejaké* $r \in \mathbb{N}^+$ *náhodného výberu* $(X_1, X_2, \ldots, X_n)^T$, *kde funkcia* $T = T(X_1, X_2, \ldots, X_n)$ *nezávislá na* θ *sa nazýva* ***štatistika*** *a hodnota* $t = T(x_1, x_2, \ldots, x_n)$ *korešpondujúca realizáciám* $x_1, x_2, \ldots, x_n$ *sa nazýva* ***realizácia štatistiky (pozorovaná hodnota štatistiky)***.

Príklad 88 (štatistika) *Majme náhodný výber* $(X_1, X_2, \ldots, X_n)^T$, *kde* $X_i \in \mathbb{R}, i = 1, 2, \ldots, n$, *potom príkladmi štatistík sú:* $T_1 = \sum_{i=1}^n X_i \in \mathbb{R}$, $T_2 = \sum_{i=1}^n X_i^2 \in \mathbb{R}^+ \cup \{0\}$, $T_3 = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2) \in \mathbb{R}^2$.

Štatistiky teda môžu byť náhodné premenné alebo náhodné vektory, ktoré sumarizujú informáciu o dátach, zjednodušujú pohľad na ne a umožňujú na ich základe dáta jednoduchšie popísať a ľahšie interpretovať.



Obr. 2.14: Histogram vygenerovaných rozdielov priemerov superponovaný teoretickou krivkou hustoty rozdelenia rozdielu výberových aritmetických priemerov

Príklad 89 (štatistika; príklady) *(a) Vypočítajte štatistiku* $T_1 = \sum_{i=1}^{n} X_i$ *pre realizácie náhodnej premennej* X *najväčšia dĺžka mozgovne (skull.L; mm; dáta: one-sample-mean-skull-mf.txt). V tomto prípade ide o čitateľ aritmetického priemeru* $\overline{x}$*. (b) vypočítajte štatistiku* $T_2 = \sum_{i=1}^{n} X_i^2$ *pre realizácie náhodnej premennej* $X = X_1 - X_2$*, kde* X *predstavuje stranový rozdiel vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej aj ľavej strane tela (length.R a length.L; v mm; dáta: paired-means-clavicle2.txt). V tomto prípade ide o sumu štvorcov stranových rozdielov.*

Definícia 16 (postačujúca štatistika) *Pre štatistický model* $\mathcal{F}$ *je štatistika* $T(x)$ ***postačujúca*** *pre parameter* θ*, ak má rovnakú hodnotu pre dva body rôzne* x_1 *a* x_2 *z výberového priestoru* $\mathcal{Y}$ *iba ak tieto body majú ekvivalentné funkcie vierohodnosti (Azzalini, 1996), t.j. pre*

$$\forall x_1, x_2 \in \mathcal{Y} : T(x_1) = T(x_2) \Rightarrow L(\theta, x_1) \approx L(\theta, x_2) \text{ pre všetky } \theta \in \Theta.$$

Detaily o funkcii vierohodnosti pozri v kapitole 2.3 *Funkcia vierohodnosti. Aj napriek tomu, že nebijektívna transformácia obsahuje „všetku informáciu o dátach“, treba mať na zreteli, že to súvisí s voľbou modelu, t.j. ak sa zmení model, štatistika už nemusí byť postačujúca (Bickel a Doksum, 2006). Preto je vo ľba modelu veľmi dôležitá. Anděl (2011) uvádza inú definíciu postačujúcej štatistiky.

Definícia 17 (postačujúca štatistika) *Štatistika* $T(\mathbf{X})$ *sa nazýva* ***postačujúca štatistika*** *pre parameter* θ*, ak podmienené rozdelenie vektora* $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$ *pri danom* $T(\mathbf{X})$ *nezávisí na* θ*.*

Predchádzajúca definícia hovorí o tom, že ak $T(\mathbf{X})$ je postačujúca štatistika pre θ, potom každá inferencia o θ závisí na $\mathbf{X}$ len cez hodnotu $T(\mathbf{X})$, t.j. ak $\mathbf{x}$ a $\mathbf{y}$ sú dve realizácie a $T(\mathbf{x}) = T(\mathbf{y})$, potom inferencia o θ bude rovnaká nezávisle na tom, či sme pozorovali $\mathbf{X} = \mathbf{x}$ alebo $\mathbf{Y} = \mathbf{y}$.

Veta 4 (postačujúca štatistika) *Ak* $f_1(\mathbf{x}|\theta)$ *je združená hustota* $\mathbf{X}$ *a* $f_2(t|\theta)$ *je hustota* $T(\mathbf{X})$*, potom* $T(\mathbf{X})$ *je postačujúca štatistika pre* θ*, ak pre všetky* $\mathbf{x}$ *je podiel* $f_1(\mathbf{x}|\theta)/f_2(t|\theta)$ *konštanta nezávislá na* θ *(Casella a Berger, 2002).*

Príklad 90 (postačujúca štatistika binomického rozdelenia) *Nech* X_i*,* $i = 1, 2, \ldots, N$*, sú iid Bernoulliho pokusy a* $X = \sum_{i=1}^{N} X_i$*. Potom* $X \sim Bin(N, p)$*. Ukážte, že* $T(\mathbf{X}) = \sum_{i=1}^{N} X_i$ *je postačujúca štatistika pre* p*.*

Riešenie

$f_1(\mathbf{x}|p)/f_2(t|p) = \prod_{i=1}^{N} p^{x_i}(1-p)^{1-x_i} / (\binom{N}{t} p^t (1-p)^{N-t}) = 1/\binom{N}{\sum x_i}$. Tento podiel je nezávislý na p, t.j. súčet jednotiek obsahuje všetku informáciu o p, ktorá je v dátach.

Príklad 91 (postačujúca štatistika normálneho rozdelenia) *Nech $X_i \sim N(\mu, \sigma^2)$, kde $i = 1, 2, \ldots, n$, sú iid premenné a σ^2 poznáme. Ukážte, že $T(\mathbf{X}) = \sum_{i=1}^N X_i/n = \overline{X}$ je postačujúca štatistika pre μ.*

Riešenie
$f_1(\mathbf{x}|\mu)/f_2(t|\mu) = \prod_{i=1}^N f(x_i, \mu)/f_2(t|\mu)$, kde $f(x_i, \mu)$ je hustota normálneho rozdelenia a$f_2(t|\mu) = (2\pi\sigma^2/n)^{-1/2} \exp(-n(\overline{x} - \mu)^2/(2\sigma^2))$. Po viacerých algebraických úpravách dostaneme

$$f_1(\mathbf{x}|\mu)/f_2(t|\mu) = n^{-1/2}(2\pi\sigma^2)^{-(n-1)/2} \exp(-\sum_{i=1}^n (x_i - \overline{x})^2/(2\sigma^2)).$$

Tento podiel je nezávislý na μ, t.j. $\overline{X}$ obsahuje všetku informáciu o μ, ktorá je v dátach.

Dá sa ukázať, že ak $\mathrm{X}_i \sim N(\mu, \sigma^2)$, kde $i = 1, 2, \ldots, n$, sú iid premenné, potom $T_1(\mathbf{X}) = (\overline{X}, \sum_{i=1}^n X_i^2)^T$ a $T_2(\mathbf{X}) = (\overline{X}, \sum_{i=1}^n (X_i - \overline{X})^2)^T$ sú postačujúce štatistiky $\boldsymbol{\theta} = (\mu, \sigma^2)^T$.

Špeciálnym prípadom štatistiky je **testovacia štatistika**, ktorá má kľúčovú úlohu pri testovaní hypotéz.

Príklad 92 (testovacia štatistika, simulačná štúdia) *Na základe simulačnej štúdie preverte, že ak náhodná premenná X má asymptoticky binomické rozdelenie $Bin(N, p)$, potom testovacia štatistika $Z_W = \frac{X/N - p}{\sqrt{(p(1-p)/N}}$, má asymptoticky normálne rozdelenie $N(0, 1)$. Použite $p = 0.1, 0.5, 0.9$, a $N = 5, 10, 30, 50$ a 100. Okomentujte výsledky v spojitosti s Haldovou podmienkou $Np(1-p) > 9$. Pre každú simuláciu X vypočítajte $z_{W,m}$, $m = 1, 2, \ldots, M$, kde $M = 1000$. Superponujte histogram vygenerovaných testovacích štatistík v relatívnej škále s teoretickou krivkou hustoty Z_W.*

Príklad 92 hovorí o použití jednovýberovej testovacej štatistiky pre parameter binomického rozdelenia (pravdepodobnosť) pre rôzne pravdepodobnosti a rôzne početnosti. Ak Haldova podmienka nie je splnená, nie je možné testovaciu štatistiku použiť.

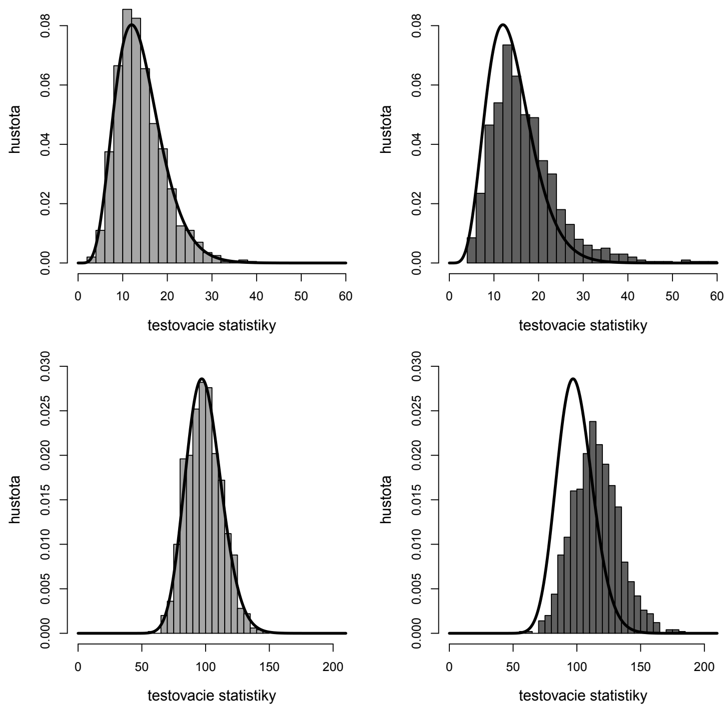
Príklad 93 (testovacia štatistika, simulačná štúdia) *Na základe simulačnej štúdie preverte, že ak (a) $X \sim N(\mu, \sigma^2)$, kde $\mu = 0, \sigma^2 = 1$ a (b) $X \sim [(1-p)N(\mu, \sigma^2) + pN(\mu, \sigma_1^2)]$, kde $p = 0.05$ a $\sigma_1^2 = 2$, potom testovacia štatistika $F = \frac{(n-1)S^2}{\sigma^2}$ má asymptoticky χ^2_{n-1} rozdelenie s $n-1$ stupňami voľnosti. Použite rozsahy náhodných výberov $n = 15$ a $n = 100$. Pre každú simuláciu X vypočítajte $F_{obs,m}$, $m = 1, 2, \ldots, M$, kde $M = 1000$. Superponujte histogram vygenerovaných testovacích štatistík v relatívnej škále s teoretickou krivkou hustoty F.*

Riešenie
Príklad hovorí o použití jednovýberovej testovacej štatistiky pre parameter normálneho rozdelenia (rozptyl) pre rôzne teoretické rozdelenia a rôzne rozsahy náhodných výberov. Ak sú výchylky od normality príliš veľké, nie je možné testovaciu štatistiku použiť. Vieme, že $E[S^2] = \sigma^2 = 1$ a $Var[S^2] = 2\sigma^4/(n-1) = 2/(n-1)$, $E[F] = n-1$ a $Var[F] = 2(n-1)$, t.j. chceme, aby sa výsledky simulačnej štúdie priblížili týmto teoretickým hodnotám (pozri tabuľku 2.13).

Tabuľka 2.13: Teoretické hodnoty stredných hodnôt a rozptylov S^2 a F a ich odhady zo simulačnej štúdie pri $n = 15$ a $n = 100$

odhady počítané pri simulácii	$E[S^2]$	$Var[S^2]$	$E[F]$	$Var[F]$
teoretické hodnoty, $n = 15$	1.0000	0.1429	14.0000	28.0000
$N(\mu, \sigma^2)$, $n = 15$	1.0003	0.1458	14.0039	28.5763
$X \sim [(1-p)N(\mu, \sigma^2) + pN(\mu, \sigma_1^2)]$, $n = 15$	1.2015	0.2943	16.8213	57.6880
teoretické hodnoty, $n = 15$	1.0000	0.0202	99.0000	198.0000
$N(\mu, \sigma^2)$, $n = 100$	0.9985	0.0198	98.8552	193.7022
$X \sim [(1-p)N(\mu, \sigma^2) + pN(\mu, \sigma_1^2)]$, $n = 100$	1.1596	0.0342	114.8005	335.5359

Pri dostatočne veľkom počte opakovaní vidíme zhodu medzi teoretickým a simulovaným rozdelením F, len ak ide o dáta z normálneho rozdelenia (pozri obrázok 2.15).



Obr. 2.15: Histogramy vygenerovaných testovacích štatistík v relatívnej škále superponované s teoretickými krivkami hustoty F; $X \sim N(0,1)$ (ľavý stĺpec) a $X \sim [(1-p)N(\mu,\sigma^2) + pN(\mu,\sigma_1^2)]$ (pravý stĺpec); $n = 15$ (horný riadok), $n = 100$ (dolný riadok)

2.3 *Funkcia vierohodnosti

Funkcia vierohodnosti je najpoužívanejšou funkciou v štatistike. Sumarizuje informácie dostupné z dát v podobe logaritmu, prvej a druhej derivácie. Používa sa nielen pri odhadovaní parametrov rozdelení pravdepodobnosti, ale aj pri testovaní hypotéz a štatistickom modelovaní.

Majme štatistický model $\mathcal{F} = \left\{ f(\cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^k \right\}$. Nech $k = 1$. Ak už bolo x pozorované, hodnota funkcie hustoty $f(\theta, \mathbf{x})$ závisí len od θ. Táto funkcia nám dáva pravdepodobnosť (hustotu) pozorovaní, a priori vo vzťahu k experimentu, ktoré sme predtým pozorovali. Ak chceme porovnať alebo zoradiť $\theta_1, \theta_2 \in \Theta$ podľa dôležitosti, použijeme podiel $f(\mathbf{x}, \theta_1)/f(\mathbf{x}, \theta_2)$, ak existuje. Tento podiel sa nezmení, ak čitateľ a menovateľ vynásobíme nejakou konštantou c nezávislou na θ. Dá sa preto povedať, že $f(\theta, \mathbf{x})$ je vhodná na porovnanie prvkov množiny Θ až na multiplikatívnu konštantu c (Pawitan, 2001).

Definícia 18 (funkcia vierohodnosti) *Pre štatistický model $\mathcal{F}$, na základe ktorého predpokladáme, že $x \in \mathbb{R}$ boli pozorované, použijeme pojem **vierohodnosť (vierohodnostná funkcia)** pre funkciu $\Theta \to \mathbb{R}^+ \cup \{0\}$ definovanú ako (Cox, 2006)*

$$L(\theta, \mathbf{x}) = c(\mathbf{x}) f(\mathbf{x}, \theta),$$

kde $c \in \mathbb{R}$ je nezávislá na θ a $f(\mathbf{x}, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$.

Zápis $L(\theta, \mathbf{x})$ neindikuje závislosť na $\mathbf{x}$, a preto sa často používa zápis $L(\theta|\mathbf{x})$; podobné platí aj pre hustotu, t.j. často sa používa $f(x_i|\theta)$. Tiež je jedno, či píšeme c alebo $c(\mathbf{x})$, keďže vierohodnosť je funkciou θ. V skutočnosti $L(\theta|\mathbf{x})$ definuje triedu funkcií, ktorej prvky sa odlišujú vďaka multiplikatívnej konštante c. Z vyššie uvedenej definície vyplýva, že dva body asociované s proporcionálnymi hustotami determinujú rovnakú vierohodnosť, t.j. ich vierohodnosti sú ekvivalentné. Vierohodnosť ale nie je pravdepodobnosťou. Vierohodnosť je nezáporná a vo väčšine prípadov pozitívna pre všetky θ, preto môžeme definovať **prirodzený logarimus funkcie vierohodnosti** ako (Brazzale a kol., 2007)

$$\ln(L(\theta|\mathbf{x})) = l(\theta|\mathbf{x}) = \ln c + \ln(f(\mathbf{x}|\theta)),$$

kde $l(\theta|\mathbf{x}) = -\infty$ ak $L(\theta|\mathbf{x}) = 0$. V zmysle c ide teda o triedu funkcií.

Definícia 19 (slabý princíp vierohodnosti) *Pre štatistický model $\mathcal{F} = \{f(\cdot;\theta) : \theta \in \Theta\}$ dva body $x_1, x_2 \in \mathbb{R}$, kde platí $L(\theta, x_1) \approx L(\theta, x_2)$, musia viesť k rovnakému záveru (inferenčnému záveru).*

Z definície 19 vyplýva, že všetka informácia o θ je obsiahnutá vo funkcii vierohodnosti pre realizácie x (z nejakého experimentu). Dve funkcie vierohodnosti parametra θ obsahujú rovnakú informáciu o θ, ak sú proporcionálne pre nejaké x (z rovnakých alebo rozdielnych experimentov).

Definícia 20 (silný princíp vierohodnosti) *Majme x_1 z modelu $\mathcal{F}_1 = \{f_1(\cdot;\theta) : \theta \in \Theta\}$ a x_2 z modelu $\mathcal{F}_2 = \{f_2(\cdot;\theta) : \theta \in \Theta\}$, kde $L_1(\theta, x_1) \approx L_2(\theta, x_2)$. Potom body $x_1, x_2 \in \mathbb{R}$ musia viesť k rovnakému záveru (inferenčnému záveru).*

Príklad 94 (princípy vierohodnosti) *Majme binomické rozdelenie (N je fixované a náhodná premenná je počet úspechov) a negatívne binomického rozdelenia (počet úspechov je fixovaný vopred a náhodná premenná je počet zlyhaní pozorovaný pred zastavením sekvencie pokusov). Ak x_1 je počet úspechov a x_2 počet neúspechov a θ pravdepodobnosť úspechu, potom (Azzalini, 1996)*

$$L_1(\theta|x_1, x_2) = \binom{N}{x_1}\theta^{x_1}(1-\theta)^{x_2}, x_1 = 1, 2, \ldots, N; x_2 = N - x_1,$$

a

$$L_2(\theta|x_1, x_2) = \binom{x_1 + x_2 - 1}{x_1}\theta^{x_1}(1-\theta)^{x_2}; x_1 = 1, 2, \ldots; x_2 = 1, 2, \ldots,$$

kde jadro funkcie vierohodnosti pre oba prípady bude rovné $L_1(\theta|x_1, x_2) = L_2(\theta|x_1, x_2) = \theta^{x_1}(1-\theta)^{x_2}$.

Časť funkcie vierohodnoti zahŕňajúca parameter sa nazýva **jadro** (**kernel**). Kedže maximalizácia funkcie je vo vzťahu k parametru, zvyšok (nejaká konštanta) nezavislý na parametri je pri maximalizácii nepotrebný. Jadro funkcie vierohodnosti je často značené rovnako ako samotná funkcia vierohodnosti (ku ktorej je proporcionálne).

Štatistická teória je kompromis rôznych logicky korektných požiadaviek použitých v kombinácii smerujúci k praktickým potrebám. V praxi slabý princíp vierohodnosti platí takmer vždy, ale silný princíp vierohodnosti iba niekedy (Cox a Donnelly, 2011).

Definícia 21 (maximálne vierohodný odhad) *Maximálne vierohodný odhad parametra θ, ozn. $\widehat{\theta}_{ML} = \widehat{\theta}$ (ozn. ML často neuvádzame a nahrádzame ho slovným spojením „MLE θ je rovné $\widehat{\theta}$“ alebo skrátene „MLE $\widehat{\theta}$“) je taká hodnota parametra θ, ktorá maximalizuje funkciu vierohodnosti $L(\theta|x)$; pozri Cox (2006); Lehmann a Casella (1998).*

Ak $X \sim N(\mu, \sigma^2)$, potom maximálne vierohodnými odhadmi parametrov μ a σ^2 sú $\widehat{\mu} = \overline{x}$ a $\widehat{\sigma}^2 = \frac{(n-1)}{n}s^2$. Ak $X \sim Bin(N, p)$, potom maximálne vierohodný odhad p je $\widehat{p} = x/N$.

Definícia 22 (funkcia vierohodnosti binomického rozdelenia a jej jadro) *Nech X je náhodná premenná, ktorá má binomické rozdelenie s parametrami N a p, t.j. $X \sim Bin(N, p)$. Realizácia X je $x = n$. Potom jadro funkcie vierohodnosti má tvar $L(p|x) = p^x(1-p)^{N-x}$ a jeho logaritmus je rovný $l(p|x) = x\ln p + (N-x)\ln(1-p)$; pozri Bickel a Doksum (2006). Binomický koeficient, kombinačné číslo $\binom{N}{x}$, nepíšeme, lebo ho pri maximalizácii nepotrebujeme.*

Definícia 23 (funkcia vierohodnosti multinomického rozdelenia a jej jadro) *Nech* $\mathbf{X}$ *je náhodná premenná, ktorá má (J-rozmerné) multinomické rozdelenie s parametrami* N *a* $\mathbf{p}$*, t.j.* $\mathbf{X} \sim Mult_J(N, \mathbf{p})$*, kde* $\mathbf{X} = (X_1, X_2, \ldots, X_J)^T$*. Realizácia* X_j *je* $x_j = n_j$*. Funkcia vierohodnosti je proporcionálna ku jadru vierohodnosti* $L(\mathbf{p}|\mathbf{x}) = \prod_{j=1}^{J} p_j^{x_j}$ *a jej logaritmus* $l(\mathbf{p}|\mathbf{x}) = \sum_{j=1}^{J} x_j \ln p_j$ *(Casella a Berger, 2002). Konštantu* $N!/\prod_j x_j!$ *nepíšeme, lebo ju pri maximalizácii nepotrebujeme.*

Definícia 24 (funkcia vierohodnosti Poissonovho rozdelenia a jej jadro) *Nech* X *je náhodná premenná, ktorá má Poissonovo rozdelenie s parametrom* λ*, t.j.* $X \sim Poiss(\lambda)$*. Potom jadro vierohodnoti (Casella a Berger, 2002)*

$$L(\lambda|\mathbf{x}) = \lambda^{\sum_{i=1}^{N} x_i} e^{-N\lambda},$$

a jeho logaritmus $l(\lambda|\mathbf{x}) = \sum_{i=1}^{N} x_i \ln\lambda - N\lambda$*. Menovateľ funkcie vierohodnosti* $x_1!x_2!\ldots x_N!$ *nepíšeme, lebo ho pri maximalizácii nepotrebujeme.*

Príklad 95 (maximálne vierohodné odhady; Poissonovo rozdelenie) *Každý rok za posledných päť rokov boli v nejakom meste registrované* 3, 2, 5, 0 *a* 4 *zemetrasenia za rok. Za predpokladu, že počet zemetrasení za rok* X *má Poissonovo rozdelenie s parametrom* λ*, t.j.* $X \sim Poiss(\lambda)$*, odhadnite* λ *(*λ *predstavuje očakávanú početnosť zemetrasení za rok).*

Riešenie
Logaritmus funkcie vierohodnosti $l(\lambda|\mathbf{x}) = \sum_{i=1}^{N} x_i \ln\lambda - N\lambda$, potom $\frac{\partial l(\lambda|\mathbf{x})}{\partial\lambda} = \frac{N\overline{x}}{\lambda} - N$, z čoho vyplýva, že $\widehat{\lambda} = \overline{x}$. Teda ak $N = 5$, vieme vypočítať $\widehat{\lambda} = \frac{\sum x_i}{N} = \overline{x}$, ktorý je rovný 2.8.

Vo všeobecnosti píšeme funkciu vierohodnosti pre Poissonove rozdelenie s parametrom λ a pozorovanými početnosťami m_n ako $L(\lambda|\mathbf{x}) = \prod_n p_n^{m_n}$, kde $p_n = \Pr(X = n) = e^{-\lambda}\lambda^n/n!$ a logaritmus jadra funkcie vierohodnosti ako $l(\lambda|\mathbf{x}) = -\lambda\sum_n m_n + \sum_n n m_n \ln\lambda$. Maximálne vierohodný odhad $\widehat{\lambda} = \frac{\sum_n n m_n}{\sum_n m_n}$.

Maximálne vierohodný odhad zjednodušuje pohľad na funkciu vierohodnosti, pretože číslo predstavujúce odhad je jednoduchšie ako funkcia. Všeobecne však jedno číslo (odhad parametra) nie je postačujúce na to, aby reprezentovalo funkciu vierohodnosti. Ak je funkcia vierohodnosti dobre **aproximovaná nejakou kvadratickou funkciou**, potom potrebujeme na jej opis najmenej dve charakteristiky – **polohu maxima** a **zakrivenie v maxime**. Presnejšie potrebujeme aproximáciu logaritmu funkcie vierodnosti **okolo maximálne vierohodného odhadu** θ **polohy maxima** nejakou kvadratickou funkciou. V tomto prípade nazývame funkciu vierohodnosti *regulárnou.* Prvú deriváciu logaritmu funkcie vierohodnosti nazývame **skóre funkcia** a označujeme ju ako $S(\theta) = \frac{\partial}{\partial\theta} l(\theta|\mathbf{x})$. Z tejto rovnosti je zrejmé, že maximálne vierohodný odhad $\widehat{\theta}$ je riešením **vierohodnostných** (**skóre**) **rovníc** $S(\theta) = 0$. V maxime je druhá derivácie logaritmu funkcie vierohodnosti záporná a zakrivenie v bode $\widehat{\theta}$ bude rovné **Fisherovej miere informácie** $\mathcal{I}(\widehat{\theta})$, kde $\mathcal{I}(\widehat{\theta}) = -\frac{\partial^2}{\partial\theta^2} l(\theta|\mathbf{x})|_{\theta=\widehat{\theta}}$. Veľké zakrivenie zodpovedá strmému a úzkemu vrcholu, čo indikuje menšiu neistotu o θ. $\mathcal{I}(\widehat{\theta})$ nazývame **pozorovaná Fisherova miera informácie. Maximálne vierohodný odhad rozptylu** odhadu parametra θ potom môžeme definovať ako $\widehat{Var[\widehat{\theta}]} = 1/\mathcal{I}(\widehat{\theta})$. **Očakávaná Fisherova miera informácie** je definovaná ako $I(\theta) = E[S^2(\theta)] = Var[S(\theta)] = -E[\frac{\partial}{\partial\theta} S(\theta)]$. Keďže $X_i, i = 1, 2, \ldots, n$, sú nezávislé, potom platí $I(\theta) = ni(\theta)$, kde $i(\theta)$ je Fisherova miera informácie jedného pozorovania.

Príklad 96 ($\mathcal{I}(\widehat{p})$ **a rozptyl pre** p**;** $X \sim Bin(N, p)$**)** *Z funkcie vierohodnosti odvoďte pozorovanú Fisherovu mieru informácie* $\mathcal{I}(\widehat{p})$ *a rozptyl* $\widehat{Var[\widehat{p}]}$*.*

Riešenie
$l(p|x) = x\ln p + (N - x)\ln(1 - p) = N\widehat{p}\ln p + N(1 - \widehat{p})\ln(1 - p)$, potom
$\frac{\partial^2}{\partial p^2} l(p|x) = -(N\widehat{p})/p^2 - N(1 - \widehat{p})/(1 - p)^2$.
Ak dosadíme $p = \widehat{p}$, dostaneme
$\frac{\partial^2}{\partial p^2} l(p|x)|_{p=\widehat{p}} = -(N\widehat{p})/\widehat{p}^2 - N(1 - \widehat{p})/(1 - \widehat{p})^2 = [-N(1 - \widehat{p}) - N\widehat{p}]/[\widehat{p}(1 - \widehat{p})]$
$= -N/[\widehat{p}(1 - \widehat{p})]$, z čoho vyplýva, že

$$\frac{1}{\mathcal{I}(\widehat{p})} = \widehat{Var[\widehat{p}]} = \frac{\widehat{p}(1-\widehat{p})}{N}.$$

Príklad 97 ($\mathcal{I}(\widehat{\lambda})$ a rozptyl pre λ; $X \sim Poiss(\lambda)$) *Každý rok za posledných päť rokov boli v nejakom meste registrované 3, 2, 5, 0 a 4 zemetrasenia za rok. Za predpokladu, že počet zemetrasení za rok $X \sim Poiss(\lambda)$, odhadnite rozptyl parametra λ a vypočítajte hodnotu tohoto odhadu rozptylu pre počet zemetrasení.*

Riešenie
$\frac{\partial^2}{\partial \lambda^2} l(\lambda|\mathbf{x}) = -\frac{N\overline{x}}{\lambda^2}$, z čoho po dosadení $\lambda = \widehat{\lambda}$ dostaneme $\frac{1}{\mathcal{I}(\widehat{\lambda})} = \widehat{Var[\widehat{\lambda}]} = \frac{\overline{x}}{N}$, ktorý je rovný 0.56.

Ako vhodný spôsob aproximácie logaritmu funkcie vierohodnosti $l(\theta|\mathbf{x})$ a $l(\boldsymbol{\theta}|\mathbf{x})$ a nejakej funkcie $g(\theta)$ a $g(\boldsymbol{\theta})$, sa javí jednorozmerný a mnohorozmerný **Taylorov rozvoj r-tého rádu**. Táto aproximácia je dostatočne dobrá z hľadiska konvergencie ku funkcii, ktorú aproximujeme.

Definícia 25 (Taylorov rozvoj r-tého rádu) *Ak existujú r-té derivácie funkcie $g(x)$, ozn. $g^{(r)}(x) = \frac{\partial^r}{\partial x^r} g(x)$, potom definujeme Taylorov rozvoj r-tého rádu pre nejakú konštantu a nasledovne (Casella a Berger, 2002)*

$$T_r(x) = \sum_{j=0}^{r} \frac{g^{(j)}(a)}{j!}(x-a)^j.$$

Pri praktických aplikáciách budeme predpokladať, že zvyšok $g(x) - T_r(x)$ bude rýchlo konvergovať k nule (pozri nasledujúcu vetu). Explicitná forma zvyšku nebude dôležitá a budeme ho zanedbávať, pretože nás bude zaujímať iba samotná aproximácia. Jedna z možných podôb zvyšku je nasledovná

$$g(x) - T_r(x) = \int_a^x \frac{g^{(r+1)}(t)}{r!}(x-t)^r dt.$$

Veta 5 (Taylorova veta) *Ak derivácia $g^{(r)}(a) = \frac{\partial r}{\partial x^r} g(x)|_{x=a}$ existuje, potom (Casella a Berger, 2002)*

$$\lim_{x \to a} \frac{g(x) - T_r(x)}{(x-a)^r} = 0.$$

V štatistických aplikáciách Taylorovej vety budeme používať Taylorov rozvoj prvého alebo druhého rádu. Rovnako budeme používať aj mnohorozmerné rozšírenie Taylorovej vety.

Majme **kvadratickú aproximáciu logaritmu funkcie vierohodnosti** pomocou Taylorovho rozvoja druhého rádu okolo $\widehat{\theta}$ definovanú ako

$$l(\theta|\mathbf{x}) \approx l(\widehat{\theta}|\mathbf{x}) + S(\widehat{\theta})(\theta - \widehat{\theta}) - \frac{1}{2}\mathcal{I}(\widehat{\theta})(\theta - \widehat{\theta})^2,$$

z ktorej dostaneme **aproximáciu logaritmu relatívnej (štandardizovanej) funkcie vierohodnosti**

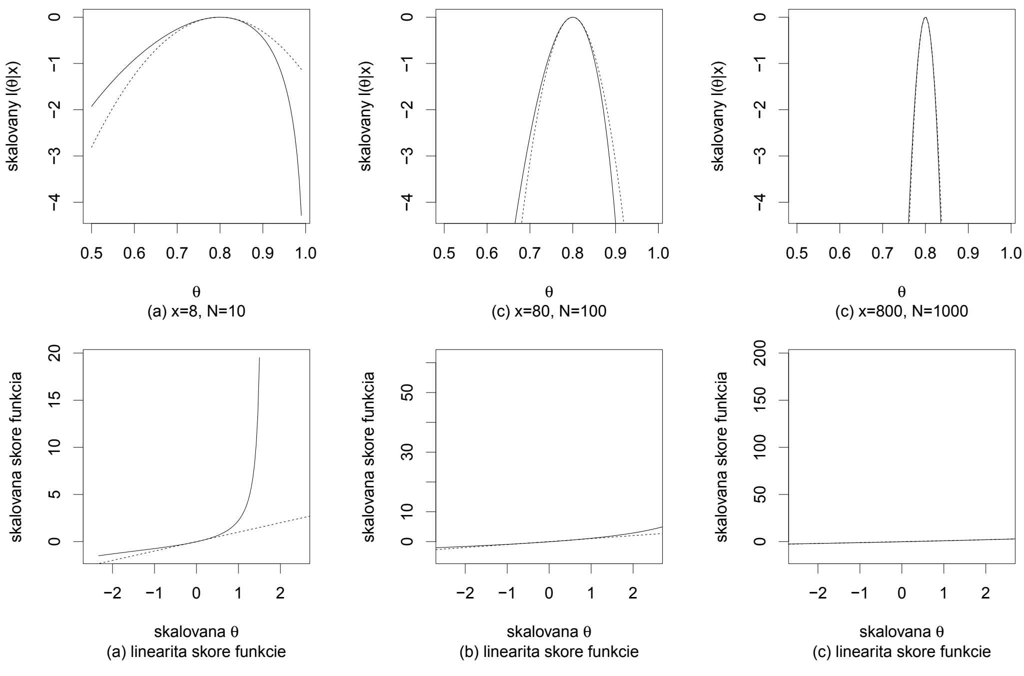
$$\ln \mathcal{L}(\theta|\mathbf{x}) = \ln \frac{L(\theta|\mathbf{x})}{L(\widehat{\theta}|\mathbf{x})} = l(\theta|\mathbf{x}) - l(\widehat{\theta}|\mathbf{x}) \approx -\frac{1}{2}\mathcal{I}(\widehat{\theta})(\theta - \widehat{\theta})^2.$$

Posledná rovnosť predstavuje kvadratickú aproximáciu normalizovaného logaritmu funkcie vierohodnosti okolo $\widehat{\theta}$. Na porovnanie skutočnej funkcie vierohodnosti a jej kvadratickej aproximácie tieto dve funkcie nakreslíme do jedného obrázka. Pri zobrazovaní fixujeme maximum logaritmu funkcie vierohodnosti do nuly a rozsah stanovíme od -4 do 0.

Praktické pravidlo – dostatočne regulárna funkcia vierohodnosti indikuje normalitu X. Alternatívne môžeme zobrať deriváciu kvadratickej aproximácie, kde dostaneme $S(\theta) \approx -\mathcal{I}(\widehat{\theta})(\theta - \widehat{\theta})$ alebo $-\mathcal{I}^{-1/2}(\widehat{\theta})$

$S(\theta) \approx \mathcal{I}^{1/2}(\widehat{\theta})(\theta - \widehat{\theta})$. Posledná rovnosť nie je závislá na mierke θ. Potom kvadratickú aproximáciu môžeme preveriť graficky zobrazením $-\mathcal{I}^{-1/2}(\widehat{\theta})S(\theta)$ oproti $\mathcal{I}^{1/2}(\widehat{\theta})(\theta - \widehat{\theta})$. Za platnosti kvadratickej aproximácie je grafom priamka s jednotkovým sklonom. Pre normálne rozdelené dáta to musí platiť presne. Keďže každá funkcia je lokálne lineárna, je potrebné preveriť, na akom intervale linearitu očakávame. V ideálnom prípade $\mathcal{I}^{1/2}(\widehat{\theta})(\theta - \widehat{\theta}) \sim N(0,1)$, preto kontrolu urobíme aspoň na intervale $\langle -2, 2\rangle$.

Príklad 98 (kvadratická aproximácia funkcie vierohodnosti) *(1) Nakreslite škálovaný logaritmus funkcie vierohodnosti binomického rozdelenia. Na osi x bude p a na osi y* $\ln \mathcal{L}(p|\mathbf{x}) = l(p|\mathbf{x}) - \max(l(p|\mathbf{x}))$. *Porovnajte* $\ln \mathcal{L}(p|\mathbf{x})$ *s kvadratickou aproximáciou vypočítanou pomocou Taylorovho rozvoja* $\ln \mathcal{L}(p|\mathbf{x}) = \ln(\frac{L(p|\mathbf{x})}{L(\widehat{p}|\mathbf{x})}) \approx -\frac{1}{2}\mathcal{I}(\widehat{p})(p-\widehat{p})^2$. *(2) Nech skóre funkcia* $S(p) = \frac{\partial}{\partial p}\ln L(p|\mathbf{x})$. *Keď zoberieme deriváciu kvadratickej aproximácie uvedenej vyššie, dostaneme* $S(p) \approx -\mathcal{I}(\widehat{p})(p-\widehat{p})$ *alebo* $-\mathcal{I}^{-1/2}(\widehat{p})S(p) \approx \mathcal{I}^{1/2}(\widehat{p})(p-\widehat{p})$. *Potom zobrazením pravej strany na osi x a ľavej strany na osi y dostaneme asymptoticky lineárnu funkciu s jednotkovým sklonom. Asymptoticky tiež platí* $\mathcal{I}^{1/2}(X/N)(p - X/N) \overset{\mathcal{D}}{\sim} N(0,1)$. *Je postačujúce mať rozsah osi x rovný* $\langle -2, 2\rangle$, *pretože funkcia je asymptoticky (lokálne) lineárna na tomto intervale. Rozumne škálujte os y. Zobrazte pre (a)* $n = 8, N = 10$, *(b)* $n = 80, N = 100$ *a (c)* $n = 800, N = 1000$ *(*$p \in (0.5, 0.99)$*). Okomentujte rozdiely medzi (a), (b) a (c). Grafické riešenie je na obrázku 2.16.*



Obr. 2.16: Porovnanie škálovaného logaritmu funkcie vierohodnosti (plná čiara) s jeho kvadratickou aproximáciou (čiarkovaná čiara) v prvom riadku a porovnanie škálovanej skóre funkcie a priamky s nulovým interceptom a jednotkovým sklonom v druhom riadku

Ak je funkcia vierohodnosti viacrozmerná, je problém ju zobraziť. Ak je $\boldsymbol{\theta}$ dvojrozmerný vektor, potom môžeme $L(\boldsymbol{\theta}|\mathbf{x})$ zobraziť ako kontúrový graf alebo perspektívny trojrozmerný graf v podobe plochy. Nech $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)^T$. Za predpokladu diferencovateľnosti $L(\boldsymbol{\theta}|\mathbf{x})$ je **skóre funkcia** definovaná ako vektor $S(\boldsymbol{\theta})$, ktorého jednotlivé členy $S(\boldsymbol{\theta}) = \frac{\partial}{\partial \theta_i} l(\boldsymbol{\theta}|\mathbf{x}), i = 1, 2, \ldots, k$, a maximálne vierohodný odhad θ_i je riešením vierohodnostných rovníc $S(\boldsymbol{\theta}) = 0$. **Pozorovaná Fisherova informačná matica** druhých derivácií $l(\boldsymbol{\theta}|\mathbf{x})$ má tvar $\mathcal{I}_{ij}(\widehat{\boldsymbol{\theta}})$ je rovný $-\frac{\partial^2}{\partial\theta_i\partial\theta_j} l(\boldsymbol{\theta}|\mathbf{x})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$. Maximálne vierohodný odhad kovariančnej matice $\widehat{Var[\widehat{\boldsymbol{\theta}}]} = \mathcal{I}^{-1}(\widehat{\boldsymbol{\theta}})$, t.j. odhad kovariančnej matice $\widehat{Var[\widehat{\boldsymbol{\theta}}]}$ je rovný inverzii pozorovanej Fisherovej informačnej matice $\mathcal{I}(\widehat{\boldsymbol{\theta}})$. **Očakávaná Fisherova informačná matica** je definovaná ako $I(\boldsymbol{\theta}) = E[S(\boldsymbol{\theta})(S(\boldsymbol{\theta}))^T] = Var[S(\boldsymbol{\theta})] = -E[\frac{\partial}{\partial\boldsymbol{\theta}}S(\boldsymbol{\theta})]$. Keďže $X_i, i = 1, 2, \ldots, n$, sú nezávislé, potom platí $I(\boldsymbol{\theta}) = ni(\boldsymbol{\theta})$, kde $i(\boldsymbol{\theta})$ je Fisherova miera informácie jedného pozorovania.

Kvadratická aproximácia logaritmu funkcie vierohodnosti pomocou Taylorovho rozvoja druhého rádu okolo $\widehat{\boldsymbol{\theta}}$ je definovaná ako

$$l(\boldsymbol{\theta}|\mathbf{x}) \approx l(\widehat{\boldsymbol{\theta}}|\mathbf{x}) + S(\widehat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}) - \frac{1}{2}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})^T\mathcal{I}(\widehat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}).$$

Pre normálne rozdelené X platí

$$\ln \mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) = \ln \frac{L(\boldsymbol{\theta}|\mathbf{x})}{L(\widehat{\boldsymbol{\theta}}|\mathbf{x})} = l(\boldsymbol{\theta}|\mathbf{x}) - l(\widehat{\boldsymbol{\theta}}|\mathbf{x}) \approx -\frac{1}{2}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})^T\mathcal{I}(\widehat{\boldsymbol{\theta}})(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}),$$

t.j. funkcia vierohodnosti a jej kvadratická aproximácia sú identické.

Definícia 26 (funkcia vierohodnosti normálneho rozdelenia) *Nech $X_1, X_2, \ldots, X_n$ sú nezávislé rovnako rozdelené premenné, $X_i \sim N(\mu, \sigma^2)$, kde $i = 1, 2, \ldots, n$, $\boldsymbol{\theta} = (\mu, \sigma^2)^T \in \Theta = \mathbb{R} \times \mathbb{R}^+$. Vďaka nezávislosti X_i dostaneme*

$$\begin{aligned} L(\boldsymbol{\theta}|\mathbf{x}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right) \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right)\right) \end{aligned}$$

a korešpondujúci logaritmus

$$l(\boldsymbol{\theta}|\mathbf{x}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right).$$

Príklad 99 ($\mathcal{I}(\widehat{\boldsymbol{\theta}})$ pre vektor $\boldsymbol{\theta} = (\mu, \sigma^2)^T$; $X_i \sim N(\mu, \sigma^2)$) *Nech $X_i \sim N(\mu, \sigma^2)$, kde $i = 1, 2, \ldots, n$. Čomu je rovná pozorovaná Fisherova informačná matica $\mathcal{I}(\widehat{\boldsymbol{\theta}})$, kde $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}, \widehat{\sigma}^2)^T$?*

Riešenie
Logaritmus funkcie vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{x}) = -\frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Derivácie funkcie vierohodnosti v μ a σ^2 budú nasledovné

$$S_1(\boldsymbol{\theta}) = \frac{\partial}{\partial\mu}l(\boldsymbol{\theta}|\mathbf{x}) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu),$$

$$S_2(\boldsymbol{\theta}) = \frac{\partial}{\partial\sigma^2}l(\boldsymbol{\theta}|\mathbf{x}) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Potom

$$\mathcal{I}(\widehat{\boldsymbol{\theta}}) = \begin{pmatrix} \frac{n}{\widehat{\sigma}^2} & 0 \\ 0 & \frac{n}{2\widehat{\sigma}^4} \end{pmatrix}.$$

Príklad 100 ($\mathcal{I}(\widehat{\mathbf{p}})$ a rozptyl pre p; $\mathbf{X} \sim Mult_J(N, \mathbf{p})$) *Z funkcie vierohodnosti odvoďte pozorovanú Fisherovu informačnú maticu $\mathcal{I}(\widehat{\mathbf{p}})$ a kovariančnú maticu $\widehat{Var[\widehat{\mathbf{p}}]}$.*

Riešenie
Označme $p_J = 1 - \sum_{j=1}^{J-1} p_j$ a predefinujme J-rozmerný vektor $\mathbf{p}$ na $(J-1)$-rozmerný vektor $\mathbf{p} = (p_1, p_2, \dots, p_{J-1})^T$. Potom $l(\mathbf{p}|\mathbf{x}) = \sum_{j=1}^{J-1} n_j \ln p_j + n_J \ln(1 - \sum_{j=1}^{J-1} p_j)$ a $\frac{\partial}{\partial p_j} l(\mathbf{p}|\mathbf{x}) = \frac{n_j}{p_j} - \frac{n_J}{p_J}$, ktoré tvoria $S(\mathbf{p})$. Fisherovu informačnú maticu dostaneme nasledovne

$$\mathcal{I}(\mathbf{p}) = -\frac{\partial}{\partial \mathbf{p}} S(\mathbf{p}) = \operatorname{diag}\left(\frac{n_1}{p_1^2}, \frac{n_2}{p_2^2}, \dots, \frac{n_{J-1}}{p_{J-1}^2}\right) + \frac{n_J}{p_J^2}\mathbf{1}\mathbf{1}^T,$$

kde $\mathbf{1}$ je $(J-1)$-rozmerný vektor jednotiek. Potom pozorovaná Fisherova informačná matica bude rovná

$$\mathcal{I}(\widehat{\mathbf{p}}) = N\left(\operatorname{diag}\left(\frac{1}{\widehat{p}_1}, \frac{1}{\widehat{p}_2}, \dots, \frac{1}{\widehat{p}_{J-1}}\right) + \frac{\mathbf{1}\mathbf{1}^T}{\widehat{p}_J}\right) = N \begin{pmatrix} \frac{1}{\widehat{p}_1} + \frac{1}{\widehat{p}_J} & \frac{1}{\widehat{p}_J} & \frac{1}{\widehat{p}_J} & \cdots & \frac{1}{\widehat{p}_J} \\ \frac{1}{\widehat{p}_J} & \frac{1}{\widehat{p}_2} + \frac{1}{\widehat{p}_J} & \frac{1}{\widehat{p}_J} & \cdots & \frac{1}{\widehat{p}_J} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{\widehat{p}_J} & \frac{1}{\widehat{p}_J} & \cdots & \frac{1}{\widehat{p}_J} & \frac{1}{\widehat{p}_{J-1}} + \frac{1}{\widehat{p}_J} \end{pmatrix}$$

a

$$\widehat{Var}[\widehat{\mathbf{p}}] = \mathcal{I}^{-1}(\widehat{\mathbf{p}}) = \frac{1}{N}\left(\operatorname{diag}(\widehat{\mathbf{p}}) - \widehat{\mathbf{p}}\widehat{\mathbf{p}}^T\right) = \frac{1}{N} \begin{pmatrix} \widehat{p}_1(1-\widehat{p}_1) & -\widehat{p}_1\widehat{p}_2 & \cdots & -\widehat{p}_1\widehat{p}_{J-1} \\ -\widehat{p}_2\widehat{p}_1 & \widehat{p}_2(1-\widehat{p}_2) & \cdots & -\widehat{p}_2\widehat{p}_{J-1} \\ \vdots & \vdots & \vdots & \vdots \\ -\widehat{p}_{J-1}\widehat{p}_1 & -\widehat{p}_{J-1}\widehat{p}_2 & \cdots & \widehat{p}_{J-1}(1-\widehat{p}_{J-1}) \end{pmatrix}.$$

Ak do $\widehat{Var}[\widehat{\mathbf{p}}]$ pridáme jeden riadok a jeden stĺpec zodpovedajúce $\widehat{p}_J$, dostaneme singulárnu kovariančnú maticu J-rozmerného vektora $\widehat{\mathbf{p}}$.

Profilová vierohodnosť. Aj napriek tomu, že funkcia vierohodnosti je často viacrozmerná, je jednoduchšie ju zobrazovať pre každý parameter θ_i zvlášť alebo pre nejakú podmnožinu parametrov vektora $\boldsymbol{\theta}$. Napr. pri modeli normálneho rozdelenia nás zaujíma len stredná hodnota μ, pričom rozptyl σ^2 je tzv. rušivý parameter (potrebný kvôli adaptácii modelu na variabilitu v dátach). Potrebujeme teda metódu, ktorá koncentruje vierohodnosť len na parameter záujmu eliminovaním rušivého parametra. Vierohodnostný prístup na elimináciu rušivého parametra pozostáva zo substitúcie jeho maximálne vierohodným odhadom v každej fixovanej hodnote parametra záujmu. Výsledkom je **profilová vierohodnostná funkcia**.

Zakrivenie profilovej vierohodnosti. Zakrivenie profilovej funkcie vierohodnosti súvisí s Fisherovou informačnou maticou. Ak napr. parametrom záujmu je θ_1 z vektora $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$, potrebujeme $\mathcal{I}(\widehat{\boldsymbol{\theta}})$ a jej inverziu $\mathcal{I}^{-1}(\widehat{\boldsymbol{\theta}})$ v nasledovnom tvare

$$\mathcal{I}(\widehat{\boldsymbol{\theta}}) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}, \mathcal{I}^{-1}(\widehat{\boldsymbol{\theta}}) = \begin{pmatrix} I^{11} & I^{12} \\ I^{21} & I^{22} \end{pmatrix}.$$

Potom zakrivenie profilovej funkcie vierohodnosti v $\widehat{\theta}_1$ nie je I_{11} ale $(I^{11})^{-1}$, kde $(I^{11})^{-1}$ je vo všeobecnosti menšie ako I_{11}. Interpretácia je nasledovná – informačné číslo I_{11} je zakrivením profilovej funkcie vierohodnosti v $\widehat{\theta}_1$, kde o θ_2 sa predpokladá, že je známe v pozorovanom odhade $\widehat{\theta}_2$; avšak $(I^{11})^{-1}$ je zakrivením profilovej funkcie vierohodnosti v $\widehat{\theta}_1$, ktoré berie do úvahy, že θ_2 je neznáme. Z toho je potom zrejmé, že $(I^{11})^{-1}$ je menšie ako I_{11}. Na základe vyššie uvedeného môžeme kvadraticky aproximovať logaritmus profilovej funkcie vierohodnosti použitím $\widehat{\theta}_i$ a $(I^{ii})^{-1}$, kde

$$\mathcal{L}(\theta_i|\mathbf{x}) = \ln\frac{L(\theta_i|\mathbf{x})}{L(\widehat{\theta}_i|\mathbf{x})} = l(\theta_i|\mathbf{x}) - l(\widehat{\theta}_i|\mathbf{x}) \approx -\frac{1}{2}(I^{ii})^{-1}(\theta_i - \widehat{\theta}_i)^2.$$

Podobným spôsobom je možné kvadraticky aproximovať aj plochu vierohodnosti.

Invariantnosť maximálne vierohodného odhadu. Invariantnosť maximálne vierohodného odhadu znamená, že ak $\widehat{\theta}$ je maximálne vierohodný odhad θ (t.j. $\widehat{\theta} = \widehat{\theta}_{ML}$) a $g(\theta)$ je funkciou θ, potom $g(\widehat{\theta})$ je tiež maximálne vierohodný odhad $g(\theta)$. Maximálne vierohodný odhad rozptylu $g(\theta)$ potom môžeme definovať ako $\widehat{Var[g(\widehat{\theta})]} = [\frac{\partial}{\partial\theta}g(\theta)|_{\theta=\widehat{\theta}}]^2/\mathcal{I}(\widehat{\theta})$. V prípade vektora $\boldsymbol{\theta}$ je $\widehat{Var[\mathbf{g}(\widehat{\boldsymbol{\theta}})]} = \Delta^T\mathcal{I}^{-1}(\widehat{\boldsymbol{\theta}})\Delta$, kde $\Delta = \frac{\partial}{\partial\boldsymbol{\theta}}\mathbf{g}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$. Tento vzorec vychádza z **delta metódy**, ktorá je postavená na Taylorovom rozvoji prvého rádu (t.j. jeho lineárnej zložky) okolo bodu $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}$, kde

$$\mathbf{g}(\widehat{\boldsymbol{\theta}}) \approx \mathbf{g}(\boldsymbol{\theta}) + \sum_{i=1}^{k}\left(\frac{\partial}{\partial\theta_i}\mathbf{g}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right)(\widehat{\theta}_i - \theta_i) = \mathbf{g}(\boldsymbol{\theta}) + \Delta^T(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}),$$

potom

$$\widehat{Var[\mathbf{g}(\widehat{\boldsymbol{\theta}})]} \approx \sum_{i=1}^{k}\left(\frac{\partial}{\partial\theta_i}\mathbf{g}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right)^2\widehat{\sigma}_i^2 + 2\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\left(\frac{\partial}{\partial\theta_i}\mathbf{g}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right)\left(\frac{\partial}{\partial\theta_j}\mathbf{g}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right)\widehat{\sigma}_{ij} = \Delta^T\widehat{\boldsymbol{\Sigma}}\Delta,$$

kde $\widehat{\boldsymbol{\Sigma}} = \widehat{Var[\boldsymbol{\theta}]}$, $\widehat{\sigma}_i^2 = \widehat{Var[\widehat{\theta}_i]}$, $\widehat{\sigma}_{ij} = \widehat{Cov[\widehat{\theta}_i, \widehat{\theta}_j]}$ a $i \neq j; i, j = 1, 2, \ldots, k$. Δ je matica $k \times k_1$, kde derivovanie prebieha po zložkách, t.j. (i,j)-ty element Δ je rovný $\frac{\partial g_j(\boldsymbol{\theta})}{\partial\theta_i}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$, $\mathbf{g}(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), \ldots, g_{k_1}(\boldsymbol{\theta}))^T$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, k_1 \leq k$. V praxi sa často vyskytuje situácia $k \neq 1$ a $k_1 = 1$, napr. $\boldsymbol{\theta} = (p_1, p_2)^T$ a $g(\boldsymbol{\theta}) = \frac{p_1}{p_2}$. Ak $k = k_1 = 1$, potom

$$\widehat{Var[g_1(\widehat{\theta}_1)]} = \widehat{Var[g(\widehat{\theta})]} = \left(\frac{\partial}{\partial\theta}g(\theta)|_{\theta=\widehat{\theta}}\right)^2\widehat{\sigma}^2,$$

kde $\widehat{\sigma}^2 = \widehat{Var[\widehat{\theta}]}$.

Príklad 101 (profilová vierohodnosť; normálne rozdelenie) *Profilová funkcia vierohodnosti pre μ je rovná $L(\mu|\mathbf{x}) = c\exp\left(-n\widehat{\sigma}_\mu^2/(2\widehat{\sigma}^2)\right)$, kde c je nejaká konštanta a $\widehat{\sigma}_\mu^2 = \frac{1}{n}\sum_{i=1}^n(x_i - \mu)^2$, t.j. ide o rez $L((\mu,\sigma^2)^T|\mathbf{x})$ v bode $\sigma^2 = \widehat{\sigma}^2$. Profilová funkcia vierohodnosti pre σ^2 je rovná $L(\sigma^2|\mathbf{x}) = c\left(\sigma^2\right)^{-n/2}\exp\left(-n\widehat{\sigma}^2/(2\sigma^2)\right)$, kde c je nejaká konštanta a $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n(x_i - \overline{x})^2$.*

Príklad 102 (kvadratická aproximácia profilovej funkcie vierohodnosti) *(1) Nakreslite škálovaný logaritmus profilovej funkcie vierohodnosti normálneho rozdelenia pre μ. Na osi x bude μ a na osi y $\ln\mathcal{L}(\mu|\mathbf{x}) = l(\mu|\mathbf{x}) - \max(l(\mu|\mathbf{x}))$. Porovnajte $\ln\mathcal{L}(\mu|\mathbf{x})$ s kvadratickou aproximáciou vypočítanou pomocou Taylorovho rozvoja $\ln\mathcal{L}(\mu|\mathbf{x}) = \ln(\frac{L(\mu|\mathbf{x})}{L(\widehat{\mu}|\mathbf{x})}) \approx -\frac{1}{2}\mathcal{I}(\widehat{\mu})(\mu - \widehat{\mu})^2$. (2) Nech skóre funkcia $S(\mu) = \frac{\partial}{\partial\mu}\ln L(\mu|\mathbf{x})$. Keď zoberieme deriváciu kvadratickej aproximácie uvedenej vyššie, dostaneme $S(\mu) \approx -\mathcal{I}(\widehat{\mu})(\mu - \widehat{\mu})$ alebo $-\mathcal{I}^{-1/2}(\widehat{\mu})S(\mu) \approx \mathcal{I}^{1/2}(\widehat{\mu})(\mu - \widehat{\mu})$. Potom zobrazením pravej strany na osi x a ľavej strany na osi y dostaneme asymptoticky lineárnu funkciu s jednotkovým sklonom. Asymptoticky tiež platí $\mathcal{I}^{1/2}(\overline{X})(\mu - \overline{X}) \stackrel{\mathcal{D}}{\sim} N(0,1)$. Je postačujúce mať rozsah osi x rovný $\langle -2, 2\rangle$, pretože funkcia je asymptoticky (lokálne) lineárna na tomto intervale. Rozumne škálujte os y. Zobrazte pre (a) $n = 10$, (b) $n = 100$ a (c) $n = 1000$. Použite (1) $X \sim N(0,1)$ a (2) $X \sim (1-p)N(0,1) + pN(0,2)$, kde $p = 0.05$. Okomentujte rozdiely medzi (a), (b) a (c), ako aj rozdiely medzi (1) a (2).*

2.4 *Maximalizácia funkcie vierohodnosti

Maximálne vierohodný odhad $\widehat{\theta}$ je možné vypočítať pomocou metód numerickej optimalizácie (pozri napr. Horová a Zelinka, 2008) aplikovaných na funkciu vierohodnosti $L(\theta|\mathbf{x})$ alebo jej logaritmus $l(\theta|\mathbf{x})$.

Newtonova (Newton-Raphsonova) metóda (metóda dotyčníc) je pomenovaná po Isaacovi Newtonovi (1643−1727) a Josephovi Raphsonovi (1648−1715). Majme kvadratickú aproximáciu logaritmu funkcie vierohodnosti pomocou Taylorovho rozvoja druhého rádu okolo nejakého bodu θ_0 definovanú ako

$$l(\theta|\mathbf{x}) \approx l(\theta_0|\mathbf{x}) + S(\theta_0)(\theta - \theta_0) - \frac{1}{2}\mathcal{I}(\theta_0)(\theta - \theta_0)^2$$

alebo lineárnu aproximácu skóre funkcie pomocou Taylorovho rozvoja prvého rádu

$$S(\theta) \approx S(\theta_0) - \mathcal{I}(\theta_0)(\theta - \theta_0).$$

Z tejto aproximácie môžeme odvodiť nasledovnú iteračnú funkciu

$$\theta_0 + \frac{S(\theta_0)}{\mathcal{I}(\theta_0)}.$$

Postup je nasledovný:

1. inicializácia metódy použitím vhodne zvoleného štartovacieho parametra $\theta^{(0)}$, pre ktorý platí $\mathcal{I}(\theta^{(0)}) \neq 0$,

2. iterácia rovnosti

$$\theta^{(i)} = \theta^{(i-1)} + \frac{S(\theta^{(i-1)})}{\mathcal{I}(\theta^{(i-1)})},$$

kde $\mathcal{I}(\theta^{(i-1)}) \neq 0$, pre $i = 1, 2, \ldots$, pokiaľ nebude $|l(\theta^{(i)}|\mathbf{x}) - l(\theta^{(i-1)}|\mathbf{x})| < \epsilon$, kde ϵ je vhodne zvolené malé číslo (prahová hodnota).

Newton-Raphsonova má jednoduchú geometrickú interpretáciu – bod $\theta^{(i)}$ je priesečník dotyčnice ku grafu skóre funkcie $S(\cdot)$ v bode $[\theta^{(i-1)}, S(\theta^{(i-1)})]$ s osou x. Táto metóda konverguje kvadraticky, t.j. počet správnych cifier odhadu sa v každom iteračnom kroku zdvojnásobí (rád metódy je rovný dvom). Konvergencia k lokálnemu extrému však nie je zaručená – metóda môže konvergovať k lokálnemu minimu alebo divergovať, ak začneme iterácie v $\theta^{(0)}$ z konvexnej časti $l(\theta|\mathbf{x})$. Aj keď je funkcia konkávna, konvergencia k lokálnemu maximu nie je zaručená (metóda nerozlišuje lokálne maximum a minimum, pretože rieši rovnosť $S(\theta) = 0$). Ak je funkcia záujmu multimodálna, nemôžeme očakávať, že metóda bude konvergovať ku globálnemu extrému. Ak je $S(\theta)$ dvakrát diferencovateľná, prvá a druhá derivácia nemenia znamienko na intervale $\langle\theta_D, \theta_H\rangle$, funkcia $S(\theta)$ má jednoduchý koreň ($S'(\theta) \neq 0$ pre každé $\theta \in \langle\theta_D, \theta_H\rangle$) a štartovací bod $\theta^{(0)}$ je ten z krajných bodov θ_D, θ_H, v ktorom je znamienko $S(\cdot)$ rovnaké ako znamienko jej druhej derivácie na intervale $\langle\theta_D, \theta_H\rangle$, potom metóda konverguje.

Newton-Raphsonova metóda je implementovaná v R vo funkcii optimize(f,interval,maximum= FALSE, tol,...), kde vstupným argumentom je buď funkcia vierohodnosti alebo jej logaritmus (f), štartovací interval (interval) a zvolená prahová hodnota (tol). Vo všeobecnosti nie je potrebné do funkcie pridávať argument obsahujúci prvú deriváciu funkcie záujmu, pretože je vypočítaná numericky. V prípade zadania argumentu hessian=TRUE, vo výsledkoch sa objaví aj $-\mathcal{I}(\widehat{\theta})$.

Alternatívnymi, avšak pomalšími, metódami sú **metóda zlatého rezu** a **metóda sukcesívnej parabolickej interpolácie**. V prvej z nich, kde derivácia funkcie záujmu nie je potrebná, sa interval $\left\langle\theta_D^{(i-1)}, \theta_H^{(i-1)}\right\rangle$, v ktorom leží maximum $l(\theta|\mathbf{x})$, v každom kroku delí v pomere zlatého rezu. Interval sa teda zužuje o $(3 - \sqrt{5})/2 \doteq 0.382$ (t.j. komplement zlatého rezu) jeho dĺžky, pričom deliaci bod intervalu je $\theta_{zr}^{(i-1)}$. Metóda zlatého rezu konverguje lineárne (rád metódy je rovný jednej), t.j. chyba sa zmenšuje v každom iteračnom kroku $(1-(3-\sqrt{5})/2)$-krát. Druhou metódou sa dopĺňa vo funkcii optimize() prvá v čase, keď je interval $\left\langle\theta_D^{(i-1)}, \theta_H^{(i-1)}\right\rangle$ príliž úzky. Bodmi $\theta_D^{(i-1)}, \theta_{zr}^{(i-1)}, \theta_H^{(i-1)}$ je preložená parabola (kvadratický interpolačný polynóm) a jej maximum bude novým bodom $\theta^{(i)}$. Prednastavené je hľadanie minima (maximum=FALSE) v štartovacom intervale (interval), čo môže byť zmenené na maximum (maximum=TRUE). Vo výstupoch funkcie optimize() bude $\widehat{\theta}$ (minimum alebo maximum) a $l(\widehat{\theta}|\mathbf{x})$ (objective).

Majme kvadratickú aproximáciu logaritmu funkcie vierohodnosti pomocou Taylorovho rozvoja druhého rádu okolo nejakého bodu $\boldsymbol{\theta}_0$ definovanú ako

$$l(\boldsymbol{\theta}|\mathbf{x}) \approx l(\boldsymbol{\theta}_0|\mathbf{x}) + S(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) - \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathcal{I}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$$

alebo lineárnu aproximáciu skóre funkcie pomocou Taylorovho rozvoja prvého rádu

$$S(\boldsymbol{\theta}) \approx S(\boldsymbol{\theta}_0) - \mathcal{I}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0).$$

Z tejto aproximácie môžeme odvodiť nasledovnú iteračnú funkciu

$$\boldsymbol{\theta}_0 + (\mathcal{I}(\boldsymbol{\theta}_0))^{-1}S(\boldsymbol{\theta}_0).$$

Postup je nasledovný:

1. inicializácia metódy použítím vhodne zvoleného štartovacieho parametra $\boldsymbol{\theta}^{(0)}$, pre ktorý platí $\mathcal{I}(\boldsymbol{\theta}^{(0)}) \neq \mathbf{0}$,
2. iterácia rovnosti
$$\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^{(i-1)} + (\mathcal{I}(\boldsymbol{\theta}^{(i-1)}))^{-1}S(\boldsymbol{\theta}^{(i-1)}),$$
$\mathcal{I}(\boldsymbol{\theta}^{(i-1)}) \neq \mathbf{0}$, pre $i = 1, 2, \ldots$, pokiaľ nebude $|l(\boldsymbol{\theta}^{(i)}|\mathbf{x}) - l(\boldsymbol{\theta}^{(i-1}|\mathbf{x})| < \epsilon$, kde ϵ je vhodne zvolené malé číslo (prahová hodnota). Vo všeobecnosti sa $-\mathcal{I}(\boldsymbol{\theta}^{(i-1)}) = l''(\boldsymbol{\theta}^{(i-1)}|\mathbf{x})$ nazýva *hesián*.

Namiesto $(\mathcal{I}(\boldsymbol{\theta}^{(i-1)}))^{-1}$ je lepšie použiť riešenie systému rovníc $(\mathcal{I}(\boldsymbol{\theta}^{(i-1)}))\mathbf{z}_{i-1} = S(\boldsymbol{\theta}^{(i-1)})$ pre nejaké $\mathbf{z}$ a potom $\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^{(i-1)} + \mathbf{z}_{i-1}$. V niektorých štatistických modeloch (napr. v logistickom regresnom modeli) sa namiesto $\mathcal{I}(\boldsymbol{\theta}^{(i-1)})$ používa $I(\boldsymbol{\theta}^{(i-1)})$, ktorá má často jednoduchší tvar. Potom hovoríme o **Fisherovej skóringovej metóde**. Ak namiesto $\mathcal{I}(\boldsymbol{\theta}^{(i-1)})$ použijeme jej pozitívne definitnú aproximáciu počítanú pomocou sukcesívne počítaných gradientov, hovoríme o **quasi Newtonovej metóde** (v angličtine nazývanej aj *variable metric method*). Fisherova skóringová metóda je potom vlastne quasi Newtonova metóda. Gradient nemusí byť špecifikovaný ako funkcia, ale môže byť počítaný numericky, napr. *centrálnou rozdielovou aproximáciou*, ako

$$\frac{\partial}{\partial \theta_i} l(\boldsymbol{\theta}|\mathbf{x}) \approx \frac{l(\boldsymbol{\theta} + \epsilon \mathbf{e}_i|\mathbf{x}) - l(\boldsymbol{\theta} - \epsilon \mathbf{e}_i|\mathbf{x})}{2\epsilon}, \text{ kde } i = 1, 2, \ldots, k,$$

i-ty komponent bazálneho vektora $\mathbf{e}_i$ obsahuje jednotku a na ostatných miestach sú nuly a ϵ je malé číslo. Derivácia v i-tom smere je potom pre $\epsilon \to 0$ nahradená jej konečnou aproximáciou. Quasi Newtonova metóda je implementovaná v R vo funkcii `optim(par,fn,gr,method,control,hessian =FALSE,...)`, kde gradient môže byť špecifikovaný voliteľným argumentom `gr` (prednastavená je metóda spomenutá vyššie[10]). Populárnou metódou aproximácie hesiánu je **Broyden-Fletcher-Goldfarb-Shannova (BFGS) metóda**, kde

$$l''(\boldsymbol{\theta}^{(i)}|\mathbf{x}) \approx l''(\boldsymbol{\theta}^{(i-1)}|\mathbf{x}) + \frac{\mathbf{y}^{(i-1)}(\mathbf{y}^{(i-1)})^T}{(\mathbf{y}^{(i-1)})^T\mathbf{s}^{(i-1)}} - \frac{l''(\boldsymbol{\theta}^{(i-1)}|\mathbf{x})\mathbf{s}^{(i-1)}(l''(\boldsymbol{\theta}^{(i-1)}|\mathbf{x})\mathbf{s}^{(i-1)})^T}{(\mathbf{s}^{(i-1)})^T l''(\boldsymbol{\theta}^{(i-1)}|\mathbf{x})\mathbf{s}^{(i-1)}},$$

kde $\mathbf{y}^{(i-1)} = S(\boldsymbol{\theta}^{(i)}) - S(\boldsymbol{\theta}^{(i-1)})$ a $\mathbf{s}^{(i-1)} = \boldsymbol{\theta}^{(i)} - \boldsymbol{\theta}^{(i-1)} = (\mathcal{I}(\boldsymbol{\theta}^{(i-1)}))^{-1}S(\boldsymbol{\theta}^{(i-1)})$. BFGS metóda je implementovaná vo funkcii `optim()` pri nastavení argumentu `method="BFGS"`. Strata kvadratickej konvergencie oproti Newtonovej metóde je spôsobená aproximáciou hesiánu. BFGS metóda patrí do tzv. *Broydenovej triedy*, kde je hodnota $l''(\boldsymbol{\theta}^{(i)}|\mathbf{x})$ modifikovaná štvrtým členom, ktorý je pri BFGS nulový (pozri Givens a Hoeting, 2005).

Vo funkcii `optim()` je prednastavená **Nelder-Meadova metóda** (nazývaná aj **metóda simplexov**; argument `method="Nelder-Mead"`), ktorá nepoužíva gradient. Je robustná na nespojité funkcie, ale konverguje pomaly. Je vytvorená na základe myšlienky „preskokov“ cez trojuholníky. V každom kroku majme trojuholník definovaný troma bodmi $\boldsymbol{\theta}_1^{(i-1)}$, $\boldsymbol{\theta}_2^{(i-1)}$, $\boldsymbol{\theta}_3^{(i-1)}$, kde platí $l(\boldsymbol{\theta}_1^{(i-1)}|\mathbf{x}) < l(\boldsymbol{\theta}_2^{(i-1)}|\mathbf{x}) < l(\boldsymbol{\theta}_3^{(i-1)}|\mathbf{x})$ a snažíme sa $\boldsymbol{\theta}_1^{(i-1)}$ nahradiť „lepším“ bodom $\boldsymbol{\theta}_1^{(i)}$, pre ktorý platí $l(\boldsymbol{\theta}_1^{(i)}|\mathbf{x}) > l(\boldsymbol{\theta}_1^{(i-1)}|\mathbf{x})$. Ak je tak možné urobiť, nový bod definujeme pomocou stredovej súmernosti a extrapolácie ako

$$\boldsymbol{\theta}_1^{(i)} = \boldsymbol{\theta}_{23}^{(i-1)} + 2\left(\boldsymbol{\theta}_{23}^{(i-1)} - \boldsymbol{\theta}_1^{(i-1)}\right),$$

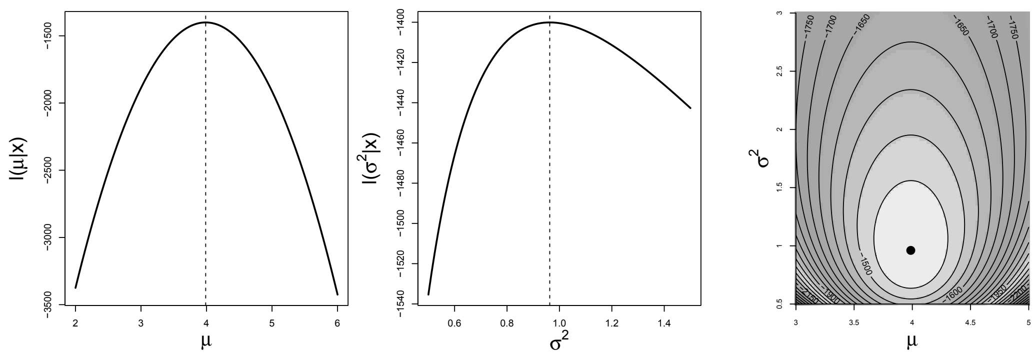
kde $\boldsymbol{\theta}_{23}^{(i-1)} = \frac{\boldsymbol{\theta}_2^{(i-1)} + \boldsymbol{\theta}_3^{(i-1)}}{2}$. Z vyššie uvedeného vyplýva, že bod $\boldsymbol{\theta}_1^{(i-1)}$zobrazujeme cez stred úsečky s krajnými bodmi $\boldsymbol{\theta}_2^{(i-1)}$ a $\boldsymbol{\theta}_3^{(i-1)}$ do bodu $\boldsymbol{\theta}_{23}^{(i-1)} + \left(\boldsymbol{\theta}_{23}^{(i-1)} - \boldsymbol{\theta}_1^{(i-1)}\right)$ a potom postupujeme po tejto polpriamke ešte ďalej do bodu $\boldsymbol{\theta}_1^{(i)}$. Ak $l(\boldsymbol{\theta}_1^{(i)}|\mathbf{x}) > l(\boldsymbol{\theta}_1^{(i-1)}|\mathbf{x})$, potom trojuholník nahradíme novým trojuholníkom definovaným bodmi $\boldsymbol{\theta}_1^{(i)}$, $\boldsymbol{\theta}_2^{(i-1)}$, $\boldsymbol{\theta}_3^{(i-1)}$.

[10] Centrálna rozdielová aproximácia hrá kľúčovú úlohu pri aproximácii hesiánu, ktorý sa používa v maximálnej vierohodnosti na odhad rozptylu.

Pri maximalizácii logaritmu funkcie vierohodnosti je vo vstupe funkcie optim() potrebné nastaviť hessian=TRUE a argument control=list(fnscale=-1), ktorý robí maximalizáciu namiesto prednastavenej minimalizácie (v argumente control je možné nastaviť aj maximálne množstvo iterácií pomocou maxit). Argument par predstavuje štartovaciu hodnotu parametra. Vo výstupoch funkcie optim() bude $\widehat{\boldsymbol{\theta}}$ (par), $l(\widehat{\boldsymbol{\theta}}|\mathbf{x})$ (value) a $l''(\widehat{\boldsymbol{\theta}}|\mathbf{x})$ (hessian). Ak je vo výstupe convergence rovné nule, potom maximalizácia skonvergovala (k lokálnemu maximu, ktoré nemusí byť globálne).

Príklad 103 (maximálne vierohodný odhad μ a σ^2) *Vygenerujte pseudonáhodné čísla z $X \sim N(4,1)$, $n = 1000$. (a) Napíšte logaritmus profilovej funkcie vierohodnosti pre μ a σ^2 a preverte, či sú maximálne vierohodné odhady μ a σ^2 dostatočne blízko k ich skutočným hodnotám. Nakreslite grafy $l(\mu|\mathbf{x})$ a $l(\sigma^2|\mathbf{x})$, kde zvýrazníte polohu maxím týchto funkcií. (b) Napíšte logaritmus funkcie vierohodnosti pre $\boldsymbol{\theta} = (\mu, \sigma^2)^T$ a preverte, či je maximálne vierohodný odhad $\widehat{\boldsymbol{\theta}}$ dostatočne blízko k jeho skutočnej hodnote. (c) Nakreslite graf $l(\boldsymbol{\theta}|\mathbf{x})$ použitím funkcie image() a superponujte ho s kontúrovým grafom použitím funkcie contour(). Zvýraznite polohu maxima.*

Riešenie (pozri obrázok 2.17)
Logaritmus funkcie vierohodnosti pre jednotlivé parametre má tvar
$l(\mu|\mathbf{x}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma_1^2 - \frac{1}{2\sigma_1^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)$, kde $\mu \in (2,6), \sigma_1 = 1$;
$l(\sigma^2|\mathbf{x}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{\sum_{i=1}^n (x_i-\mu_1)^2}{2\sigma^2}$, kde $\mu_1 = 4, \sigma \in (0.5, 1.5)$;
$l(\boldsymbol{\theta}|\mathbf{x}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}$, kde $\mu \in (2,6)$ a $\sigma \in (0.5, 1.5)$.
Výsledky simulácie: $\widehat{\mu} = 4.019708$ a $\widehat{\sigma}^2 = 1.000038$.



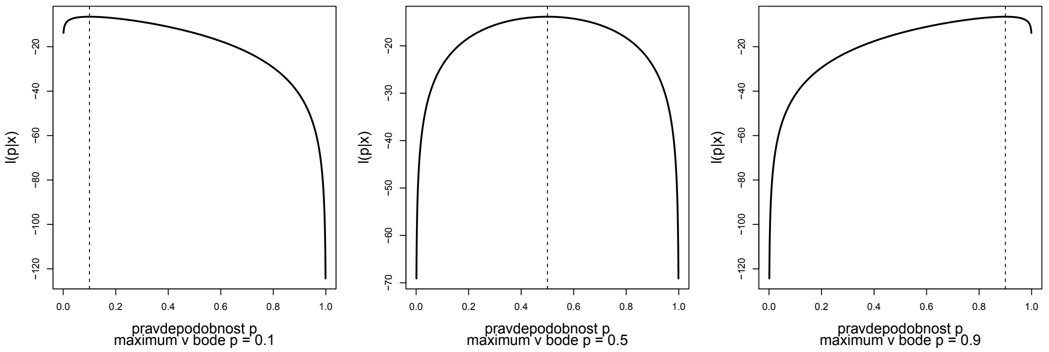
Obr. 2.17: Profilová funkcia vierohodnosti pre μ (vľavo), σ^2 (uprostred) a funkcia vierohodnosti pre oba parametre (vpravo); $X \sim N(4,1)$); maximálne vierohodné odhady strednej hodnoty a rozptylu sú označené zvislou čiarkovanou čiarou (vľavo a uprotred) a maximálne vierohodný odhad vektora parametrov je označený • (vpravo)

Ak náhodná premenná X nebude mať normálne rozdelenie, funkcia vierohodnosti pre strednú hodnotu nemusí mať symetrický parabolický tvar okolo strednej hodnoty. Odhad strednej hodnoty môže byť potom vychýlený.

Príklad 104 (maximálne vierohodné odhady) *Za predpokladu normality rozdelenia náhodnej premennej X vypočítajte maximálne vierohodné odhady strednej hodnoty μ (ozn. $\widehat{\mu}$) a rozptylu σ^2 (ozn. $\widehat{\sigma}^2$) pomocou logaritmov funkcií vierohodnosti $l(\mu|\mathbf{x})$, resp. $l(\sigma^2|\mathbf{x})$. Porovnajte tieto odhady s aritmetickým priemerom $\overline{x}$ a rozptylom s^2. Musí platiť $\widehat{\mu} = \overline{x}$ a $\widehat{\sigma}^2 = \frac{(n-1)}{n}s^2$. Realizáciami náhodnej premennej X sú hodnoty $x_i, i = 1, 2, \ldots, n$, premenných: (a) dĺžka pravej kľúčnej kosti (length.R; dáta: paired-means-clavicle2.txt); (b) morfologická výška tváre (face.H; dáta: one-sample-correlation-skull-mf.txt); (c) šírka lebky (skull.B; dáta: one-sample-mean-skull-mf.txt).*

Príklad 105 (binomické rozdelenie, maximálne vierohodný odhad p) *Nech $X \sim Bin(N,p)$ a realizácie X sú $x = n$. Predpokladajme, že sme pozorovali (a) $x = 2$, (b) $x = 10$ a (c) $x = 18$ úspechov v $N = 20$ pokusoch. Pomocou R vypočítajte maximálne vierohodný odhad p. Výsledok zobrazte do grafu spolu s funkciou vierohodnosti.*

Riešenie (pozri obrázok 2.18)
Logaritmus funkcie vierohodnosti pre p má tvar $l((p|\mathbf{x}) = x\log(p) + (N-x)\log(1-p)$, kde $p \in (0,1)$.



Obr. 2.18: Funkcia vierohodnosti pre $X \sim Bin(N,p)$ ($p = 0.1, 0.5, 0.9$ a $N = 20$); odhady $\widehat{p}$ sú označené zvislou čiarkovanou čiarou

Ďalej derivujeme
$\partial l\,(p|\mathbf{x})\,|\partial p = x/p - (N-x)\,/\,(1-p) = [x\,(1-p) - (N-x)\,p]\,/\,[p\,(1-p)] = (x - Np)\,/\,[p\,(1-p)] = 0$,
potom $\widehat{p} = x/N$.
(a) $\widehat{p} = x/N = 2/20 = 0.1$,
(b) $\widehat{p} = x/N = 10/20 = 0.5$,
(c) $\widehat{p} = x/N = 18/20 = 0.9$.

Z grafov na obrázku 2.18 je zreteľné, že funkcia vierohodnosti pre p je symetrická len pre $p = 0.5$, pre ostatné p je asymetrická. Naviac pre p a $1-p$ dostaneme grafy, ktoré možno transformovať jeden na druhý pomocou osi zrkadlenia definovanej ako vertikálna priamka v $p = 0.5$.

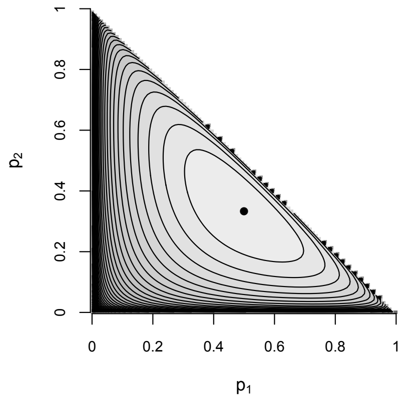
Príklad 106 (maximálne vierohodné odhady; binomické rozdelenie) *Za predpokladu, že náhodná premenná X má binomické rozdelenie, vypočítajte maximálne vierohodný odhad $\widehat{p}$ pomocou logaritmu funkcie vierohodnosti $l(p|\mathbf{x})$. Porovnajte tento odhad s výrazom $\sum_{i=1}^{N} x_i/N$. Realizáciami náhodnej premennej X sú nasledujúce binárne premenné: (a) pohlavie (sex; dáta: one-sample-probability-sexratio.txt), kde ozn. pohlavia dievča „f“ preznačíme na 1 a ozn. pohlavia chlapec „m“ preznačíme na 0; (b) pohlavie (sex; dáta: two-samples-probabilities-sexratio.txt), kde ozn. pohlavia muž „m“ preznačíme na 1 a ozn. pohlavia žena „f“ preznačíme na 0. V prípade (a) počítame pravdepodobnosť výskytu dievčat a v prípade (b) pravdepodobnosť výskytu chlapcov.*

Príklad 107 (maximálne vierohodné odhady; multinomické rozdelenie) *Nech náhodná premenná X má multinomické rozdelenie. Potom vypočítajte maximálne vierohodné odhady $\widehat{p}_1$ a $\widehat{p}_2$ pomocou logaritmu funkcie vierohodnosti $l(\mathbf{p}|\mathbf{x})$. Porovnajte odhad p_1 s odhadom p z príkladu 106, kde pravdepodobnosť $\widehat{p}_1$ bola označená ako $\widehat{p}$. Realizáciami X sú binárne premenné: (a) pohlavie (sex; dáta: one-sample-probability-sexratio.txt), kde ozn. pohlavia dievča „f“ preznačíme na 1 a ozn. pohlavia chlapec „m“ preznačíme na 0; (b) pohlavie (sex; dáta: two-samples-probabilities-sexratio.txt), kde ozn. pohlavia muž „m“ preznačíme na 1 a ozn. pohlavia žena „f“ preznačíme na 0. Pravdepodobnosť $\widehat{p}_1$ je (a) pravdepodobnosť výskytu dievčat a (b) pravdepodobnosť výskytu chlapcov.*

Príklad 107 hovorí o tom, že parameter p_1 dvojrozmerného multinomického rozdelenia je parametrom p binomického rozdelenia.

Príklad 108 (maximálne vierohodné odhady; multinomické rozdelenie) *Majme dáta more-samples-probabilities-pubis.txt. Nakreslite logaritmus štandardizovanej $\mathcal{L}(\boldsymbol{\theta}|\mathbf{x})$, kde $\boldsymbol{\theta} = (p_1,\, p_2)^T$, Európskej populácie ($n_1 = 30$, $n_2 = 20$ a $n_3 = 10$) pomocou funkcie contour(). Dokreslite do obrázku jej maximum v bode $\widehat{\boldsymbol{\theta}} = (\widehat{p}_1, \widehat{p}_2)^T$.*

Riešenie (pozri obrázok 2.19)



Obr. 2.19: Logaritmus štandardizovanej funkcie vierohodnosti multinomického rozdelenia v parametroch p_1 a p_2 (Európska populácia) s maximom označeným •

Príklad 109 (overdispersion v Poissonovom modeli, pokrač.) *Majme početnosti úrazov n medzi m_n robotníkmi v továrni, pozri tabuľku 2.10 (Greenwood a Yule, 1920). Vypočítajte očakávané m_n za predpokladu, že početnosti úrazov na robotníka X majú negatívne binomické rozdelenie s parametrami α a π.*

Riešenie (pozri tabuľku 2.14)
Aby sme mohli fitovať negatívne binomické rozdelenie, potrebujeme funkciu vierohodnosti

$$L(\alpha, \pi|\mathbf{x}) = \prod_{n=0}^{4} (\Pr(X=n))^{m_n} \left(1 - \sum_{n=0}^{4} \Pr(X=n)\right)^{m_{\geq 5}}.$$

a jej logaritmus

$$l(\alpha, \pi|\mathbf{x}) = \sum_{n=0}^{4} m_n \ln \Pr(X=n) + m_{\geq 5} \ln \left(1 - \sum_{n=0}^{4} \Pr(X=n)\right).$$

Numerickou optimalizáciou dostaneme $\widehat{\alpha} = 0.84$ a $\widehat{\pi} = 0.64$. Pomer zlyhaní $\widehat{\mu} = \frac{1-\widehat{\pi}}{\widehat{\pi}}\widehat{\alpha} = 0.47$. Keď porovnáme pozorované m_n a vypočítané (teoretické) m_n zistíme, že početnosti sú veľmi podobné (pozri tabuľku 2.14).

Tabuľka 2.14: Očakávané početnosti robotníkov m_n (zaokrúhlené na nula desatinných miest) s n úrazmi v továrni (negatívne binomické rozdelenie)

n	0	1	2	3	4	≥ 5
očakávané m_n	446	134	44	15	5	3

2.5 Kritériá klasifikácie štatistických modelov

Kritérá delenia štatistických modelov (ako aj k nim prislúchajúcich testov) sú nasledovné:

- **množstvo výberov** – jeden, dva alebo viac ako dva výbery [jedno-, dvoj- a viacvýberový test/model];
- **závislosť výberov** – nezávislé a závislé výbery (opakované merania na subjekte v čase, merania na párových orgánoch) [test/model dvoch a viacerých nezávislých výberov, párový test];
- **množstvo premenných** – jedna, dve alebo viac premenných [jedno-, dvoj- a viac-premenný test/model];

- **typ premenných** – kvalitatívne alebo kvantitatívne premenné [binomický, Poissonov, multinomický, súčinový multinomický model, model normálneho rozdelenia, rôzne modely kauzality];

- **rozmer endpointov/závislých premenných** – jedna, dve alebo viac premenných [jedno-, dvoj- a viac-rozmerný test/model – lineárny regresný model (LRM) vs. mnohorozmerný LRM (MLRM), analýza rozptylu ANOVA vs. mnohorozmerná ANOVA (MANOVA), analýza kovariancie ANCOVA vs. mnohorozmerná ANCOVA (MANCOVA)];

- **typ náhodnosti efektov** – fixné, náhodné alebo zmiešané [(M)ANOVA, (M)ANCOVA model a test/testy v nich];

- **typ kauzálneho vzťahu** – lineárny (priamka) alebo nelineárny (polynóm ľubovoľného stupňa – kvadratický, kubický a iný, alebo ľubovoľná funkcia);

- **typ vzťahu parametrov modelu** – lineárny [LRM – priamka, polynóm ľubovoľného stupňa a pod.] a nelineárny [nelineárny regresný model (NLRM)]; toto kritérium sa často zamieňa s predchádzajúcim, ale tieto dve kritéria nie sú totožné;

- **prítomnosť odľahlých pozorovaní**;

- **typ hypotézy** – jednostranná, obojstranná, stochasticky usporiadaná.

2.6 Praktické dôsledky odchýlok od normality

Základným predpokladom mnohých modelov je normalita rozdelenie spojitých premenných, čo však nemusí byť v praxi splnené. Napr. hodnoty často používanej telesnej hmotnosti nemajú v západnej civilizácii normálne rozdelenie, ale negatívne zošikmené (pozri kapitolu 3.2 Charakteristiky variability). Odchýlky od normality môžu mať mnoho príčin, od **neošetrených metodických nezrovnalostí** až ku **skutočným biologickým procesom v populácii**, z ktorých (náhodný) výber (vzorka) pochádza. Predpoklad normality by nikdy nemal byť samozrejmý (ani u premenných, ktoré obvykle normálne rozdelenie majú), dáta by mali byť pred štatistickými analýzami vhodne zobrazené (pozri kapitolu 3.6 Štatistická grafika) a ich normalita by mala byť v každej vzorke testovaná (pozri kapitolu 5.2 Kolmogorov-Smirnovov test dobrej zhody). Normalita rozdelenia by mala byť starostlivo posudzovaná a v prípade zachytenia odchýlkok od normality v nejakom znaku v mnohých nezávislých výberoch by príčiny odchýlky mali byť bližšie skúmané. Malo by sa zistiť, či ide o metodickú chybu (odľahlé hodnoty žiadajúce odstránenie, nenáhodný výber, atď.) alebo skutočný trend, ktorý má svoje biologické príčiny (*rozdielna regulácia hornej a dolnej medze intenzity nejakého metabolického procesu* alebo *smerová/direkcionálna selekcia* u zošikmeného rozdelenia, *sledovaným znakom obmedzené vzorkovanie* (nenáhodný výber) alebo *stabilizačná selekcia u leptokurtického rozdelenia* (pozri kapitolu 3.2 Charakteristiky variability), atď.). Zošikmené rozdelenie majú často *inkrementálne (prírastkové) dáta* merané v priebehu ontogenézy, keďže je ich dolná hranica prirodzene obmedzená (nulový prírastok), zatiaľ čo horná môže u každého jedinca dosiahnúť rôzne veľké hodnoty (Garn a Rohmann, 1963). Treba mať na pamäti, že našim hlavným a najdôležitejším cieľom je zistenie podstaty procesov, ktoré prebiehajú v populácii, z ktorej pochádza naša vzorka a ktorej je obrazom.

V bežnej praxi je však časté, že sú používané relatívne malé vzorky, v ktorých sa odchýlky od normality ľahko prejavia. Môžu však nastať aj *kontroverzné situácie*, že rozdelenie dát (pri relatívne malej vzorke) sa nepodobá na normálne, ale test dobrej zhody s normálnym rozdelením hypotézu o zhode zamietne. Rovnako sa často stáva, že meranie alebo experiment nie sú dopredu naplánované z hľadiska náhodného výberu, použitia nejakého modelu, z hľadiska minimalizácie variability a odhadu minimálneho rozsahu súboru. Dôležitosť takéhoto postupu je bežne aplikovanými výskumníkmi ignorovaná, čo môže viesť až do situácie nemožnosti použitia dát samotných alebo nemožnosti použitia akejkoľvek štatistickej metódy.

Súčasne sa utvrdzuje predstava o ideálnej podstate normality v prírode a tiež všeobecnej normalite rozdelenia často testovaných znakov, takže akékoľvek ďalšie, novozískané odchýlky od normality sa považujú za chybu. Skutočnosť ale často ukazuje, že so zvyšujúcim sa počtom prípadov vo vzorke (virtuálne až k celkovej veľkosti populácie) celý rad (empirických) veličín v skutočnosti nesmeruje ku stále dokonalejšej Gausovej krivke, ale má rozdelenie od normálneho viac či menej sa vzďaľujúce alebo nejakým spôsobom vychýlené. V týchto prípadoch sa odporúča

1. použitie nejakej **transformácie dát** (v prípade zošikmenia; logaritmická, odmocninová, Box-Coxova a pod.), ktorá závisí na samotných dátach a nedá sa použiť univerzálne;
2. použitie **urezávania alebo winsorizácie dát** na jednom alebo oboch koncoch rozdelenia (v prípade prítomnosti odľahlých pozorovaní alebo zošikmenia; pozri kapitolu 3.1 Charakteristiky polohy, 3.2 Charakteristiky variability a 3.3 Detekcia odľahlých pozorovaní);
3. **nahradenie asymptotického rozdelenia testovacej štatistiky bootstrapovým alebo permutačným** (v prípade nedostatočného alebo relatívne malého rozsahu vzorky alebo pochybností o asymptotickom rozdelení testovacej štatistiky aj pri väčších rozsahoch).

Po aplikovaní prvých dvoch metód je možné použiť asymptotické testy v nezmenenej podobe a v treťom prípade sa použijú len samotné testovacie štatistiky.

3 Charakteristiky polohy a variability a štatistická grafika

Každé rozdelenie pravdepodobnosti býva charakterizované parametrami, ktoré sa odhadujú z realizácií (dát). Tieto parametre označujeme ako **číselné charakteristiky rozdelenia** (**štatistiky**). Patria medzi ne **charakteristiky polohy a variability** ako napr. stredná hodnota, medián a ostatné kvantily, rozptyl, rozpätie a pod.

Stredná (očakávaná) hodnota (prvý začiatočný moment) rozdelenia náhodnej veličiny X, ozn. $E[X]$, vypočítame ako (Casella a Berger, 2002)

$$E[X] = \begin{cases} \sum_{i=1}^{n} x_i \Pr(X = x_i), & \text{ak } X \text{ je diskrétna náhodná premenná s realizáciami } x_1, x_2, \ldots, x_n, \\ \int_{-\infty}^{\infty} x f_X(x) dx, & \text{ak } X \text{ je spojitá náhodná premenná s hustotou } f_X(x). \end{cases}$$

Rozptyl (druhý centrálny moment) $Var[X] = E[(X - E[X])^2]$, ktorého kladnú odmocninu $SD[X]$, vyjadrujúcu „rozptýlenie" X okolo $E[X]$, nazývame **smerodajná odchýlka**.

Príklad 110 (charakteristiky normálneho rozdelenia) *Nech X_i sú iid z normálneho rozdelenia $N(\mu, \sigma^2)$, t.j. $X_i \sim N(\mu, \sigma^2)$. Potom $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ nazývame **výberová stredná hodnota** a $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ **výberový rozptyl***

Kedže X je náhodná premenná, $\overline{X}$ nie je aritmetický priemer, ale tzv. výberový priemer; aritmetický priemer $\overline{x}$ je realizácia výberového priemeru. Podobne S^2 je tzv. výberový rozptyl; $\widehat{\sigma}^2 = s^2$ je jeho realizácia. Tiež $SD[X] = S$ je výberová smerodajná odchýlka a nie smerodajná odchýlka; smerodajná odchýlka $\widehat{\sigma} = s$. Podiel $V_k = S^2/\overline{X}$ nazývame **výberový koeficient variácie**.

Príklad 111 (charakteristiky binomického rozdelenia) *Ak X pochádza z binomického rozdelenia, ozn. $Bin(N, p)$, potom $E[X] = Np$ je **stredná hodnota** a $Var[X] = Np(1-p)$ **rozptyl** náhodnej veličiny X.*

Nech $X_1, X_2, \ldots, X_n$ je náhodný výber z nejakého rozdelenia a $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ **usporiadaný náhodný výber** (vo vzrastajúcom poradí) s rozsahom n, kde $X_{(i)}$ nazývame **poriadkové štatistiky**. Sú to náhodné premenné, pre ktoré platí $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$. Potom (Casella a Berger, 2002)

$$X_{(1)} = \min_{1 \le i \le n} X_i,$$

$$X_{(2)} = \text{druhé najmenšie } X_i \ (X_i \text{ v poradí druhé})$$

$$\vdots$$

$$X_{(n-1)} = X_i \text{ v poradí } (n-1)\text{-vé},$$

$$X_{(n)} = \max_{1 \le i \le n} X_i.$$

Výberový medián je charakterizovaný číslom $Q_2 = \widetilde{X}$, ktoré rozdeľuje náhodný výber tak, že približne polovica X_i je menšia ako táto hodnota a polovica je väčšia ako táto hodnota. Definujeme ho ako[1]

$$\widetilde{X} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{ak } n \text{ je nepárne}, \\ \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right) & \text{ak } n \text{ je párne}. \end{cases}$$

[1] Ak n je nepárne, potom medián je prostredná hodnota, t.j. hodnota nachádzajúca sa v poradí ako prostredná alebo na mieste X, ktoré zodpovedá poradiu $(n+1)/2$. Ak n je nepárne, potom medián je uprostred zoradených hodnôt, t.j. hodnota na mieste X, ktoré zodpovedá priemeru X s poradiami $n/2$ a $n/2+1$.

Majme číslo $p \in (0,1)$. Potom **výberovým $100p$-tym percentilom náhodného výberu** $X_{(\{np\})}$ bude také X_i, kde približne np hodnôt X_i bude menších ako $X_{(\{np\})}$ a $n(1-p)$ väčších ako $X_{(\{np\})}$. Ak $p = 0.5$, ide o 50. percentil alebo medián. Označenie $\{x\}$ v indexe znamená *najbližšie celé číslo*, kde $i - 0.5 \leq x < i + 0.5$, kde $\{x\} = i$. Potom

$$\widetilde{X}_p = X_{(\{np\})} = \begin{cases} X_{(\{np\})}, & \text{ak } \frac{1}{2n} < p \leq 0.5, \\ X_{(n+1-\{n(1-p)\})}, & \text{ak } 0.5 < p < 1 - \frac{1}{2n}. \end{cases}$$

Príklad 112 (výberový $100p$-ty percentil) *Ak $n = 12$, potom výberovým 65. percentilom je $X_{(9)}$, pretože $n(1-p) = 12 \times (1 - 0.65) = 4.2$ a $n + 1 - \{n(1-p)\} = 12 + 1 - 4 = 9$.*

Často používanými percentilmi[2] sú **výberový dolný kvartil** (25. percentil) $Q_1 = \widetilde{X}_{0.25}$ a **výberový horný kvartil** (75. percentil) $Q_3 = \widetilde{X}_{0.75}$.

Na výpočet rozptylu nejakého percentilu musíme poznať pravdepodobnostnú funkciu poriadkovej štatistiky, jej hustotu a distribučnú funkciu. Výpočet môžeme výhodne zjednodušiť, keď vieme, aké je asymptotické rozdelenie poriadkovej štatistiky.

Definícia 27 (pravdepodobnostná funkcia poriadkovej štatistiky) *Majme náhodný výber $X_1, X_2, \ldots, X_n$ z nejakého diskrétneho rozdelenia s pravdepodobnostnou funkciou $f_X(x_i) = p_i$, kde $x_1 < x_2 < \ldots < x_n$. Nech $P_0 = 0$, $P_1 = p_1$, $P_2 = p_1 + p_2, \ldots, P_i = p_1 + p_2 + \ldots + p_i$. Nech $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ sú poriadkové štatistiky. Potom (Bickel a Doksum, 2006)*

$$\Pr(X_{(j)} \leq x_i) = \Pr(Y \geq j) = \sum_{k=j}^{n} \binom{n}{k} P_i^k (1 - P_i)^{n-k}$$

a

$$\Pr(X_{(j)} = x_i) = \Pr(X_{(j)} \leq x_i) - \Pr(X_{(j)} \leq x_{i-1}) = \sum_{k=j}^{n} \binom{n}{k} \left[P_i^k (1 - P_i)^{n-k} - P_{i-1}^k (1 - P_{i-1})^{n-k}\right],$$

kde Y je náhodná premenná počtu hodnôt z $X_1, X_2, \ldots, X_n$, ktoré sú menšie alebo rovné ako x_i. Nech $\{X_j \leq x_i\}$ je priaznivá udalosť a $\{X_j > x_i\}$ je nepriaznivá udalosť. Ak $i = 1$, potom $\Pr(X_j = x_i) = \Pr(X_j \leq x_i)$, pretože $\Pr_0 = 0$. Teda Y je počet priaznivých udalostí v n pokusoch, t.j. $Y = card\{X_j \leq x_i\}$. Pravdepodobnosť úspechu je potom $\Pr(X_j \leq x_i)$ pre každý pokus, pretože pokusy sú rovnako rozdelené. Priaznivá a nepriaznivá udalosť j-teho pokusu je nezávislá od výsledku iného pokusu, pretože X_j sú nezávislé od ostatných X_i. Teda $Y \sim Bin(n, P_i)$.

Ozn. card$\{\cdot\}$ znamená veľkosť (kardinalitu) množiny.

Ak $X_1, X_2, \ldots, X_n$ je náhodný výber zo spojitého rozdelenia, zhody neprichádzajú do úvahy a pravdepodobnosť, že nejaké dve alebo viaceré X_j sú rovnaké, je nula. Teda $\Pr(X_{(1)} < X_{(2)} < ... < X_{(n)}) = 1$ a výberový priestor pre $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ je $\mathcal{X} = \{(x_1, x_2, \ldots, x_n) : x_1 < x_2 < \ldots < x_n\}$.

Definícia 28 (hustota a distribučná funkcia poriadkovej štatistiky) *Nech $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ sú poriadkové štatistiky náhodného výberu $X_1, X_2, \ldots, X_n$ s distribučnou funkciou $F_X(x)$ a hustotou $f_X(x)$. Potom hustota poriadkovej štatistiky $X_{(j)}$ je rovná (Bickel a Doksum, 2006)*

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x) \left[F_X(x)\right]^{j-1} \left[1 - F_X(x)\right]^{n-j}$$

a distribučná funkcia

$$F_{X_{(j)}}(x) = \Pr(X_{(j)} \leq x) = \Pr(Y \geq j) = \sum_{k=j}^{n} \binom{n}{k} \left[F_X(x)\right]^k \left[1 - F_X(x)\right]^{n-k}.$$

[2] Všeobecne definujeme **kvantil** ako hodnotu skúmanej veličiny, ktorá delí náhodný výber na dve časti pod a nad kvantilom; číslo v dolnom indexe hovorí o časti náhodného výberu pod kvantilom. Podľa toho rozlišujeme aj jeho typ, napr. kvartil, decil, percentil a pod.

Definícia 29 (asymptotické rozdelenie poriadkovej štatistiky) *Nech $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ sú poriadkové štatistiky náhodného výberu $X_1, X_2, \ldots, X_n$. Majme pravdepodobnosť α, kde $F(t_\alpha) = \alpha$. Asymptoticky platí, že $\sqrt{n}(\frac{j}{n} - \alpha)$ konverguje k 0. Potom je poriadková štatistika $X_{(j)}$ normálne rozdelená so strednou hodnotou $E[X_{(j)}] = t_\alpha$ a rozptylom $\sigma^2_{X_{(j)}} = \frac{\alpha(1-\alpha)}{f^2(t_\alpha)n}$. Ak $X \sim N(\mu, \sigma^2)$, potom $\sigma^2_{X_{(j)}} = \sigma^2 \frac{\pi^2}{24 \ln n}$ (Casella a Berger, 2002).*

Príklad 113 (rozptyl poriadkovej štatistiky) *Pomocou delta metódy odvoďte rozptyl poriadkovej štatistiky v definícii 29.*

Príklad 114 (rozptyl poriadkovej štatistiky, $X \sim N(\mu, \sigma^2)$) *Pomocou definície 29 odvoďte rozptyl poriadkovej štatistiky, ak $X \sim N(\mu, \sigma^2)$.*

Definícia 30 (stredná hodnota a rozptyl mediánu) *Stredná hodnota mediánu $X_{(\frac{n+1}{2})}$ je rovná $E[X_{(\frac{n+1}{2})}] = \widetilde{\mu}$ a rozptyl mediánu $\sigma^2_{X_{(\frac{n+1}{2})}} = \frac{1}{4f^2(\widetilde{\mu})n}$, kde n je nepárne. Ak $X \sim N(\mu, \sigma^2)$, potom $\sigma^2_{X_{(\frac{n+1}{2})}} = \sigma^2 \frac{\pi}{2n}$ (Casella a Berger, 2002).*

Príklad 115 (rozptyl mediánu) *Pomocou delta metódy odvoďte rozptyl poriadkovej štatistiky v definícii 30.*

Príklad 116 (rozptyl mediánu, $X \sim N(\mu, \sigma^2)$) *Pomocou definície 30 odvoďte rozptyl poriadkovej štatistiky, ak $X \sim N(\mu, \sigma^2)$.*

Ak má náhodná premenná X normálne rozdelenie, výpočet rozptylu mediánu sa zjednoduší, t.j. stačí poznať σ a n. Znalosť rozptylu mediánu je potrebná na výpočet $100 \times (1 - \alpha)\%$ intervalu spoľahlivosti pre medián (pozri kapitolu 4 Testovanie hypotéz). Popis rozdelenia mediánu, ak n je párne, je nad rámec tejto knihy.

Rozpätie náhodného výberu je vzdialenosť medzi najmenšou a najväčšou poriadkovou štatistikou, t.j. $R = X_{(n)} - X_{(1)}$.

Definícia 31 (hodnoty distribučnej funkcie v kvantiloch) *Empirická distribučná funkcia $F_n(x)$ je definovaná nasledovne*

$$F_n(x) = \begin{cases} 0, & ak\ x < X_{(1)}, \\ \frac{i}{n}, & ak\ X_{(i)} \le x < X_{(i+1)}, \\ 1, & ak\ x \ge X_{(n)}. \end{cases}$$

*Majme transformáciu $T_{(1)} = F_n(X_{(1)})$, $T_{(2)} = F_n(X_{(2)}), \ldots, T_{(n)} = F_n(X_{(n)})$. Potom $T_{(1)}$, $T_{(2)}$, $\ldots, T_{(n)}$ sú **poriadkové štatistiky**. Potom platí*

$$\lim_{n\to\infty} \Pr(\sup_{\forall x \in \mathcal{Y}} [F_n(x) - F(x)] n^{1/2} \le \lambda) = \Phi(\lambda),$$

kde $F(X)$ je teoretická distribučná funkcia a $\Phi(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2\lambda^2}$. Potom $100 \times (1 - \alpha)\%$ pás spoľahlivosti pre $F(x)$ definujeme ako $F_n(x) \pm \lambda_\alpha 1/n^{1/2}$, kde $\Phi(\lambda_\alpha) = 1 - \alpha$ a $Var[F_n(x)] = 1/n$ (Kolmogorov, 1933; Smirnov, 1933; Wilks, 1948). Potom môžeme tvrdiť, že $F(X)$ patrí do $100 \times (1 - \alpha)\%$ pásu spoľahlivosti a zároveň je medzi nulou a jednotkou s pravdepodobnosťou $1 - \alpha$.

Príklad 117 (graf distribučnej funkcie a jej IS) *Nakreslite graf distribučnej funkcie $N(\mu, \sigma^2)$, kde $\mu = 0$ a $\sigma^2 = 1$. Do grafu dokreslite 95% pás spoľahlivosti pre $F(x)$. Jeho hranice vypočítajte pomocou simulácie pseudonáhodných čísel z $N(0, 1)$ pri $n = 50$, kde $F_n(x)$ je odhadnutá z dát. Teoretická distribučnú funkciu $\Phi(\lambda)$ naprogramujte v R použitím definície 31 alebo knižnice kolmim a funkciu pkolm().*

3.1 Charakteristiky polohy

Realizácie budeme označovať ako $x_1, x_2, \ldots, x_n$, **usporiadané realizácie** budú $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. Potom môžeme definovať nasledovné odhady charakteristík polohy (výberové charakteristiky polohy) spolu s ich anglickými ekvivalentami:

- **výberové minimum** $X_{\min}$, ktorého realizácia $x_{\min} = x_{(1)}$;
- **výberové maximum** $X_{\max}$, ktorého realizácia $x_{\max} = x_{(n)}$;
- **výberový aritmetický priemer** $\overline{X}$, ktorého realizácia $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{j=1}^{n_j} x_j f_j, n_j \leqslant n$, kde f_j sú frekvencie (počty) prislúchajúcich x_j a $n = \sum_j f_j$;
- **výberový modus** X_{mod}, ktorého realizácia x_{mod} je najčastejšie sa vyskytujúca hodnota (pri diskrétnej premennej ide o hodnotu x, v ktorej má pravdepodobnmostná funkcia svoje maximum; pri spojitej premennej ide o hodnotu x, v ktorej má hustota svoje maximum);
- **výberový medián** $\widetilde{X}$(robustný odhad polohy), ktorého realizácia

$$\widetilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{ak } n \text{ je nepárne,} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right) & \text{ak } n \text{ je párne;} \end{cases}$$

  rozdelenie je *symetrické*, ak $\overline{x} = \widetilde{x} = x_{\text{mod}}$, rozdelenie je *pozitívne zošikmené* (pravostranne), ak $\overline{x} > \widetilde{x} > x_{\text{mod}}$ a rozdelenie je *negatívne zošikmené* (ľavostranne), ak $\overline{x} < \widetilde{x} < x_{\text{mod}}$;

- **výberové kvartily** poznáme tri
  - **výberový prvý (dolný) kvartil** Q_1, ktorého realizácia $\widetilde{x}_{0.25}$ predstavuje hodnotu, od ktorej je 1/4 dát menšia a 3/4 dát sú väčšie,
$$\Pr[x_{\min}, \widetilde{x}_{0.25}] = \Pr[X \leq \widetilde{x}_{0.25}] = \frac{1}{4}, \Pr[\widetilde{x}_{0.25}, x_{\max}] = \Pr[X \geq \widetilde{x}_{0.25}] = \frac{3}{4};$$
  - **výberový druhý kvartil** (medián) Q_2, ktorého realizácia $\widetilde{x}_{0.5} =\widetilde{x}$ je hodnota, od ktorej je 1/2 dát menšia a 1/2 dát je väčšia,
$$\Pr[x_{\min}, \widetilde{x}_{0.5}] = \Pr[X \leq \widetilde{x}_{0.5}] = \frac{1}{2}, \Pr[\widetilde{x}_{0.5}, x_{\max}] = \Pr[X \geq \widetilde{x}_{0.5}] = \frac{1}{2};$$
  - **výberový tretí (horný) kvartil** Q_3, ktorého realizácia $\widetilde{x}_{0.75}$ predstavuje hodnotu, od ktorej je 1/4 dát väčšia a 3/4 dát sú menšie,
$$\Pr[x_{\min}, \widetilde{x}_{0.75}] = \Pr[X \leq \widetilde{x}_{0.75}] = \frac{3}{4}, \Pr[\widetilde{x}_{0.75}, x_{\max}] = \Pr[X \geq \widetilde{x}_{0.75}] = \frac{1}{4};$$
- **výberové decily** $\widetilde{X}_k$, ktorých realizácie $\widetilde{x}_k$ delia súbor na desatiny, t.j. $k/10$ dát je pod decilom a $(10-k)/10$ nad decilom, kde $k \in \{0, 1, \ldots, 10\}$;
- **výberové percentily** $\widetilde{X}_p$ (čítame ako $100p$-percentil[3]), ktorých realizácie $\widetilde{x}_p$ definujeme ako

$$\widetilde{x}_p = \begin{cases} x_{(k+1)} & \text{pre } k \neq np, \\ \frac{1}{2}\left(x_{(k)} + x_{(k+1)}\right) & \text{pre } k = np, \end{cases}$$

  kde $k = \lfloor np \rfloor$, čo je celá časť čísla np (niekedy sa používa definícia cez $\{x\}$ v indexe, čo znamená najbližšie celé číslo);

- **výberový päťčíselný súhrn** $(X_{\min}, Q_1, Q_2, Q_3, X_{\max})^T$, ktorého realizáciu označujeme $(x_{\min}, \widetilde{x}_{0.25}, \widetilde{x}_{0.50}, \widetilde{x}_{0.75}, x_{\max})^T$.

[3]Výraz „$100p$-percentil“ čítame ako „stokrát p-ty percentil“. Ak $p = 0.75$, potom $100 \times 0.75 = 75$, čo čítame ako „75. percentil“.

Príklad 118 (výšky 10-ročných dievčat) *Majme výšky $n = 12$ náhodne vybraných 10-ročných dievčat v cm usporiadaných podľa veľkosti (**poradia** ozn. ako r_i pre $x_{(i)}$; pri rovnakých pozorovaniach hovoríme o **strednoporadiach**; strednoporadie sa vypočíta ako priemer poradí realizácií s rovnakou hodnotou).*

Riešenie (pozri tabuľku 3.1)

Tabuľka 3.1: Zoradné realizácie x_i a ich poradia r_i pre výšky 10-ročných dievčat

i	1	2	3	4	5	6	7	8	9	10	11	12
$x_{(i)}$	131	132	135	141	141	141	141	142	143	146	146	151
r_i	1	2	3	5.5	5.5	5.5	5.5	8	9	10.5	10.5	12

$\overline{x} \doteq 140.83$, $\widetilde{x} = \frac{1}{2}\left(x_{(6)} + x_{(7)}\right) = 141$, $\widetilde{x}_{0.25} = \frac{1}{2}\left(x_{(3)} + x_{(4)}\right) = 138$, kde $k = \lfloor 12 \times 0.25 \rfloor = 3$, $Q_3 = \widetilde{x}_{0.75} = \frac{1}{2}\left(x_{(9)} + x_{(10)}\right) = 144.5$, kde $k = \lfloor 12 \times 0.75 \rfloor = 9$.

Čo sa stane so spomínanými charakteristikami polohy, keď zmeníme mierku (škálu), napr. gramy na kilogramy alebo namiesto hmotnosti použijeme logaritmus hmotnosti?

Nech $a, b \in \mathbb{R}$ sú nejaké dané konštanty, a je posunutie a b škála. Potom $y_i = a + bx_i$ a pre priemer[4]

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n}(a + bx_i) = \frac{1}{n}\left(na + b\sum_{i=1}^{n} x_i\right) = a + b\overline{x} = \overline{a + bx}.$$

Pokiaľ $b \in \mathbb{R}^+$, usporiadanie hodnôt x_i sa pri transformácii na $y_i = a + bx_i$ nezmení, teda

$$a + bx_{(1)} \leq a + bx_{(2)} \leq ... \leq a + bx_{(n)}.$$

Pre medián v tomto prípade platí[5]

$$\widetilde{y} = a + b\widetilde{x} = \widetilde{a + bx}.$$

Ľahko sa dá nahliadnuť, že usporiadanie zachová každá rastúca funkcia $g(x)$, teda platí

$$g\left(x_{(1)}\right) \leq g\left(x_{(2)}\right) \leq ... \leq g\left(x_{(n)}\right)$$

a pre medián bude platiť $g(\widetilde{x}) \doteq \widetilde{g}(x)$. Pre nepárne n platí predchádzajúci vzťah presne, označenie „približnosti“ potrebujeme pre párne n, kde $x_{\left(\frac{n}{2}\right)} < x_{\left(\frac{n}{2}+1\right)}$. V tomto prípade je však 1/2 hodnôt $g(x_i)$ menšia ako $\widetilde{g}(x)$. Teda špeciálne môžeme medián logaritmu (napr. hmotnosti) spočítať ako logaritmus mediánu (napr. hmotnosti). Pokiaľ dôjde v pozorovaniach k posunutiu, dôjde k rovnakému posunutiu aj v charakteristike polohy. Ak zmeníme mierku, potom stačí urobiť rovnakú úpravu aj u charakteristiky polohy.

Robustnou charakteristikou strednej hodnoty (odolnejšou na odľahlé pozorovania) je (Tukey, 1962)

- **výberový γ-urezaný aritmetický priemer** $\overline{X}_g$, ktorého realizáciou je $\overline{x}_g$ a vypočíta sa ako

$$\overline{x}_g = \frac{1}{n - 2g}\left(x_{(g+1)} + x_{(g+2)} + \ldots + x_{(n-g)}\right),$$

  kde $g = \{\gamma n\}, g = \lfloor \gamma n \rfloor, \gamma = 0.1, 0.2$. Viac ako $\gamma 100$ % pozorovaní[6] musí byť nahradených, aby sa tento priemer zmenil na malý alebo veľký v porovnaní s pôvodným [[7] *breakdown point* $\overline{x}_g$ je teda γ],

[4] Rovnosť znamená, že priemer posunutej a preškálovanej veličiny y je rovný posunutému a preškálovanému priemeru pôvodnej veličiny x.

[5] Rovnosť znamená, že medián posunutej a preškálovanej veličiny y je rovný posunutému a preškálovanému mediánu pôvodnej veličiny x.

[6] Výraz „$\gamma 100$ % pozorovaní“ čítame ako „gama krát stopercent pozorovaní“.

[7] *Breakdown point* hovorí o počte pozorovaní, ktoré potrebujeme na to, aby sme výrazne zmenili hodnotu charakteristiky polohy. Pre γ-urezaný a γ-winsorizovaný aritmetický priemer ide o γn pozorovaní, pre medián ide o $n/2$ pozorovaní a pre aritmetický priemer stačí iba jedno pozorovanie (preto hovoríme, že aritmetický priemer je veľmi citlivý na odľahlé pozorovania).

- **výberový γ-winsorizovaný priemer** $\overline{X}_w$, ktorého realizácia $\overline{x}_w$ je definovaná ako

$$\overline{x}_w = \frac{1}{n}\left((g+1)x_{(g+1)} + x_{(g+2)} + \ldots + (g+1)x_{(n-g)}\right).$$

Viac ako $\gamma 100$ % pozorovaní musí byť nahradených, aby sa tento priemer zmenil na malý alebo veľký v porovnaní s pôvodným [*breakdown point* $\overline{x}_w$ je teda γ].

3.2 Charakteristiky variability

Definujeme nasledovné základné odhady charakteristík variability (výberové charakteristiky variability) spolu s ich anglickými ekvivalentami:

- **výberový rozptyl** S^2, ktorého realizáciou je

$$s^2 = s_{n-1}^2 = s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2;$$

pri lineárnej transformácii sa rozptyl mení nasledovne[8]

$$s_y^2 = s_{a+bx}^2 = b^2 s_x^2,$$

t.j.

$$\begin{aligned} s_y^2 &= s_{a+bx}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(a + bx_i - \overline{a+bx}\right)^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\left(a + bx_i - (a + b\overline{x})\right)^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\left(b\left(x_i - \overline{x}\right)\right)^2 = b^2 s_x^2; \end{aligned}$$

výpočtová podoba rozptylu

$$s_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) = \frac{1}{n-1}\left(\sum_{j=1}^{n_j} x_j^2 f_j - n\overline{x}^2\right), n_j \leq n,$$

kde f_j sú frekvencie (počty) prislúchajúcich x_j a $n = \sum_j f_j$;

- **výberová smerodajná odchýlka** S, ktorej realizáciou je

$$s = s_{n-1} = s_x = \sqrt{s_x^2};$$

pri lineárnej transformácii sa smerodajná odchýlka mení nasledovne[9]

$$s_y = s_{a+bx} = |b|\, s_x,$$

teda, ak pripočítame ku všetkým pozorovaniam rovnakú konštantu, miera variability sa nezmení; zmena mierky (u pomerovej mierky zmena jednotiek) má za následok rovnakú úpravu jednotlivých pozorovaní i miery variability v podobe smerodajnej odchýlky;

- **výberový koeficient variácie** V_k, ktorého realizácia v_k predstavuje normalizovanú podobu výberového rozptylu (inverzia *signal-to-noise ratio*; podiel variability na priemere)

$$v_k = \frac{s_x}{\overline{x}};$$

[8] Rovnosť znamená, že rozptyl posunutej a preškálovanej veličiny y je rovný násobku druhej mocniny škály a rozptylu pôvodnej veličiny x.

[9] Rovnosť znamená, že smerodajná odchýlka posunutej a preškálovanej veličiny y je rovná násobku absolútnej hodnoty škály a smerodajnej odchýlky pôvodnej veličiny x.

bezrozmerná veličina, zvyčajne vyjadrovaná v percentách, t.j. $100 \times (s_x/\overline{x})$ % a môže sa používať len pre realizácie, ktorých rozsah nadobúda kladné hodnoty; používa sa pri porovnávaní variability súborov s nerovnakými priemermi (napr. pri porovnaní variability výšky detí určitého veku s výškou dospelých určitého veku alebo pri porovnaní variability premenných meraných v rôznych jednotkách);

- **výberový rozptyl aritmetického priemeru** $S^2_{\overline{X}}$, ktorého realizáciou je

$$s^2_{\overline{x}} = \frac{s^2_x}{n};$$

- **výberová stredná chyba priemeru** (**štandardná chyba**) $S_{\overline{X}}$, ktorej realizáciou je

$$s_{\overline{x}} = \frac{s_x}{\sqrt{n}};$$

- **výberový koeficient šikmosti** B_1, ktorého realizáciou je

$$b_1 = \frac{n^{-1}\sum_{i=1}^n (x_i - \overline{x})^3}{\left[n^{-1}\sum_{i=1}^n (x_i - \overline{x})^2\right]^{3/2}} = \frac{\sqrt{n}\sum_{i=1}^n (x_i - \overline{x})^3}{\left[\sum_{i=1}^n (x_i - \overline{x})^2\right]^{3/2}},$$

kde rozdelenie je *symetrické*, ak $b_1 = 0$, *pozitívne zošikmené* (hustota na ľavej strane stúpa strmšie ako na pravej), ak $b_1 > 0$ a *negatívne zošikmené* (hustota na pravej strane stúpa strmšie ako na ľavej), ak $b_1 < 0$;

- **výberový koeficient špicatosti** B_2, ktorého realizáciou je

$$b_2 = \frac{n^{-1}\sum_{i=1}^n (x_i - \overline{x})^4}{\left[n^{-1}\sum_{i=1}^n (x_i - \overline{x})^2\right]^{2}} - 3 = \frac{n\sum_{i=1}^n (x_i - \overline{x})^4}{\left[\sum_{i=1}^n (x_i - \overline{x})^2\right]^{2}} - 3,$$

kde rozdelenie je normálne (*mezokurtické*), ak $b_2 = 0$, špicaté (*leptokurtické*), ak $b_2 > 0$ a ploché (*platykurtické*), ak $b_2 < 0$;

- **výberová suma štvorcov** $SS = \sum_{i=1}^n \left(X_i - \overline{X}\right)^2$, ktorej realizácia je

$$SS_{\text{obs}} = \sum_{i=1}^n (x_i - \overline{x})^2,$$

kde sa tento čitateľ rozptylu používa napr. v lineárnom regresnom modeli, v modeli ANOVA a pod.;

- **výberová suma absolútnych odchýlok** $SAD = \sum_{i=1}^n |X_i - \widetilde{X}_{0.5}|$, ktorej realizácia je

$$SAD_{\text{obs}} = \sum_{i=1}^n |x_i - \widetilde{x}_{0.5}|;$$

- **výberový priemer absolútnych odchýlok** $MAD = \frac{1}{n}\sum_{i=1}^n |X_i - \widetilde{X}_{0.5}|$, ktorého realizácia je

$$MAD_{\text{obs}} = SAD_{\text{obs}}/n;$$

- **výberové rozpätie** $D = X_{\max} - X_{\min}$, ktorého realizáciou je

$$d_{\text{obs}} = x_{\max} - x_{\min};$$

- **výberové medzikvartilové rozpätie** $D_Q = Q_3 - Q_1$, ktorého realizáciou je

$$d_Q = \widetilde{x}_{0.75} - \widetilde{x}_{0.25};$$

kde rozdelenie je (medzi kvartilmi) *symetrické*, ak $\widetilde{x}_{0.75}-\widetilde{x}_{0.50}=\widetilde{x}_{0.50}-\widetilde{x}_{0.25}$, *pozitívne zošikmené*, ak $\widetilde{x}_{0.75}-\widetilde{x}_{0.50}>\widetilde{x}_{0.50}-\widetilde{x}_{0.25}$ a *negatívne zošikmené*, ak $\widetilde{x}_{0.75}-\widetilde{x}_{0.50}<\widetilde{x}_{0.50}-\widetilde{x}_{0.25}$;

- **výberové decilové rozpätie** $D_D=\widetilde{X}_{0.9}-\widetilde{X}_{0.1}$, ktorého realizáciou je
$$d_D=\widetilde{x}_{0.9}-\widetilde{x}_{0.1};$$
- **výberové percentilové rozpätie** $D_P=\widetilde{X}_{0.99}-\widetilde{X}_{0.01}$, ktorého realizáciou je
$$d_P=\widetilde{x}_{0.99}-\widetilde{x}_{0.01}.$$

Robustnými charakteristikami variability sú (Tukey, 1962)

- **výberový γ-urezaný rozptyl** S_g^2, ktorého realizácia s_g^2 sa vypočíta ako
$$s_g^2=\frac{1}{n-2g-1}\sum_{i=g+1}^{n-g}x_{(i)};$$
viac ako $\gamma 100$ % pozorovaní musí byť nahradených, aby sa tento rozptyl zmenil na veľký v porovnaní s pôvodným s^2 [*breakdown point* s_g^2 je γ]; platí $s_g^2<s^2$ pretože urezanie odstráni odľahlé hodnoty;
- **výberový γ-winsorizovaný rozptyl** S_w^2, ktorého realizáciu označujeme ako s_w^2; viac ako $\gamma 100$ % pozorovaní musí byť nahradených, aby sa tento rozptyl zmenil na veľký v porovnaní s pôvodným s^2 [*breakdown point* s_w^2 je γ]; platí $s_w^2<s^2$ pretože winsorizovanie priťahuje extrémne hodnoty bližšie k priemeru;
- **výberový kvartilový koeficient variácie** $V_{k,Q}=(Q_3-Q_1)/Q_2$, ktorého realizáciu $v_{k,Q}$ vypočítame ako
$$v_{k,Q}=\frac{\widetilde{x}_{0.75}-\widetilde{x}_{0.25}}{\widetilde{x}}.$$

Ďalšie robustné charakteristiky variability (výberové rozpätie) charakterizované pomocou upravených hraníc sú

- **výberové robustné minimum a maximum** („vnútorné hradby“) $X^*_{\min}=B_D=Q_1-1.5D_Q$ a $X^*_{\max}=B_H=Q_1+1.5D_Q$, ktorých realizácie sú definované ako
$$x^*_{\min}=b_D=\widetilde{x}_{0.25}-1.5\left(\widetilde{x}_{0.75}-\widetilde{x}_{0.25}\right),$$
$$x^*_{\max}=b_H=\widetilde{x}_{0.75}+1.5\left(\widetilde{x}_{0.75}-\widetilde{x}_{0.25}\right),$$
kde prvky vybočujúce z hradieb sa považujú za *podozrivé, potencionálne odľahlé pozorovania*;
- **výberové robustné minimum a maximum** („vonkajšie hradby“) definované ako $B^*_H=Q_1-3(Q_3-Q_1)$, $B^*_H=Q_3+3(Q_3-Q_1)$, ktorých realizácie sú $b^*_D=\widetilde{x}_{0.25}-3d_Q$, $b^*_H=\widetilde{x}_{0.75}+3d_Q$;
  - pokiaľ sú nejaké $x_i<b^*_D\vee x_i>b^*_H$, hovoríme, že ide o *vzdialené body*[10],
  - ak $x_i\in\langle b^*_D,b_D)\vee(b_H,b^*_H\rangle$, ide o *body vonkajšie*,
  - ak $x_i\in\langle b_D,b_H\rangle$, ide o *body vnútorné* alebo *body priľahlé mediánu*;
  - pre normálne rozdelenie platí $B_H-B_D=Q_3+1.5D_Q-Q_1+1.5D_Q=4D_Q\doteq 4.2$; pravdepodobnosť, že $x_i\notin\langle B_D,B_H\rangle$ je potom 0.04;
- **výberové robustné miery šikmosti** B_{1Q} a B_{1O} a ich rozptyly za podmienky asymptotickej normality $B_{1\cdot}$, kde $\cdot=Q$ alebo O, ktorých realizácie sú definované nasledovne
  - kvartilový koeficient šikmosti
$$b_{1Q}=\frac{(\widetilde{x}_{0.75}-\widetilde{x}_{0.50})-(\widetilde{x}_{0.50}-\widetilde{x}_{0.25})}{\widetilde{x}_{0.75}-\widetilde{x}_{0.25}}, Var_{as}(b_{1Q})=1.84,$$

[10] Ozn. $\vee$ znamená „alebo“ a ozn. $\wedge$ znamená „a súčasne“.

– oktilový koeficient šikmosti

$$b_{1O} = \frac{(\widetilde{x}_{0.875} - \widetilde{x}_{0.50}) - (\widetilde{x}_{0.50} - \widetilde{x}_{0.125})}{\widetilde{x}_{0.875} - \widetilde{x}_{0.125}}, Var_{as}(b_{1O}) = 1.15.$$

3.3 Detekcia odľahlých pozorovaní

Homogénny náhodný výber je taký výber, v ktorom všetky $x_i, i = 1, 2, \ldots, n$, sú realizácie rovnakého rozdelenia pravdepodobnosti s konštantným rozptylom σ^2. K **nehomogenitám výberu** dochádza všade tam, kde sa vyskytujú výrazné nerovnomernosti v realizáciách, náhle sa menia podmienky experimentu a pod. Nehomogenita môže byť spôsobená aj nevhodne zvoleným výberom subjektov.

Špeciálnym prípadom ovplyvňujúcim homogenitu výberu sú **odľahlé pozorovania** (outliers). Takéto pozorovania skresľujú odhady polohy (špeciálne aritmetického priemeru) a variability (hlavne rozptylu), takže môžu znehodnotiť ďalšiu štatistickú analýzu. Pri ich overovaní sa používa mnoho idealizovaných predpokladov. Musíme poznať ich predpokladaný počet, ich rozdelenie a tiež rozdelenie ostatných prvkov náhodného výberu. Navyše je nutné zostrojiť štatistický alebo pravdepodobnostný model, podľa ktorého sa odľahlé pozorovania chovajú. Testovanie odľahlých pozorovaní bez doplnkových informácií je teda málo spoľahlivé.

Jednoduchou technikou, kedy sa predpokladá, že dáta majú normálne rozdelenie, je **modifikácia vnútorných hradieb** b_D a b_H na

$$b_D^{mod} = \widetilde{x}_{0.25} - kd_Q, b_H^{mod} = \widetilde{x}_{0.75} + kd_Q,$$

kde sa parameter k volí tak, aby pravdepodobnosť $\Pr(n, k)$ bola dostatočne vysoká, napr. 0.95. $\Pr(n, k)$ je pravdepodobnosť, že žiaden prvok z náhodného výberu z normálneho rozdelenia s rozsahom n nebude mimo intervalu $I = \langle b_D^{mod}, b_H^{mod} \rangle$. Ak $\Pr(n, k) = 0.95$ a $n \in \langle 8, 100 \rangle$ použijeme aproximáciu $k \approx 2.25 - 3.6/n$. Teda prvky mimo I sa považujú za odľahlé. Postup spomenutý vyššie je dostatočne robustný.

Definujme mieru rozptylu ako **kvantilovú odchýlku** $d_{Q^*} = 2d_Q$. Ak urobíme štandardizáciu, dostaneme $d_{Q^*_{st}} = 1$ a **štandardizovaný medián** bude $\widetilde{x}_{st} = 0$ a **štandardizovaná kvantilová funkcia** indikujúca tvar bude (Meloun a Militký, 2004)

$$q_{st}(p) = \frac{\widetilde{x}_p - \widetilde{x}_{0.5}}{d_{Q^*}}.$$

Hodnoty kvantilov, pre ktoré platí $|q_{st}(p)| \geq 1$, sú považované za vybočujúce (pre normálne rozdelenie) a hovoríme potom o **identifikátoroch dlhých „chvostov“** (koncov). Hodnoty $q_{st}(p)$ môžeme použiť na

– identifikáciu miery šikmosti $SQ_{\text{obs}} = q_{st}(0.25) + q_{st}(0.75)$, kedy je rozdelenie pravdepodobnosti symetrické medzi kvartilmi, ak SQ_{obs} je rovné nule,

– identifikáciu dĺžky koncov, kedy

- $q_{st}(0.95) < 0.5$ hovorí o krátkych koncoch,
- $q_{st}(0.95) > 1$ hovorí o dlhých koncoch a
- pre stredne dlhé konce bude platiť $q_{st}(0.95) \in \langle 0.5, 1.0 \rangle$.

Okrem vonkajších a vnútorných hradieb, môžeme odľahlé pozorovania jednoducho detegovať aj nasledovne (Rousseeuw a van Zomeren, 1990)

- $|x - \overline{x}| > 2s$,
- $\widehat{\sigma} = MAD_{\text{obs}}/0.6745$, $|x - \widetilde{x}_{0.5}| > k\frac{MAD_{\text{obs}}}{0.6745}$ (ako k sa najčastejšie používa 2 alebo 2.24).

Príklad 119 (odľahlé pozorovania; dĺžka kľúčnej kosti) *Nech náhodná premennná X najväčšia dĺžka kosti kľúčnej z pravej strany (cla.L; dáta: more-samples-variances-clavicle.txt) má asymptoticky normálne rozdelenie, t.j. $X \sim N(\mu, \sigma^2)$. Potom identifikujte odľahlé pozorovania pomocou (a) vnútorných hradieb b_D a b_H, (b) modifikovaných vnútorných hradieb b_D^{mod} a b_H^{mod}, (c) identifikujte dlhé „chvosty“ rozdelenia tejto premennej na základe štandardizovanej kvantilovej funkcie.*

3.4 Z-skóre

V antropológii nás veľmi často zaujímajú normované realizácie, nazývané tiež normované veličiny, a to z-**skóre**, ktoré definujeme ako

$$z_i = \frac{x_i - \overline{x}}{s_x}.$$

Dostaneme ich ako špeciálny prípad lineárnej transformácie $y = a + bx$, kde voľbou $b = 1/s_x$ a $a = -\overline{x}/s_x$. Potom aritmetický priemer z-skóre

$$\overline{z} = -\frac{\overline{x}}{s_x} + \frac{1}{s_x}\overline{x} = 0$$

a rozptyl z-skóre

$$s_z^2 = \left(\frac{1}{s_x}\right)^2 s_x^2 = 1,$$

Pomocou z-skóre môžeme vyjadriť aj koeficient šikmosti a špicatosti, kde

$$b_1 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{s_x}\right)^3 = \frac{1}{n}\sum_{i=1}^{n} z_i^3,$$

a za podmienky asymptotickej normality

$$E\left[b_1\right] = 0, Var\left[b_1\right] = \frac{n-2}{(n+1)(n+3)};$$

$$b_2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{s_x}\right)^4 = \frac{1}{n}\sum_{i=1}^{n} z_i^4$$

a za podmienky asymptotickej normality

$$E\left[b_2\right] = 3 - \frac{6}{n+1}, Var\left[b_2\right] = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}.$$

Pokiaľ dáta pochádzajú z normálneho rozdelenia, budú mať oba koeficienty hodnoty približne nulové (pri b_2 po odčítaní konštanty 3).

Musíme si uvedomiť, že prítomnosť odľahlých pozorovaní v realizáciách ovplyvňuje výpočet priemeru a rozptylu, ktoré sú potrebné na výpočet z-skóre, a následne tak ovplyvňuje tiež hodnotu vlastného z-skóre. Ak je rozdelenie dát zošikmenené alebo je normalita porušená inak, nebude z-skóre odrážať situáciu vierohodne a jeho ďalšie použitie je problematické. Ak predpokladáme, že je rozdelenie znaku v populácii normálne, obmedzuje sa použie z-skóre napr. na zistenie, či nejaké pozorovanie (pacient) patrí svojimi charakteristikami do zdravej populácie.

Príklad 120 (z-skóre; šírka lebky) *Majme náhodnú premennú X šírka lebky (skull.B; mm; dáta: one-sample-mean-skull-mf.txt) u mužov. Za predpokladu asymptotickej normality X, t.j. $X \sim N(\mu, \sigma^2)$, vypočítajte z-skóre tejto premennej pomocou funkcie mean() a sd(). Výsledok skontrolujte pomocou funkcie scale().*

3.5 Príklady na charakteristiky polohy a variability

Príklad 121 (argument minima) *Vygenerujte pseudonáhodné čísla* $X \sim N(\mu, \sigma^2)$, *kde* $\mu = 0, \sigma^2 = 1$ *a* $n = 1000$. *Vygenerované čísla ozn.* $x_i, i = 1, 2, \ldots, 1000$. *Nájdite numericky také* c, *ktoré minimalizuje (a) sumu štvorcov odchýlok* $\sum_{i=1}^{1000}(x_i - c)^2$, *t.j.* $c_1 = \arg\min_{\forall c} \sum_{i=1}^{1000}(x_i - c)^2$ *a (b) sumu absolútnych odchýlok* $\sum_{i=1}^{100}|x_i - c|$, *t.j.* $c_2 = \arg\min_{\forall c} \sum_{i=1}^{1000}|x_i - c|$. *Za* c *dosadzujte postupne (1) všetky* $x_{(j)}$ *(*$x_{(j)}$ *sú usporiadané* x_i *podľa veľkosti od najmenšieho po najväčšie) a vybrané charakteristiky polohy ako (2) aritmetický priemer, (3) nejaké kvantily* $\widetilde{x}_p$, *kde* $p \in \langle 0, 1\rangle$ *a pod. Nakreslite obrázok závislosti (a) sumy štvorcov odchýlok na* $x_{(j)}$, *t.j. body* $[x_j, y_j]$, *kde* $y_j = \sum_{i=1}^{1000}(x_i - x_{(j)})^2$ *a (b) sumy absolútnych odchýlok na* $x_{(j)}$, *t.j. body* $[x_{(j)}, y_j]$, *kde* $y_j = \sum_{i=1}^{1000}|x_i - x_{(j)}|$. *Podobné obrázky nakreslite aj pre* $\widetilde{x}_p$ *namiesto* $x_{(j)}$.

Príklad 122 (výšky 10-ročných dievčat, pokrač.) *Vypočítajte základné charakteristiky polohy a variability.*

Riešenie v (pozri tabuľku 3.2 a 3.3)
Dáta:

```
x <- c(131, 132, 135, 141, 141, 141, 141, 142, 143, 146, 146, 151)
```

Minimum, maximum, medián, aritmetický priemer, prvý kvartil, druhý kvartil, tretí kvartil, kvartily (pomocou jednej funkcie), päťčíselný súhrn, rozptyl a smerodajná odchýlka:

```
min(x) # 131
max(x) # 151
median(x) # 141
mean(x) # 140.8333
priemer <- sum(x)/length(x) # 140.8333
q1 <- quantile(x,0.25,type=2) # 138
q2 <- quantile(x,0.50,type=2) # 141
q3 <- quantile(x,0.75,type=2) # 144.5
quantile(x,c(0.25,0.5,0.75),type=2) # 138.0 141.0 144.5
quantile(x,c(0,0.25,0.5,0.75,1),type=2) # 131.0 138.0 141.0 144.5 151.0
var(x) # 33.78788
sd(x) # 5.812734
```

Funkcie na výpočet rozptylu, smerodajnej odchýlky, štandardnej chyby, šikmosti a špicatosti:

```
"rozptyl" <- function(x) sum((x-mean(x))^2)/(length(x)-1)
"smerodch" <- function(x) sqrt(var(x))
"SE" <- function(x) sqrt(var(x) / length(x))
SE(x) # 1.677992
"sikmost" <- function(x) {(1/length(x))*sum(((x-mean(x))/(sqrt(var(x))))^3)}
sikmost(x) # -0.2121993
"spicatost" <- function(x) {(1/length(x))*sum(((x-mean(x))/
                          (sqrt(var(x))))^4) - 3}
spicatost(x) # -0.9029347
```

Suma štvorcov, priemerná absolútna odchýlka (priemer absolútnych odchýlok), suma absolútnych odchýlok, rozpätie, medzikvartilové rozpätie, zisťovanie symetrie, robustný výpočet minima a maxima („vnútorné hradby“), robustné rozpätie:

```
var(x)*(length(x)-1) # 371.6667
mad(x) # 5.1891
mad(x)*length(x) # 62.2692
c(min(x),max(x)) # 131 151
range(x) # 131 151
Dq <- quantile(x,0.75,type=2)-quantile(x,0.25,type=2) # 6.5
c(q3 - q2, q2 - q1) # 3.5 3.0
Bd <- q1-1.5*Dq; Bh <- q3+1.5*Dq
c(Bd,Bh) # 128.25 154.25
```

Funkcia na výpočet kvartilovej šikmosti:

```
"kvart.sikmost" <- function(x) {
                    q1 <- quantile(x,0.25,type=2)
                    q2 <- quantile(x,0.5,type=2)
                    q3 <- quantile(x,0.75,type=2)
                    ((q3 - q2) - (q2 - q1))/(q3 - q1)
                    }
kvart.sikmost(x) # 0.07692308
```

Funkcia na výpočet oktilovej šikmosti:

```
"oktil.sikmost" <- function(x)  {
                        q125 <- quantile(x,0.125,type=2)
                        q2 <- quantile(x,0.5,type=2)
                        q875 <- quantile(x,0.875,type=2)
                        ((q875 - q2) - (q2 - q125))/(q875  - q125)
                        }
oktil.sikmost(x) # -0.2857143
```

Funkcia na výpočet výberového γ-urezaného aritmetického priemeru a rozptylu (vytvorenie dátového vektora na ich výpočet):

```
"urezanie" <- function(x, gama = 0.1){
x <- na.omit(x) # odstranenie NA, ak sa v datach nachadzaju
n <- length(x)
g.min <- floor(gama*n)  # najvacsie cele cislo mensie ako gama*n
g.max <- floor((1-gama)*n) # najvacsie cele cislo mensie ako (1-gama)*n
x.sort <- sort(x) # zoradenie podla velkosti
x.min <- x.sort[g.min] # vybratie dolnej hranice
x.max <- x.sort[g.max] # vybratie hornej hranice
xg <- x[x > x.min & x < x.max]
return(xg)
}
```

Funkcia na výpočet výberového aritmetického priemeru a rozptylu winsorizovaného pomocou „vnútorných hradieb“ (vytvorenie dátového vektora na ich výpočet):

```
"winsorizacia" <- function(x){
x <- na.omit(x) # odstranenie NA
q1 <- quantile(x,0.25)
q3 <- quantile(x,0.75)
Dq <- q3 - q1
min.x <- q1 - 1.5*Dq
max.x <- q3 + 1.5*Dq
xw <- x
for (i in 1: length(x)) if (x[i] >= max.x) xw[i] <- max.x
for (i in 1: length(x)) if (x[i] <= min.x) xw[i] <- min.x
return(xw)
}
```

Porovnanie troch typov aritmetických priemerov a rozptylov:

```
xg <- urezanie(x)
xw <- winsorizacia(x)
TAB <- rbind(c(length(x), mean(x), sd(x)),c(length(xg), mean(xg), sd(xg)),c(length(xw), mean(xw), sd(xw)))
dimnames(TAB)[[1]] <- c("surove_data","urezane_data","winsorizovane_data")
dimnames(TAB)[[2]] <- c("rozsah","priemer","sd")
TAB <- round(TAB,2)
```

Tabuľka 3.2: Rozsah, aritmetický priemer a smerodajná odchýlka pre surové, urezané a winsorizované dáta (výšky 10-ročných dievčat)

	rozsah	aritmetický priemer	smerodajná odchýlka
surové dáta	$n = 12$	$\overline{x} = 140.83$	$s^2 = 5.81$
urezané dáta	$n_g = 8$	$\overline{x}_g = 139.50$	$s_g^2 = 3.85$
winsorizované dáta	$n_w = 12$	$\overline{x}_w = 141.03$	$s_w^2 = 5.21$

Funkcia na výpočet niektorých základných charakteristík:

```
"zakl.char" <- function(x, type = 7){
# odstranenie chybajucich pozorovani
x <- x[!is.na(x)]
n <- length(x) # rozsah
# kvantily p = 0,0.25,0.5,0.75 a 1
kvantily <- quantile(x,c(0,0.25,0.5,0.75,1),type=type)
priemer <- mean(x) # priemer
SD <- sd(x) # smerodajna odchylka
StEr <- SE(x) # standardna chyba
sikm <- sikmost(x) # sikmost
spic <- spicatost(x) # spicatost
# vsetky vysledky spolu
vysledky <- c(n,priemer,kvantily,sikm,spic,SD,StEr)
# priradenie nazvov vysledkom
names(vysledky) <- c("n","priem",names(kvantily),"sik","spic","sd","se")
# zaokruhlenie na dve desatinne miesta
vysledky <- round(vysledky,2)
return(vysledky)
}
zakl.char(x, type = 2)
```

Tabuľka 3.3: Vybrané charakteristiky polohy a variability pre surové dáta (výšky 10-ročných dievčat)

charakteristika	n	$\overline{x}$	$\widetilde{x}_{\min}$	$\widetilde{x}_{0.25}$	$\widetilde{x}_{0.50}$	$\widetilde{x}_{0.75}$	$\widetilde{x}_{\max}$	b_1	b_2	s_x	$s_{\overline{x}}$
hodnota	12	140.8	131.0	138.0	141.0	144.5	151.0	−0.21	−0.90	5.81	1.68

Príklad 123 (základné charakteristiky polohy a variability) *Vypočítajte základné charateristiky polohy a variability pre premennú najväčšia dĺžka lebky (skull.L) a najväčšia šírka lebky (skull.B) u mužov; dáta: one-sample-mean-skull-mf.txt. Výsledok uložte pomocou funkcie write.table().*

Riešenie v R (pozri tabuľku 3.4)

```
DATA <- read.table("one-sample-mean-skull-mf.txt",header=TRUE)
names(DATA) ## "id"        "pop"       "sex"       "skull.L"  "skull.B"
attach(DATA)
ZCH1 <- zakl.char(skull.L[sex=="m"])
ZCH2 <- zakl.char(skull.B[sex=="m"])
ZCH <- rbind(ZCH1,ZCH2)
dimnames(ZCH)[[1]] <- c("skull.L","skull.B")
ZCH
write.table(ZCH,"skull-tab-01.txt")
```

Tabuľka 3.4: Vybrané charakteristiky polohy a variability pre najväčšiu dĺžku lebky

	n	$\overline{x}$	$\widetilde{x}_{\min}$	$\widetilde{x}_{0.25}$	$\widetilde{x}_{0.50}$	$\widetilde{x}_{0.75}$	$\widetilde{x}_{\max}$	b_1	b_2	s_x	$s_{\overline{x}}$
skull.L	217	182.04	164.00	177.00	182.00	187.00	199.00	-0.06	-0.48	6.37	0.43
skull.B	216	137.19	124.00	134.00	137.00	140.00	149.00	0.08	-0.30	4.82	0.33

Príklad 124 (šikmosť a špicatosť) *Naprogramujte funkcie na výpočet rozptylu šikmosti a špicatosti.*

Príklad 125 (základné charakteristiky polohy a variability) *Vypočítajte základné charakteristiky polohy a variability pre nasledujúce premenné:*
(a) šírko-dĺžkový index lebky vypočítaný ako podiel premenných šírka lebky (skull.B; v mm) a dĺžka lebky (skull.L; v mm; dáta: one-sample-mean-skull-mf.txt) u mužov;
(b) stranový rozdiel vertikálneho priemeru diafýzy kľúčnej kosti (simd.R a simd.L; v mm) na pravej aj ľavej strane tela (dáta: paired-means-clavicle2.txt);
(c) najväčšia výška mozgovne (skull.pH; mm) a morfologická výška tváre (face.H; mm; dáta: one-sample-correlation-skull-mf.txt).

3.6 Štatistická grafika

Pokiaľ chceme zobraziť základné a relevantné grafy (spolu s výpočtom základných charakteristík polohy a variability), hovoríme o **exploratórnej analýze dát (EDA)**; pozri Murrell (2011) a Kabacoff (2011). Grafická interpretácia výberového súboru je možná pomocou stĺpcového diagramu, spojnicového grafu (polygónu početnosti, frekvenčnej krivky a polygónu kumulatívnych početností), bodového grafu, kruhového diagramu, histogramu, empirickej distribučnej funkcie, krabicového diagramu a kvantilového diagramu.

3.6.1 Stĺpcový diagram

Stĺpcový diagram – vyjadruje číselné hodnoty pomocou obdĺždnikových stĺpcov, obyčajne v zvislej, no niekedy aj vo vodorovnej polohe; môže byť neškálovaný, t.j. v *absolútnej škále*, alebo škálovaný, t.j. v *relatívnej škále* (lepšie porovnanie v prípade sledovania viacerých súborov); špeciálnymi variantmi sú **veková pyramída** (strom života, znázorňuje vekové zloženie obyvateľstva) a **histogram**. Stĺpcový diagram nakreslíme pomocou funkcie barplot(x).

Kombinácii stĺpcových diagramov usporiadaných nad sebou pre škálované kategoriálne dáta, t.j. dáta v podobe pravdepodobností, ktorých suma je rovná jednej, sa hovorí aj **spinogram**. Spinogram nakreslíme pomocou funkcií

```
library(vcd)
spine(x)
```

Príklad 126 (stĺpcový diagram; oči vs. vlasy) *Kontingenčná tabuľka predstavuje pravdepodobnosti výskytu rôznych farieb vlasov a očí v populácii. Ide o model multinomického rozdelenia, ozn.* $Mult(\mathbf{p}, N)$*, pretože všetky pravdepodobnosti dávajú v sume jednotku. Vypočítajte početnosti všetkých buniek tabuľky, ak máme v populácii* 1000 *jedincov. Prepočítajte pravdepodobnosti na model súčinového multinomického rozdelenia (pravdepodobnosti výskytu pre dve farby očí a tri farby vlasov pozri v tabuľke 3.5). Použite základné funkcie (pozri Spector, 2008; Venables a kol., 2013; Verzani, 2005) a skontrolujte pomocou funkcií* `margin.table(oci)` *a* `prop.table(oci)`*. Nakreslite stĺpcové diagramy pre oba typy súčinových multinomických rozdelení – pre početnosti, ako aj pre pravdepodobnosti. [Marginálne súčty možno pridať pomocou funkcie* `addmargins(oci)`*; keby sme mali surové dáta (hodnoty* 0 *a* 1 *pre každý subjekt), kontingenčnú tabuľku vytvoríme pomocou funkcie* `ftable(data)`*.]*

Tabuľka 3.5: Kontingenčná tabuľka 2 × 3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov

oči/vlasy	blond	hnedé	ryšavé
modré	0.12	0.22	0.06
hnedé	0.15	0.34	0.11

Riešenie v R (pozri tabuľky 3.6 až 3.9 a obrázok 3.1)

```
modre.oci <- c(0.12,0.22,0.06) # vektor pravdepodobnosti
hnede.oci <- c(0.15,0.34,0.11) # vektor pravdepodobnosti
oci <- rbind(modre.oci,hnede.oci) # dva vektory spojene do matice
sum(oci) # celkova (totalna) suma = 1
# nazvy riadkov a stlpcov
dimnames(oci)[[1]] <- c("modre","hnede")
dimnames(oci)[[2]] <- c("blond","hnede","rysave")
round(addmargins(oci),2) # marginalne pravdepodobnosti
oci.pocty <- oci*1000 # fiktivne pocetnosti
addmargins(oci.pocty) # marginalne pocetnosti
sumy <- apply(oci.pocty,2,sum) # sumy po stlpcoch
oci.prav <- oci.pocty
oci.prav[1,] <- oci.pocty[1,]/sumy
oci.prav[2,] <- oci.pocty[2,]/sumy
apply(oci.prav,2,sum) # suma po stlpcoch = 1
round(addmargins(oci.prav,1),2) # marginalne pocetnosti po riadkoch
sumy1 <- apply(oci.pocty,1,sum) # sumy po riadkoch
oci.prav1 <- oci.pocty
oci.prav1[,1] <- oci.pocty[,1]/sumy1
oci.prav1[,2] <- oci.pocty[,2]/sumy1
oci.prav1[,3] <- oci.pocty[,3]/sumy1
apply(oci.prav1,1,sum) # suma po riadkoch = 1
round(addmargins(oci.prav1,2),4) # marginalne pocetnosti po stlpcoch
par(mfrow=c(2,2))
par(mar=c(4,2,1,1))
barplot(oci.pocty,space=0,cex.names=1.5)
barplot(t(oci.pocty),space=0,cex.names=1.5)
par(mar=c(4,2,0,0))
barplot(oci.prav,space=0,cex.names=1.5)
barplot(t(oci.prav1),space=0,cex.names=1.5)
# legend("topright",c("modre oci","hnede oci"),fill=c("black","grey")) # mozne pridanie legendy (nezobrazena)
```

Tabuľka 3.6: Kontingenčná tabuľka 2 × 3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi pravdepodobnosťami (multinomické rozdelenie)

oči/vlasy	blond	hnedé	ryšavé	suma
modré	0.12	0.22	0.06	0.40
hnedé	0.15	0.34	0.11	0.60
suma	0.27	0.56	0.17	1.00

Tabuľka 3.7: Kontingenčná tabuľka 2×3 početností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi početnosťami (multinomické rozdelenie)

oči/vlasy	blond	hnedé	ryšavé	suma
modré	120	220	60	400
hnedé	150	340	110	600
suma	270	560	170	1000

Tabuľka 3.8: Kontingenčná tabuľka 2×3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi stĺpcovými početnosťami (súčinové multinomické rozdelenie; po stĺpcoch)

oči/vlasy	blond	hnedé	ryšavé
modré	0.4444	0.3929	0.3529
hnedé	0.5556	0.6071	0.6471
suma	1.0000	1.0000	1.0000

Tabuľka 3.9: Kontingenčná tabuľka 2×3 pravdepodobností výskytu pre dve farby očí a tri farby vlasov spolu s marginálnymi riadkovými početnosťami (súčinové multinomické rozdelenie; po riadkoch)

oči/vlasy	blond	hnedé	ryšavé	suma
modré	0.3000	0.5500	0.1500	1.0000
hnedé	0.2500	0.5667	0.1833	1.0000

3.6.2 Spojnicový graf, polygón početnosti a frekvenčná krivka

Spojnicový graf – znázorňuje priebeh časového radu a jeho špeciálnymi prípadmi sú **polygón početnosti**, **frekvenčná krivka**, **polygón kumulatívnych početností**.

Polygón početnosti – spojnicový diagram, v ktorom nad stredmi triednych intervalov I_i vztýčime kolmice, s výškou úmernou príslušným triednym početnostiam a koncové body kolmíc pospájame. Ich súradnice sú $[x_i^*, n_i]$. Ide o zobrazenie priebehu početností vnútri každého intervalu, ale plocha uzavretá spojnicou polygónu nie je presne (len približne) úmerná počtu pozorovaní v intervale.

Frekvenčná krivka – vznikne, keď koncové body polygónu početnosti pospájame hladkou krivkou; vystihuje celkom presne priebeh rozdelenia početnosti a plocha v každom mieste ohraničená krivkou je priamo úmerná počtu pozorovaní.

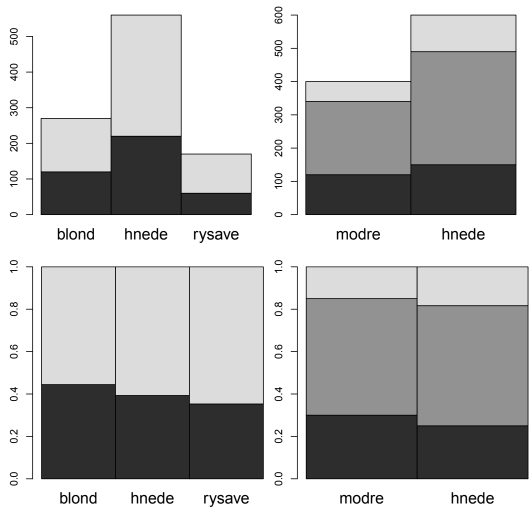
3.6.3 Bodový graf

Bodový (**rozptylový**) **graf** – zobrazuje namerané hodnoty v pravouhlej súradnicovej sústave (dvojdimenzionálnej, 2D; trojdimenzionálnej, 3D), pričom jednotlivé kategórie sa odlišujú pomocou rôznych značiek, farieb a pod.; často sa používa na zobrazenie závislosti dvoch znakov. Dvojdimenzionálny a trojdimenzionálny rozptylový graf nakreslíme pomocou funkcií

```
plot(x,y) # 2D
library(scatterplot3d)
scatterplot3d(x,y,z) # 3D
```

Argumenty funkcie **plot(x)**:

- type= argument kontrolujúci typ grafu
  - type="p" – body (prednastavená hodnota; *default*),
  - type="l" – čiary,
  - type="b" – body pospájané čiarami,
  - type="s" – schodovitá funkcia,
  - type="n" – prázdny obrázok;
- argumenty popisu osí a grafu
  - xlab="retazec" – popis osi x,
  - ylab="retazec" – popis osi y,



Obr. 3.1: Stĺpcové diagramy – početnosti (prvý riadok) pre vlasy (vľavo), pre oči (vpravo); pravdepodobnosti (druhý riadok) pre vlasy (vľavo), pre oči (vpravo)

  - main="retazec" – hlavný nadpis,
  - sub="retazec" – podnadpis pod osou x;

- farba col="anglicky.nazov" alebo kódovanie v RGB-škále (rgb(cislo1,cislo2,cislo3) vytvorí RGB-vektor z hodnôt intenzity); transformácia col2rgb(anglicky.nazov) vytvorí RGB-vektor z anglického názvu farby, kde RGB-vektor je textový vektor so 7 alebo 9 elementmi, kde za "#"nasleduje *red*, *blue* a *green* farebný kanál a voliteľne aj koeficient transparencie α v hexadecimálnej sústave (po preškálovaní na hodnoty $0, \ldots, 255$; default je "black";
- veľkosť cex=k, $k \in \mathbb{R}$, prednastavená hodnota je 1;
- typ bodov pch=k, číslo $k = 1, 2, \ldots, 20$, prednastavená hodnota je 1 (prázdny krúžok);
- typ čiar lty=k, číslo $k = 1, 2, \ldots, 20$, prednastavená hodnota je 1 (plná čiara).

Ďalšie funkcie súvisiace s bodovým grafom:

- points(x,y) – pridávanie bodov do obrázka;
- lines(x,y) – pridávanie čiar do obrázka;
- text(x, y, labels) – pridanie textu v bodoch špecifikovaných súradnicami x a y, kde popis labels[i] sa zobrazuje v bodoch (x[i],y[i]), prednastavené hodnoty sú 1:length(x);

  plot(x,y, type="n"); text(x, y, names);
- title(main,sub,xlab,ylab) – dodatočné pridanie nadpisov a popisov osí;
- legend(x,y,legend) – dodatočné pridanie legendy, špecificky umiestnenej v súradniciach x a y, kde sú prednastavené nasledovné polohy "bottomright", "bottom", "bottomleft", "left", "topleft", "top", "topright", "right" a "center"' argumenty funkcie
  - fill="retazec" – farba výplne (ne)orámovanej legendy,
  - col="retazec" – farba nakreslených bodov alebo čiar,
  - lty=k – typ čiar, $k = 1, 2, \ldots, 20$,
  - lwd=k – šírka čiar, $k \in \mathbb{N}$, prednastavená hodnota je 1 (plná čiara),

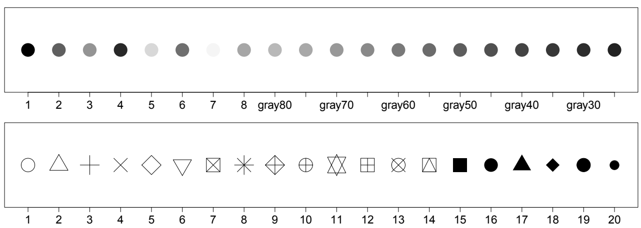
  – pch=k – typ bodov, $k = 1, 2, \ldots, 20$, prednastavená hodnota je 1 (prázdny krúžok);

- locator(n,type) – určenie polohy konkrétneho bodu v grafe (napr. odľahlé pozorovanie) pomocou jedného bodového kliknutia myšou v jeho blízkosti, pričom funkcia bod nielen označí, ale aj vypočíta jeho súradnice text(locator(1),"retazec"); použitie v legende legend(locator(1),...);
- identify(x,y,labels) identifikácia bodov, ak poznáme ich súradnice.

Príklad 127 (typy bodov a základné farby) *Nakreslite obrázok (a) základných dvadsiatich typov bodov a (b) ôsmich typov farieb a dvanástich odtieňov sivej.*

Riešenie v R (pozri obrázok 3.2)

```
windows(14,2.5)
par(mar=c(3,0.1,0.1,0.1))
plot(1:20,rep(0:1,10),type="n",sub="",xlab="",ylab="",bty="n",axes=FALSE)
points(1:20,rep(0.5,20),pch=1:20,cex=4)
axis(1,at=1:20,labels=1:20,cex.axis=1.5)
box()
windows(14,2.5)
par(mar=c(3,0.1,0.1,0.1))
plot(1:20,rep(0:1,10),type="n",sub="",xlab="",ylab="",bty="n",axes=FALSE)
points(1:8,rep(0.5,8),pch=16,col=1:8,cex=4)
siva <- paste("gray",rev(seq(25,80,by=5)),sep="")
points(9:20,rep(0.5,12),pch=16,col=siva ,cex=4)
axis(1,at=1:20,labels=c(1:8,siva),cex.axis=1.5)
box()
```



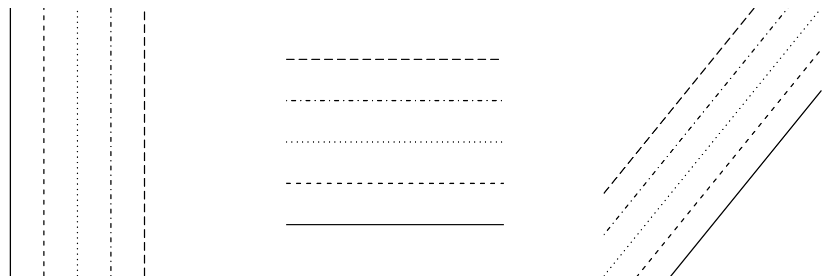
Obr. 3.2: Základné typy bodov (dolný riadok) a farieb (horný riadok)

Príklad 128 (typy bodov a základné farby) *Nakreslite obrázok zpiatich základných typov čiar, ktoré (a) smerujú zvislo, (b) smerujú vodorovne, (c) zvierajú s osou x uhol* $45°$.

Riešenie v R (pozri obrázok 3.3)

```
windows(12,4)
par(mfcol=c(1,3))
plot(1,1,type="n",xlab="",ylab="",bty="n",axes=FALSE,xlim=c(-3,3),ylim=c(-3,3))
abline(v=-2:2,lty=1:5)
plot(1,1,type="n",xlab="",ylab="",bty="n",axes=FALSE,xlim=c(-3,3),ylim=c(-3,3))
abline(h=-2:2,lty=1:5)
plot(1,1,type="n",xlab="",ylab="",bty="n",axes=FALSE,xlim=c(-3,3),ylim=c(-3,3))
abline(a=-2,b=tan(pi/4),lty=1)
abline(a=-1,b=tan(pi/4),lty=2)
abline(a=0,b=tan(pi/4),lty=3)
abline(a=1,b=tan(pi/4),lty=4)
abline(a=2,b=tan(pi/4),lty=5)
```

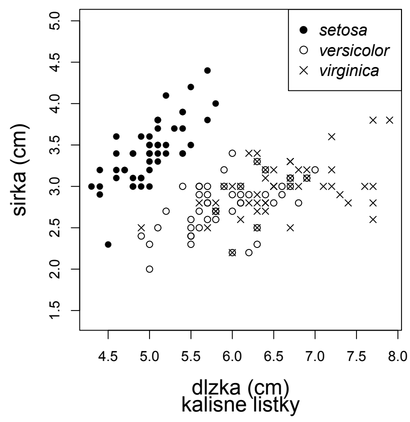
Príklad 129 (základy grafiky; dáta iris) *Nakreslite rozptylový graf dĺžky a šírky kališných lístkov pre všetky tri taxóny kosatcov pomocou (a) rôznych typov bodov a (b) rôznych farieb. Dokreslite do obrázku (c) regresné priamky pre každý taxón použitím rôzneho typu čiary. Pozri* help(iris) *ohľadom popisu premenných a ďalších detailov o dátach (Fisher, 1936/1971).*



Obr. 3.3: Základné typy čiar – zvislo, vodorovne a v uhle 45° (zľava doprava)

Riešenie v R (riešenie (a) pozri na obrázku 3.4)

```
irisDATA <- as.matrix(iris[,1:4]) # zmena datoveho ramca na maticu
dimnames(irisDATA)[[2]]
# [1] "Sepal.Length"  "Sepal.Width"
# [3] "Petal.Length"  "Petal.Width"
irisLABELS <- iris[,5]
levels(irisLABELS) # "setosa"     "versicolor" "virginica"
plot(irisDATA[,"Sepal.Length"],irisDATA[,"Sepal.Width"]) # rozptylovy graf
x.rozs <- range(irisDATA[,"Sepal.Length"]) # rozsah
y.rozs <- range(irisDATA[,"Sepal.Width"]) # rozsah
# typy bodov podla skupin
windows(5,5)
par(mar=c(5,5,1,1))
plot(irisDATA[,"Sepal.Length"],irisDATA[,"Sepal.Width"],type="n",xlab="",ylab="",asp=1)
points(irisDATA[irisLABELS=="setosa","Sepal.Length"],irisDATA[irisLABELS=="setosa","Sepal.Width"],pch=16)
points(irisDATA[irisLABELS=="versicolor","Sepal.Length"],irisDATA[irisLABELS=="versicolor","Sepal.Width"],
       pch=1)
points(irisDATA[irisLABELS=="virginica","Sepal.Length"],irisDATA[irisLABELS=="virginica","Sepal.Width"],pch=4)
legend("topright",c("setosa","versicolor","virginica"),text.font=3,pch=c(16,1,4),cex=1.2)
title(xlab="dlzka (cm)",ylab="sirka (cm)",sub="kalisne listky",cex.lab=1.5,cex.sub=1.5)
```



Obr. 3.4: Rozptylový graf pre dĺžku a šírku kališných lístkov

3.6.4 Kruhový diagram

Kruhový (výsekový, koláčový) diagram – zachytáva štruktúru dát takým spôsobom, že celá plocha kruhu predstavuje celý súbor a kruhové výseky jej jednotlivé časti, pričom polomery zvierajúce uhol 3.6° vymedzujú plochu odpovedajúcu 1 % celého obsahu. Kruhový diagram nakreslíme pomocou funkcie pie(x).

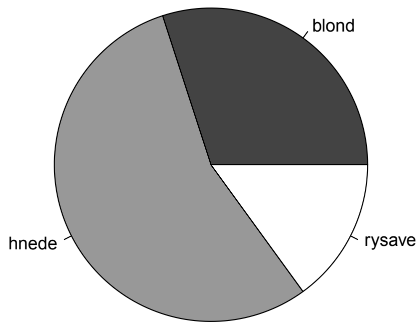
Argumenty funkcie **pie(x)**:

- x vektor relatívnych hodnôt (pravdepodobností), ktoré v súčte dávajú 1, teda i-ta položka bude x[i]/sum(x) kruhu (ale aj početností); graf začína horizontálnou čiarou doprava a pokračuje proti smeru hodinových ručičiek;
- names vektor mien prislúchajúcich jednotlivým položkám grafu;
- col vektor farieb, ktorými sú jednotlivé položky vyfarbené;
- labels="retazec" – pomenovania kruhových výsekov.

Príklad 130 (kruhový diagram) *Vytvorte kruhový diagram znázorňujúci pravdepodobnosti výskytu rôznych farieb vlasov so súčasným výskytom modrých očí (dáta z príkladu 126, tabuľka 3.9). Vytvorte rovnaký diagram v odtieňoch sivej farby (gray()).*

Riešenie v R (pozri obrázok 3.5)

```
windows(4,4)
par(mar=c(0,0,0,0))
pie(oci.prav1[1,]) # nezobrazene
pie(oci.prav1[1,], col=gray(seq(0.4,1.0, length=3)))
```

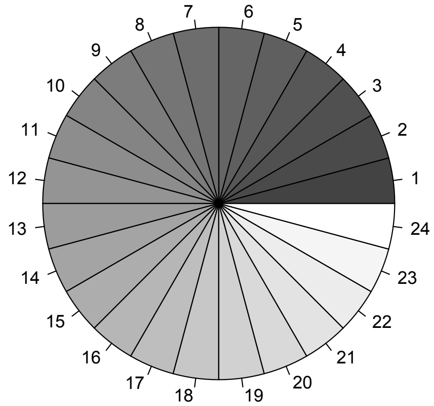


Obr. 3.5: Kruhový diagram (dáta oči vs. vlasy)

Príklad 131 (kruhový diagram) *Nakreslite tiež kruhový diagram 24 (1)* ***odtieňov sivej*** *(gray(sekvencia), kde sekvencia sú čísla z intervalu $\langle 0, 1\rangle$), (2)* ***odtieňov farieb dúhy*** *(rainbow(k)), (3)* ***teplých farieb*** *(heat.colors(k)) a (4)* ***topografických farieb*** *(topo.colors(k) alebo terrain.colors(k)).*

Riešenie v R (pozri obrázok 3.6)

```
windows(4,4)
par(mar=c(0,0,0,0))
pie(rep(1,24), col=gray(seq(0.4,1.0, length=24)), radius=0.9)
pie(rep(1,24), col=rainbow(24), radius=0.9) # nezobrazene
pie(rep(1,24), col=heat.colors(24), radius=0.9) # nezobrazene
pie(rep(1,24), col=topo.colors(24), radius=0.9) # nezobrazene
```



Obr. 3.6: Kruhový diagram (odtiene sivej)

3.6.5 Histogram

Histogram – predstavuje stĺpcový diagram s k stĺpcami, ktorých základňa sa rovná šírke intervalu $I_i = (x_i, x_{i+1}\rangle$ a výška i-teho stĺpca jeho početnosti ($i = 1, 2, \ldots, k$). Zobrazuje početnosti pozorovaní v jednotlivých intervaloch v *absolútnej škále* (na osi y sú zobrazené početnosti) a v *relatívnej škále* (obsah histogramu je rovný jednej). Histogram možno opísať pomocou *frekvenčnej tabuľky*, ktorá obsahuje početnosti a relatívne početnosti. Množstvo intervalov volí príslušný štatistický softvér alebo aj sám užívateľ. Potrebných je aspoň 12 triednych intervalov (ich počet nesmie klesnúť pod 6). Šírka jedného je minimálne $h_{\min} = 0.08(x_{\max} - x_{\min})$. Musí obsahovať minimálne 5 meraní. *Počet tried* histogramu je rovný $k = \log_2 n + 1 \doteq 2 + 3.3 \log_{10} n$ (**Sturgesova formula**; Becker a kol. (1988)), intervaly sú definované ako $\langle x_0, x_1\rangle, (x_1, x_2\rangle, \ldots, (x_{n-1}, x_n\rangle$. Šírka intervalov je teda $h = d_{\text{obs}} / (\log_2 n + 1)$ *pre realizácie z normálneho rozdelenia.* Teraz už vlastne nepracujeme s realizáciami x_i, ale so stredmi intervalov $x_i^* = (x_i + x_{i+1})/2$. Počty hodnôt n_i, ktoré sa v intervale I_i nachádzajú, sa nazývajú *triedne početnosti.* Pokiaľ realizácie nemajú normálne rozdelenie, treba použiť robustné algoritmy. Taktiež odľahlé pozorovania môžu dramaticky nafúknuť rozpätie, čo môže spôsobiť nárast šírky intervalov. Preto sa využívajú dva algoritmy ako kompromis medzi výchylkou (bias) a rozptylom realizácií pochádzajúcich z normálneho rozdelenia. Potom šírky triednych intervalov budú $h_1 = 3.49\widehat{\sigma} n^{-1/3}$, $\widehat{\sigma} = s$ (**Scottova formula**; Becker a kol. (1988)), $h_2 = 2d_Q n^{-1/3}$ (robustnejšia, **Freedman-Diaconisova formula**, ktorá je nezavislá od odľahlých pozorovaní a vyberá menšie intervaly ako Scottova formula; Venables a Ripley (2002)). Pre symetrické rozdelenia platí $h_3 = \lfloor 2\sqrt{n} \rfloor$ alebo $h_4 = \left\lfloor 2.46 \times (n-1)^{0.4} \right\rfloor$. Pokiaľ sa neočakáva príliš zošikmené rozdelenie, je šírka triednych intervalov h konštantná. V prípade komplikovanejších tvarov výberových rozdelení treba zväčšiť počet triednych intervalov alebo použiť špeciálne postupy na hľadanie nekonštantne dlhých triednych intervalov (Meloun a Militký, 2004). Histogram nakreslíme pomocou funkcie hist(x).

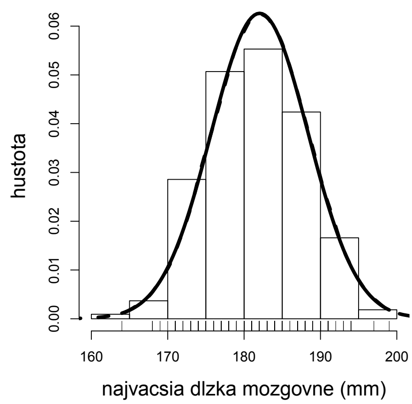
Argumenty funkcie **hist(x)**:

- prob=FALSE (prednastavená hodnota) – **v absolútnej škále**, kde na osi y sú početnosti;
- prob=TRUE – **v relatívnej škále** – suma obsahov stĺpcov (obdĺždnikov) je rovná jednej;
- breaks="Sturges" (**Sturgesova formula**, prednastavená hodnota), ďalšie možnosti sú "Scott" (**Scottova formula**) a "FD" alebo "Freedman-Diaconis" (**Freedman-Diaconisova formula**);
- nclass=k, $k \in \mathbb{N}$ – počet triednych intervalov;
- plot=FALSE – ak chceme vypísať číselné detaily, ale nechceme obrázok.

Príklad 132 (histogram a hustota normálneho rozdelenia) *Nakreslite histogram najväčšej dĺžky lebky (v mm)* skull.L *u mužov (dáta* one-sample-mean-skull-mf.txt*) a superponujte ho s (a) krivkou hustoty normálneho rozdelenia (plná čiara) a (b) krivkou hustoty vypočítanej z dát (empirická hustota; čiarkovaná čiara). Pod bázu histogramu nakrelite tzv. „koberec“, ktorý charakterizuje polohu realizácií.*

Riešenie v R (pozri obrázok 3.7)

```
DATA <- read.table("one-sample-mean-skull-mf.txt", header=TRUE)
attach(DATA)
x <- na.omit(skull.L[sex=="m"]) # odstranenie chybajuceho pozorovania NA
priemer <- mean(x)
SD <- sd(x)
# 100 kvantilov od minima po maximum x a hodnoty hustoty v nich
kvant <- seq(min(x), max(x), length=100)
# doplnenie hustoty normalneho rozdelenia
hust <- dnorm(kvant, mean = priemer, sd=SD)
hist(x, ylim=c(0, range(hust)[2]), prob=TRUE, main="", xlab="x", ylab="hustota")
lines(kvant, hust, col="red", lwd=2)
# alternativne cez MC simulaciu
x.seq <- rnorm(100000, mean=mean(x), sd=sd(x))
lines(density(x.seq), lwd=2)
# koberec
rug(x)
```

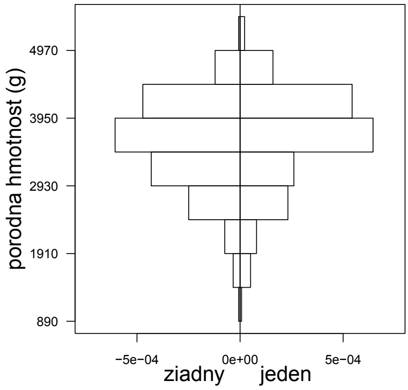


Obr. 3.7: Histogram so superponovanou krivkou hustoty normálneho rozdelenia (plná čiara) a hustoty vypočítanej z dát (empirická hustota; čiarkovaná čiara); pod histogramom je tzv. „koberec“

Príklad 133 (dva histogramy) *Nakreslite dva histogramy tak, aby sa svojimi bázami dotýkali. Aplikujte na dáta* two-samples-means-birth.txt*.*

Riešenie v R (pozri obrázok 3.8; obe krivky hustoty sa prekrývajú)

```
DATA <- read.table("two-samples-means-birth.txt", header=TRUE)
names(DATA)
# [1] "o.sib.N" "birth.W"
attach(DATA) # zmena typu objektu na faktor
o.sib.N.faktor <- as.factor(o.sib.N) # hladiny faktora
# oznacenie hladin a ich zmena
levels(o.sib.N.faktor)
o.sib.N.faktor1 <- factor(o.sib.N.faktor, labels=c("ziadny","jeden"))
rozsah <- range(birth.W)
library(Hmisc) # nacitanie kniznice
windows(5,5)
par(mar=c(4.5,4.5,1,1))
xx <- histbackback(split(birth.W, o.sib.N.faktor1), probability=TRUE, xlab="", ylab="", axes=FALSE)
title(ylab="porodna hmotnost (g)", cex.lab=1.5)
axis(1, cex.axis=0.9)
axis(2, at=seq(0,8, length=5), labels=seq(rozsah[1], rozsah[2], length=5), cex.axis=0.9, las=1)
mtext("ziadny", side=1, line=2, at=-mean(xx$left), font=1, cex=1.5)
mtext("jeden", side=1, line=2, at=mean(xx$right), font=1, cex=1.5)
```

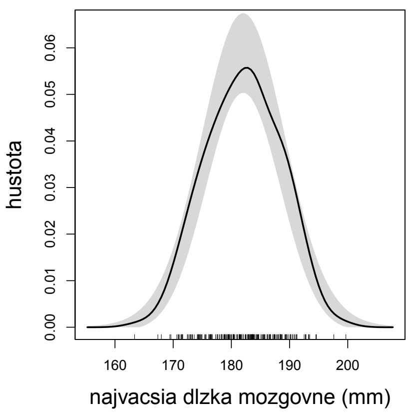


Obr. 3.8: Dva histogramy s priloženými bázami (základňami)

Príklad 134 (MC simulácia hustoty normálneho rozdelenia) *Superponujte hustotu najväčšej dĺžky lebky (v mm)* skull.L *(dáta* one-sample-mean-skull-mf.txt*) s hustotou normálneho rozdelenia vypočítanou pomocou MC simulácie s* 95% *pásom spoľahlivosti normálneho rozdelenia so strednou hodnotou rovnou aritmetickému priemeru a rozptylom rovným výberovému rozptylu (Bowman a Azzalini, 1997).*

Riešenie v R (pozri obrázok 3.9)

```
DATA <- read.table("one-sample-mean-skull-mf.txt",header=TRUE)
attach(DATA)
library(sm)
x <- na.omit(skull.L[sex=="m"])
yy <- sm.density(x, model="Normal")
windows(5,5)
par(mar=c(4.5,4.5,1,1))
sm.density(x,model="Normal",ylim=c(0,0.0655),xlab="",ylab="")
lines(yy$eval.points,yy$estimate,lwd=2)
title(xlab="najvacsia dlzka mozgovne (mm)",ylab="hustota",cex.lab=1.5)
```



Obr. 3.9: Hustota so superponovanou hustotou normálneho rozdelenia v podobe 95% pásu spoľahlivosti

3.6.6 Empirická distribučná funkcia

Histogram kumulatívnych početností (**súčtový histogram**) – namiesto početností n_j budeme nad jednotlivými intervalmi I_i zakreslovať obdĺžniky s výškou rovnajúcou sa príslušným kumulatívnym početnostiam $N_i = \sum_{j=1}^{i} n_j$. Kumulatívne relatívne početnosti definujeme ako N_i/n, čomu zodpovedá **empirická distribučná funkcia**, definovaná pre zvolené číslo x ako relatívna početnosť v intervale $(-\infty, x\rangle$, teda ako N_i-tina hodnôt x_i menších alebo rovných ako x, t.j. (Wasserman, 2006)

$$\widehat{F}_n(x) = \frac{\# x_i \leq x}{n} = \sum_{i=1}^{n} I(x_i \leq x)/n,$$

kde $I(\cdot)$ je indikačná funkcia. Obrázok empirickej distribučnej funkcie nakreslíme pomocou funkcie plot(ecdf(x),verticals = TRUE,do.points=FALSE).

Argumenty funkcie **plot(ecdf(x))**:

- verticals=FALSE je prednastavená hodnota; ak verticals=TRUE, je nakreslená schodovitá funkcia;
- do.points=TRUE je prednastavená hodnota; vkreslí do obrázka aj body, v ktorých je funkcia počítaná;
- col.01line="gray70" (prednastavená hodnota) – farba horizontálnych priamok v bodoch 0 a 1.

Príklad 135 (distribučná funkcia a hustota normálneho rozdelenia) *Nakreslite empirickú distribučnú funkciu najväčšej dĺžky lebky (v mm)* skull.L *u mužov (dáta* one-sample-mean-skull-mf.txt*) a superponujte ju s krivkou distribučnej funkcie normálneho rozdelenia červenou farbou.*

Riešenie v R (pozri obrázok 3.10)

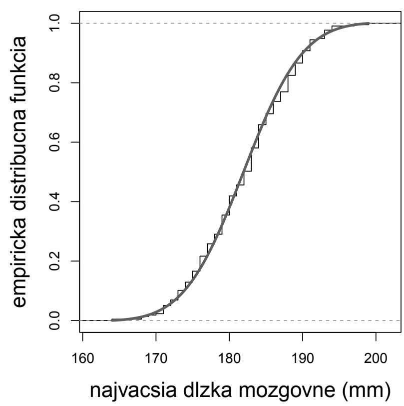
```
DATA <- read.table("one-sample-mean-skull-mf.txt",header=TRUE)
attach(DATA)
x <- na.omit(skull.L[sex=="m"])
priemer <- mean(x)
SD <- sd(x)
ROZP <- range(x)
# zoradenych 200 hodnot (ekvidistantne) medzi min a max x
x1 <- seq(ROZP[1],ROZP[2],length=200)
# teoreticka CDF (normalne rozdelenie)
y <- dnorm(x1,mean=priemer,sd=SD)
y <- cumsum(y)/sum(y)
# zobrazenie krivky distribucnej funkcie
# teoretickeho normalneho rozdelenia
windows(5,5)
par(mar=c(5,5,1,1))
plot(ecdf(x),verticals=TRUE,do.points=FALSE,xlab="najvacsia dlzka mozgovne (mm)",
      ylab="empiricka distribucna funkcia",main="",cex.lab=1.5)
lines(x1,y,col="red",lwd=3)
```

3.6.7 Krabicový diagram

Krabicový diagram – predstavuje grafické znázornenie päťčíselného súhrnu, t.j. zobrazuje v poradí zdola nahor hodnoty $x^*_{\min}, \widetilde{x}_{0.25}, \widetilde{x}_{0.50}, \widetilde{x}_{0.75}, x^*_{\max}$. Umožňuje tiež doplnenie hodnoty aritmetického priemeru a tak zvýrazniť prípadné odchýlky od normality, identifikovať symetriu rozdelenia medzi kvartilmi, symetriu rozdelenia v koncoch rozdelenia, či odhaliť odľahlé pozorovania. Ak $\widetilde{x}_{0.50} = \overline{x}$ ide o **symetrické rozdelenie**, ak $\widetilde{x}_{0.50} < \overline{x}$, ide o **pravostranne zošikmené rozdelenie**, ak $\widetilde{x}_{0.50} > \overline{x}$, ide o **ľavostranne zošikmené rozdelenie**. Často sa používa na grafické porovnanie dvoch a viacerých skupín. *Šírka krabičiek* je proporčná k odmocnine z rozsahu výberového súboru $\sqrt{n}$. Ak hovoríme o **krabicových diagramoch so zárezom**, ide o zárez charakterizujúci 95% *empirický interval spoľahlivosti* (pozri kap. Testovanie hypotéz) mediánu $\widetilde{\mu}$, ozn. (d, h), kde

$$d = \widetilde{x} - 1.57\frac{d_Q}{\sqrt{n}}, h = \widetilde{x} + 1.57\frac{d_Q}{\sqrt{n}}.$$

Odhadom rozptylu mediánu je $\widehat{\sigma}^2_{\widetilde{x}} = d_Q/1.349$, kde vo všeobecnosti platí (pre akékoľvek rozdelenie pravdepodobnosti), že $\sigma^2_{\widetilde{x}} = \frac{1}{4nf^2(\widetilde{x})}$, kde f je hustota rozdelenia pravdepodobnosti (Casella a Berger,



Obr. 3.10: Emirická distribučná funkcia (schodovitá krivka) superponovaná krivkou distribučnej funkcie normálneho rozdelenia (hladká krivka)

2002). Pre normálne rozdelenie bude platiť $\sigma^2_{\widetilde{x}} = \sigma^2_x \frac{\pi}{2n}$, kde $\widetilde{X} \sim N\left(\widetilde{\mu}, \sigma^2_{\widetilde{x}}\right)$, pozri definíciu 30. Obrázok krabicového diagramu nakreslíme pomocou funkcie boxplot(x).

Argumenty funkcie **boxplot(x)**:

- varwidth je argument relatívnej šírky jednotlivých krabičiek; prednastavená hodnota je FALSE s rovnakou šírkou všetkých krabičiek; ak ju zmeníme na TRUE, šírka krabičiek bude proporcionálna druhej odmocnine z počtu pozorovaní;
- notch=TRUE znamená zobrazenie zárezov krabičiek, ktoré odpovedajú 95% intervalom spoľahlivosti pre medián (prednastavená hodnota je FALSE);
- col určuje farbu vnútri krabičiek;
- border určuje farbu hraníc krabičiek;
- names je vektor pomenovaní pre jednotlivé zobrazované skupiny, ak ho vynecháme, použijú sa názvy z atribútu names z dátového rámca;
- pch typ bodu na zobrazenie odľahlých pozorovaní; prednastavená hodnota je 1;
- horizontal=FALSE je prednastavená hodnota;
- plot=FALSE – ak chceme iba vypísať číselné charakteristiky (prednastavená hodnota je TRUE).

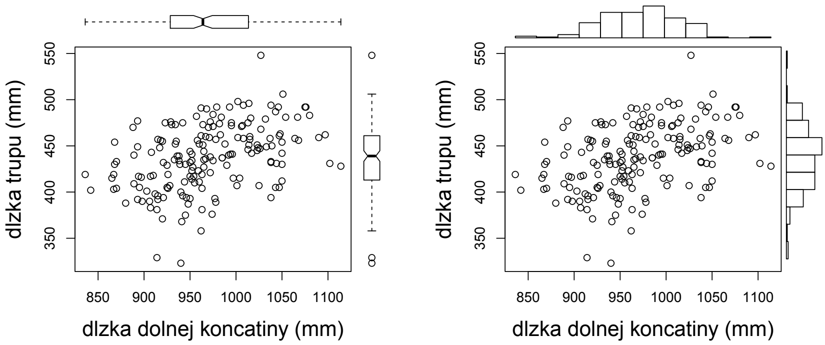
Najčastejšie používanou kombináciou argumentov je boxplot(x,varwidth=TRUE,notch=TRUE,outpch=16). Hodnoty aritmetického priemeru sa do krabicových diagramov dokresľujú pomocou príkazu points(priemer, pch=16, col="red").

Príklad 136 (bodový graf a marginálne grafy) *Naprogramujte bodový graf dvoch premenných (1) s krabicovými diagramami a (2) histogramami pre marginálne dáta. Aplikujte na dáta two-samples-correlations-trunk.txt.*

Riešenie v R (pozri obrázok 3.11)

```
DATA <- read.table("two-samples-correlations-trunk.txt",header=TRUE)
names(DATA) # [1] "sex" "lowex.L" "tru.L"
attach(DATA)
windows(5,5)
par(fig=c(0,0.9,0,0.9))
plot(lowex.L, tru.L, xlab="dlzka dolnej koncatiny (mm)",ylab="dlzka trupu (mm)")
par(fig=c(0,0.9,0.55,1),new=TRUE)
boxplot(lowex.L,horizontal=TRUE,axes=FALSE)
```

```
par(fig=c(0.65,1,0,0.9),new=TRUE)
boxplot(tru.L, axes=FALSE)

H <- hist(tru.L, plot=FALSE)
k1 <- 0
k2 <- 5
par(fig=c(0,0.9,0,0.9))
plot(lowex.L, tru.L, xlab="dlzka dolnej koncatiny (mm)", ylab="dlzka trupu (mm)")
par(fig=c(0,0.9,0.55,1), new=TRUE)
plot(NULL, type="n", ylim=c(0,max(H$counts)+k1), xlim=c(range(H$breaks)), xlab="", ylab="", main="", bty="n",
     axes=FALSE)
rect(H$breaks[1:(length(H$breaks)-1)]+k1,0+k1,H$breaks[2:length(H$breaks)]+k1,H$counts+k1)
par(fig=c(0.65,1,0,0.9),new=TRUE)
plot(NULL, type="n", xlim=c(0,max(H$counts)+k2), ylim=c(range(H$breaks)), xlab="", ylab="", main="", bty="n",
     axes=FALSE)
rect(0+k2,H$breaks[1:(length(H$breaks)-1)]+k2,H$counts+k2,H$breaks[2:length(H$breaks)]+k2)
```



Obr. 3.11: Bodový graf s marginálnymi krabicovými diagramami (vľavo) a s histogramami (vpravo)

3.6.8 Kvantilový diagram

Kvantilový diagram (*qq*-diagram) – zobrazuje body so súradnicami $\left[\Phi^{-1}\left(i/\left(n+1\right)\right), x_{(i)}\right]$, kde $\Phi^{-1}\left(p\right)$ je kvantilová funkcia normovaného normálneho rozdelenia definovaná nasledovne

$$\Pr\left(Z \leq \Phi^{-1}\left(p\right)\right) = p.$$

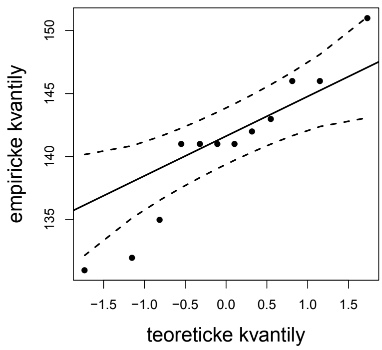
Šikmosť $b_1 > 0$ sa prejaví v podobe **konvexného** usporiadania hodnôt, $b_1 < 0$ sa prejaví v podobe **konkávneho** usporiadania hodnôt. Taktiež zviditeľňuje **dĺžku „chvostov“** (ľavého a pravého konca krivky) rozdelenia, pričom **esovitým usporiadaním** bodov sa prejavujú krátke „chvosty“ a **inverzne esovitým usporiadaním** bodov dlhé „chvosty“. Zreteľná je tiež prípadná bimodalita rozdelenia. Štatisticky možno testovať normalitu rozdelenia pomocou simulácií z $N\left(0,1\right)$ a vytvorením 95% **Atkinsonovej obálky spoľahlivosti**. Pre normálne rozdelenie bude platiť $\sigma^2_{\widetilde{x}_p} = \sigma^2_x \frac{\pi^2}{24 \ln n}$, kde $\widetilde{X}_p \sim N\left(\widetilde{\mu}_p, \sigma^2_{\widetilde{x}_p}\right)$, pozri definíciu 29.

Kvantilový diagram vytvoríme pomocou funkcie qqnorm(x) a qqline(x). Druhá funkcia dokreslí do grafu priamku prechádzajúcu bodmi charakterizujúcimi prvý a tretí kvartil teoretických a empirických kvantilov.

Príklad 137 (Atkinsonova obálka *qq*-diagramu) *Nakreslite* 95% *Atkinsonovu obálku spoľahlivosti (Atkinson, 1981; Flack a Flores, 1989). Aplikujte na dáta výšky* 10*-ročných dievčat.*

Riešenie v R (pozri obrázok 3.12)

```
library(car)
qqPlot(x, distribution="norm", xlab="kvantily N(0,1)", ylab="kvantily dat ",
          main="", envelope=.95, col.lines="red", lwd=2, pch=16, cex=1, grid=FALSE)
```



Obr. 3.12: *qq*-diagram normovanej spojitej premennej výška 10-ročných dievčat (mm) so superponovanou obálkou normálneho rozdelenia

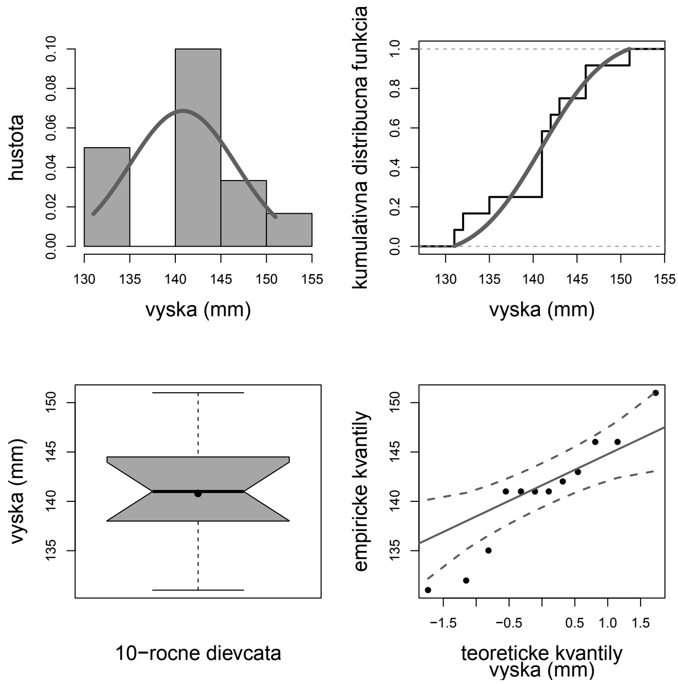
3.7 Príklady zo štatistickej grafiky

Príklad 138 (všetky základné grafy pre jeden výber ako funkcia) *Do jedného obrázka* 2×2 *naprogramujte v R štvoricu nasledovných grafov (1) histogram v relatívnej škále so superponovanou krivkou hustoty normálneho rozdelenia, (2) krabicový diagram so zakresleným priemerom, (3) empirickú (kumulatívnu) distribučnú funkciu superponovanú s teoretickou distribučnou funkciou normálneho rozdelenia a (4) qq-diagram so superponovanou qq-priamkou. Aplikujte na dáta výška* 10*-ročných dievčat.*

Riešenie v R (pozri obrázok 3.13)

```
"grafy.jeden.vyber" <- function(x,name1="vyska (mm)",name2="10-rocne dievcata"){
require(car)
par(mfcol=c(2,2),mar=c(5,5,3,1)) # nastavenie typu okna a jeho okrajov
priemer <- mean(x)
SD <- sd(x)
kvant <- seq(min(x),max(x),length=100)
hust <- dnorm(kvant,priemer,SD)
ROZP <- range(x)
y1 <- seq(ROZP[1],ROZP[2],length=200)
y <- dnorm(y1,mean=mean(y1),sd=sd(y1))
y <- cumsum(y)/sum(y)
# histogram a teoreticka hustota
hist(x,prob=TRUE,main="",xlab=name1,ylab="hustota",col="gray",cex.lab=1.5)
lines(kvant,hust,col="red",lwd=4)
# krabicovy diagram
boxplot(x,varwidth=TRUE,notch=TRUE,outpch=16,xlab=name2,ylab=name1,col="gray",cex.lab=1.5)
points(priemer,pch=16,cex=1.2)
# empiricka a teoreticka CDF
plot(ecdf(x),verticals=TRUE,do.points=FALSE,lwd=2,main="",xlab=name1,ylab="kumulativna distribucna funkcia",
     cex.lab=1.5)
lines(y1,y,col="red",lwd=4)
# kvantilovy diagram
qqPlot(x,distribution="norm",main="",xlab="",ylab="",envelope=.95,col.lines="red",lwd=2,pch=16,cex=1,grid=
     FALSE)
title(sub=name1,xlab="teoreticke kvantily",ylab="empiricke kvantily",cex.lab=1.5,cex.sub=1.5)
}
# graf
windows(7,7)
par(mar=c(5,5,1,1))
grafy.jeden.vyber(x)
```

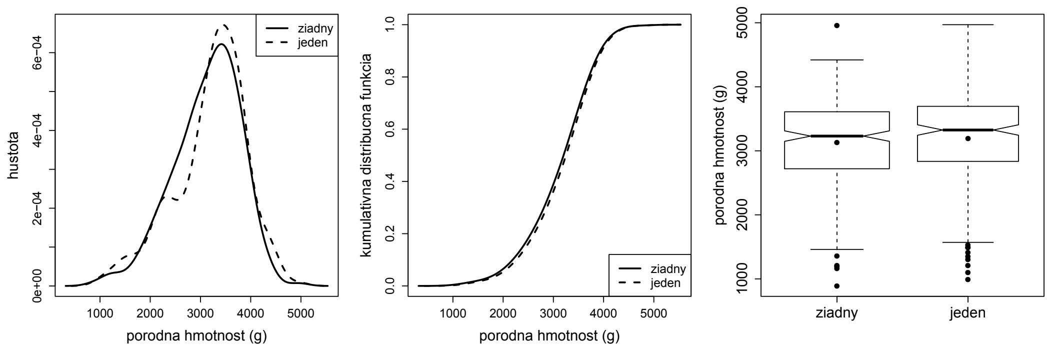
Príklad 139 (všetky základné grafy pre dva výbery ako funkcia) *V R naprogramujte do jedného obrázka* 1×3 *trojicu nasledovných grafov (1) superponované krivky hustôt, (2) superponované krivky empirických (kumulatívnych) distribučných funkcií, (3) krabicové diagramy so zakresleným priemerom. Aplikujte na dáta* two-samples-means-birth.txt*.*



Obr. 3.13: Základná štvorica grafov pre spojitú premennú výška 10-ročných dievčat (mm) – hustota (vľavo hore), kumulatívna distribučná funkcia (vpravo hore), krabicový diagram (vľavo dole), kvantilovy diagram s Atkinsonovou obálkou (vpravo dole)

Riešenie v R (pozri obrázok 3.14)

```
"grafy.dva.vybery" <- function(x1,x2,name1="ziadny",name2="jeden",name3="porodna hmotnost (g)"){
hust.1 <- density(x1)$y
x1.sekv <- density(x1)$x
CDF.1 <- cumsum(hust.1)/sum(hust.1)
hust.2 <- density(x2)$y
x2.sekv <- density(x2)$x
CDF.2 <- cumsum(hust.2)/sum(hust.2)
obe.sekv <- c(x1.sekv,x2.sekv)
obe.hust <- c(hust.1,hust.2)
obe.CDF <- c(CDF.1,CDF.2)
## dve hustoty
par(mfcol=c(1,3),mar=c(5,5,1,1))
plot(obe.sekv,obe.hust,type="n",xlab=name3,ylab="hustota",cex.lab=1.5,cex.axis=1.2)
lines(x1.sekv,hust.1,lty=1,lwd=2)
lines(x2.sekv,hust.2,lty=2,lwd=2)
legend("topright",c(name1,name2),lty=c(1,2),lwd=2,cex=1.2)
## dve CDF
plot(obe.sekv,obe.CDF,type="n",xlab=name3,ylab="kumulativna distribucna funkcia",cex.lab=1.5,cex.axis=1.2)
lines(x1.sekv,CDF.1,lty=1,lwd=2)
lines(x2.sekv,CDF.1,lty=2,lwd=2)
legend("bottomright",c(name1,name2),lty=c(1,2),lwd=2,cex=1.2)
x <- c(x1,x2)
kod <- factor(c(rep(name1,length(x1)),rep(name2,length(x2))),labels=c(name1,name2),levels=c(name1,name2))
## dva krabicove diagramy
boxplot(x~kod,ylab=name3,varwidth=TRUE,notch=TRUE,pch=16,cex=1.2,cex.lab=1.5,cex.axis=1.5)
priem <- tapply(x,kod,mean)
points(priem,pch=16,cex=1.2)
}
# grafy
DATA <- read.table("two-samples-means-birth.txt",header=TRUE)
names(DATA) # [1] "o.sib.N" "birth.W"
attach(DATA)
o.sib.N.faktor <- as.factor(o.sib.N) # zamena typu objektu na faktor
levels(o.sib.N.faktor) # kontrola hladin faktora
windows(12,4)
grafy.dva.vybery(birth.W[o.sib.N.faktor==0],birth.W[o.sib.N.faktor==1])
```



Obr. 3.14: Základná trojica grafov pre pôrodnú hmotnosť birth.W (dáta two-samples-means-birth.txt) – hustoty (vľavo), kumulatívne distribučné funkcie (uprostred), krabicové diagramy (vpravo)

Príklad 140 (stĺpcový diagram) *(a) Vypočítajte podmienené pravdepodobnosti (pre každý riadok kontingenčnej tabuľky); t.j. za predpokladu súčinového multinomického rozdelenia. Premennú v riadkoch budeme označovat X (prediktor) a premennú v stĺpcoch ako Y (závisle premenná).*
(1) Y: farba dúhovky, X: radiálne útvary v štruktúre dúhovky (dáta: multinom-iris-color.txt),
(2) Y: zakončenie troch hlavných dlaňových línií, X: farba vlasov (dáta: multinom-palmar-lines.txt),
(3) Y: priľahlosť ušného laloka, X: pohlavie (dáta: multinom-earlobe.txt),
(4) Y: krvná skupina, X: mesto (dáta: multinom-blood-groups.txt).
(b) Nakreslite stĺpcové diagramy podmienených pravdepodobností pre (1) až (4).

Príklad 141 (štatistická grafika) *Nakreslite (1) histogram v relatívnej škále so superponovanou krivkou hustoty normálneho rozdelenia, (2) krabicový diagram so zakresleným priemerom, (3) empirickú (kumulatívnu) distribučnú funkciu superponovanú s teoretickou distribučnou funkciou normálneho rozdelenia a (4) qq-diagram so superponovanou qq-priamkou pre nasledujúce premenné: (a) stranový rozdiel vertikálneho priemeru diafýzy kľúčnej kosti (simd.R a simd.L; v mm) na pravej a ľavej strane tela (dáta: paired-means-clavicle2.txt), (b) najväčšia výška mozgovne (skull.pH; v mm; dáta: one-sample-correlation-skull-mf.txt), (c) morfologická výška tváre (face.H; v mm; dáta: one-sample-correlation-skull-mf.txt).*

Príklad 142 (bodový graf a krabicové diagramy) *Nakreslite bodový graf spolu s krabicovými diagramami pre marginálne dáta pre premenné vertikálny priemer diafýzy kľúčnej kosti na pravej a ľavej strane tela (simd.R a simd.L; v mm; dáta: paired-means-clavicle2.txt).*

Príklad 143 (bodový graf a krabicové diagramy) *Nakreslite bodový graf spolu s krabicovými diagramami pre marginálne dáta pre (a) premenné najväčšia výška mozgovne (skull.pH; v mm) a morfologická výška tváre (face.H; v mm) a (b) ich z-skóre (dáta: one-sample-correlation-skull-mf.txt).*

Príklad 144 (krabicové diagramy) *Nakreslite krabicové diagramy pre obe pohlavia pre nasledovné premenné: (a) dĺžky lebky (skull.H; v mm; sex; dáta: two-samples-means-skull.txt), (b) dĺžka dolnej končatiny (lowex.L; v mm; sex; dáta two-samples-correlations-trunk.txt), (c) dĺžka trupu (tru.L; v mm; sex; dáta: two-samples-correlations-trunk.txt), (d) najväčsia dĺžka kľúčnej kosti pravej strany (cla.L; v mm; population; dáta: more-samples-variances-clavicle.txt).*

4 Testovanie hypotéz

Aby sme mohli z dát vyvodiť nejaké interpretovateľné závery, musíme najprv ciele (účely, biologicky formulované hypotézy) výskumu preformulovať do matematicko-štatistickej podoby. Takáto formulácia je dôležitá pri výbere správneho modelu na dáta. V prípade parametrického modelu máme jeho parametre θ, kde hypotézy sú o týchto parametroch. Jednou hypotézou je tzv. nulová hypotéza, čo je tvrdenie o θ, ktoré testujeme prostredníctvom nejakého **štatistického testu**. Ďalšou je alternatívna hypotéza, ktorá je doplnkom nulovej hypotézy v priestore, z ktorého pochádza θ. Na základe výsledku testu vyvodíme záver o θ. Proces testovania hypotéz nazývame aj **štatistická inferencia**, ktorej základom je **štatistická teória testovania hypotéz**.

Definícia 32 (štatistický test) *Ak parametrický priestor Θ, do ktorého patrí parameter θ modelu $\mathcal{F}$, rozdelíme na dve podmnožiny Θ_0 a Θ_1, kde $\Theta = \Theta_0 \cup \Theta_1$, potom* ***štatistický test*** $\mathcal{T} = \mathcal{T}(X)$ *predstavuje funkciu X z výberového priestoru $\mathcal{Y}$ do priestoru $\Theta_0 \cup \Theta_1$. Štatistický test je teda pravidlo výberu jednej z dvoch možností, nulovej a alternatívnej, na základe dát.* ***Nulovú hypotézu*** *definujeme ako $H_0 : \theta \in \Theta_0$ a* ***alternatívnu hypotézu*** *ako $H_1 : \theta \in \Theta_1$. Prvky množiny $\mathcal{Y}$, pre ktoré nezamietame H_0, predstavujú tzv.* ***oblasť (obor) nezamietania nulovej hypotézy****. Tieto prvky označujeme $\mathcal{Y}_0 = \mathcal{A}_T = \mathcal{A}$. V tomto prípade H_0 nezamietame. Prvky množiny $\mathcal{Y}$, pre ktoré zamietame H_0, predstavujú tzv.* ***oblasť (obor) zamietania nulovej hypotézy*** *alebo* ***kritickú oblasť****. Tieto prvky označujeme $\mathcal{Y}_1 = \mathcal{W}_T = \mathcal{W}$. V tomto prípade H_0 zamietame. Rozdelenie priestoru Θ na obor zamietania $\mathcal{W}_T$ a nezamietania $\mathcal{A}_T$ prebieha na základe* ***testovacej štatistiky*** $T(X)$.

Stretávame sa s nasledovnými štyrmi možnosťami:

A) ak platí H_0, tak naše rozhodutie je nezamietnuť H_0 (správne);

B) ak platí H_0, tak naše rozhodutie je zamietnuť H_0 (nesprávne);

C) ak neplatí H_0, tak naše rozhodutie je nezamietnuť H_0 (nesprávne);

D) ak neplatí H_0, tak naše rozhodutie je zamietnuť H_0 (správne).

V prípade B, C je naše rozhodnutie nesprávne, v prípade A, D správne. V prípade B sa dopúšťame **chyby prvého druhu** (CHPD), kde Pr(CHPD) $\leq \alpha$ a α sa nazýva **hladina významnosti**. Doplnok ku α je pravdepodobnosť $1 - \alpha$ a nazýva sa **koeficient spoľahlivosti** (prípad A). V prípade C sa dopúšťame **chyby druhého druhu** (CHDD), kde Pr(CHDD) $= \beta$. Doplnok ku Pr(CHDD) je pravdepodobnosť $1 - \beta$ a nazýva sa **sila testu** (pri nejakej alternatíve; prípad D). Sila testu závisí na zvolenej testovacej metóde a hlavne na tom, aké je skutočné rozdelenie náhodnej premennej X, aký je typ použitej testovacej štatistiky alebo aké sú skutočné hodnoty parametrov θ.

Teda

rozhodnutie/skutočnosť	H_0 platí	H_0 neplatí
H_0 nezamietnuť	správne rozhodnutie	chyba II. druhu
H_0 zamietnuť	chyba I. druhu	správne rozhodnutie

Vyššie uvedené pravdepodobnosti môžeme zosumarizovať nasledovne:

A) $1 - \alpha \leq \Pr(\text{nezamietni } H_0 | H_0 \text{ platí}) = \Pr(\text{nezamietni } H_0 | H_1 \text{ neplatí})$;

B) $\alpha \geq \Pr(\text{CHPD}) = \Pr(\text{zamietni } H_0 | H_0 \text{ platí}) = \Pr(\text{zamietni } H_0 | H_1 \text{ neplatí})$;

C) $\beta = \Pr(\text{CHDD}) = \Pr(\text{nezamietni } H_0 | H_0 \text{ neplatí}) = \Pr(\text{nezamietni } H_0 | H_1 \text{ platí})$;

D) $1 - \beta = \Pr(\text{zamietni } H_0 | H_0 \text{ neplatí}) = \Pr(\text{zamietni } H_0 | H_1 \text{ platí})$.

V prípade (A) a (B) hovoríme o teste na hladine významnosti α (α-level test). Ak v (A) a (B) nahradíme znamienko nerovnosti rovnosťou, hovoríme o teste úrovne α (size α test). Test $\mathcal{T}(\mathbf{x})$ charakterizujeme jeho **silofunkciou** $\beta^*(\theta) = 1-\beta(\theta) = \Pr(\theta \in \Theta : \mathcal{T}(\mathbf{x}) = \Theta_1)$. Táto je závislá na θ a niekoľkých iných argumentoch (vždy sú špecifikované pri konkrétnych testoch) ako aj na type alternatívnej hypotézy. **Hladinu významnosti** α je možné definovať tiež pomocou $\beta^*(\theta)$, kde $\alpha = \sup_{\theta\in\Theta_0} \beta^*(\theta)$, t.j. ide o najväčšiu možnú pravdepodobnosť chyby I. druhu.

Hladina významnosti α je daná (určená štatistikom, experimentátorom) vopred, testovacia štatistika $T(\mathbf{x})$ sa vypočíta na základe realizácií (odpovedí) x náhodnej premennej X, kritická hodnota (alebo kvantil) sa nájde v štatistických tabuľkách alebo ju vypočítame pomocou R.

V praxi teda voľba $\mathcal{W}$ závisí na požiadavke, aby pravdepodobnosť chyby I. druhu bola menšia alebo rovná zvolenému kladnému číslu α, kde $\alpha \in (0, 1/2)$, najčastejšie $\alpha = 0.05$, 0.01 alebo 0.001. Súčasne volíme $\mathcal{W}$ tak, aby pravdepodobnosť chyby I. druhu bola čo najmenšia. Optimálne by sme chceli mať $\beta^*(\theta)$ (pri nejakej hodnote θ) tak veľké číslo (blížiace sa jednotke) ako je možné, keď $\theta \in \Theta_1$ a také malé číslo, keď $\theta \in \Theta_0$. Tieto dve požiadavky sú v konflikte, t.j.

- zväčšovanie oboru nezamietania H_0 zmenšuje pravdepodobnosť chyby I. druhu, ale zväčšuje pravdepodobnosť chyby II. druhu;
- zmenšovanie oboru nezamietania H_0 zväčšuje pravdepodobnosť chyby I. druhu, ale zmenšuje pravdepodobnosť chyby II. druhu.

Extrémnymi prípadmi sú dve situácie, kde $\mathcal{T}_0(\mathbf{x}) = \Theta_0$ a $\mathcal{T}_1(\mathbf{x}) = \Theta_1$, t.j. každý z týchto dvoch testov je najlepší možný, keď θ patrí určitej podmnožine Θ, ale najhorší možný v opačnom prípade.

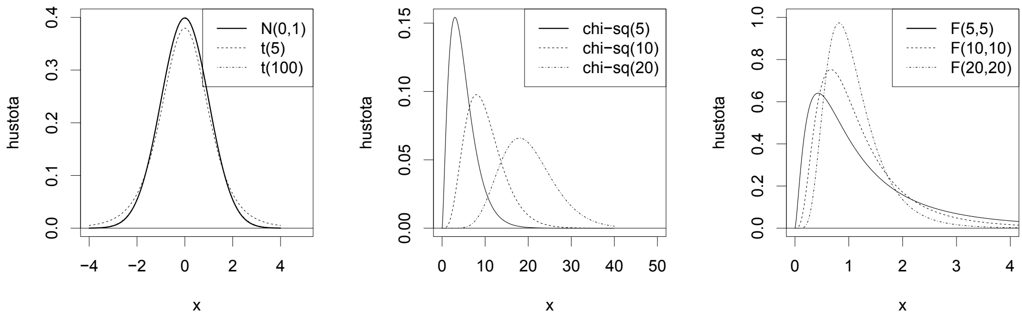
Klasickým prístupom k tejto dileme je **Neyman-Pearsonov prístup**, kde dopredu fixujeme α a vyberieme taký test, ktorý má pri danej alternatíve najväčšiu silu $1-\beta$ pre $\theta \in \Theta_1$ medzi všetkými možnými na teoreticky stanovenej (danej, nominálnej) hladine významnosti α. Testy spomínané v kapitolách 5 až 8 sú najsilnejšími testami pri daných alternatívach za určitých predpokladov (bližšie pozri kap. 5 až 8). V reálnych situáciách (alebo často aj za porušenie predpokladov testov) sa nominálna hladina významnosti líši od aktuálnej (skutočnej) hladiny významnosti (pozri definíciu 42).

Definícia 33 (kvantil) *Nech F_X je distribučná funkcia náhodnej premennej X a $\alpha \in (0,1)$, kde $F_X(x_{1-\alpha}) = 1-\alpha$. Potom číslo $x_{1-\alpha} = x(1-\alpha) = F_X^{-1}(1-\alpha)$ sa nazýva α-**kvantil** príslušného rozdelenia. Pre vybrané rozdelenia platí*

- $\Pr(X < u_{1-\alpha}) = 1-\alpha$, *kde $x_{1-\alpha} = u_{1-\alpha}$ pre $X \sim N(0,1)$ (**normálne rozdelenie** s parametrami $\mu = 0$ a $\sigma^2 = 1$; hovorí sa mu aj **normálne normované rozdelenie** alebo **štandardizované normálne rozdelenie**);*
- $\Pr\left(X < \chi^2_{df}(1-\alpha)\right) = 1-\alpha$, *kde $x_{1-\alpha} = \chi^2_{df}(1-\alpha)$ pre $X \sim \chi^2_{df}$ (**chi-kvadrát rozdelenie** s df stupňami voľnosti);*
- $\Pr(X < t_{df}(1-\alpha)) = 1-\alpha$, *kde $x_{1-\alpha} = t_{df}(1-\alpha)$ pre $X \sim t_{df}$ (**Studentovo t-rozdelenie** s df stupňami voľnosti);*
- $\Pr(X < F_{df_1,df_2}(1-\alpha)) = 1-\alpha$, *kde $x_{1-\alpha} = F_{df_1,df_2}(1-\alpha)$ pre $X \sim F_{df_1,df_2}$ (**Fisherovo F-rozdelenie** s df_1 a df_2 stupňami voľnosti).*

Hustoty vyššie spomenutých rozdelení pozri na obrázku 4.1. Tiež platí

- $\Pr\left(x_{\alpha/2} < X < x_{1-\alpha/2}\right) = F_X\left(x_{1-\alpha/2}\right) - F_X\left(x_{\alpha/2}\right) = 1-\alpha$, kde $\alpha \in (0,1)$;
- $\Pr(X < Q_1) = 1/4$, kde $\mathrm{Q}_1 = F_X^{-1}(1/4)$ sa nazýva **výberový prvý kvartil** (**dolný kvartil**);
- $\Pr(X < Q_2) = 1/2$, kde $\mathrm{Q}_2 = F_X^{-1}(1/2)$ sa nazýva **výberový druhý kvartil** (**medián**);
- $\Pr(X < Q_3) = 3/4$, kde $\mathrm{Q}_3 = F_X^{-1}(3/4)$ sa nazýva **výberový tretí kvartil** (**horný kvartil**);
- $Q_d = F_X^{-1}(k/10)$ sa nazýva **výberový k-ty decil**, $Q_p = F_X^{-1}(k/100)$ sa nazýva **výberový k-ty percentil**.



Obr. 4.1: Grafické znázornenie hustôt normálneho rozdelenia, t-rozdelenia, χ^2-rozdelenia a F-rozdelenia pri rôznych stupňoch voľnosti

Definícia 34 (kritická hodnota) ***Kritická hodnota*** *príslušného rozdelenia je hodnota x^*_α, ktorú náhodná premenná X prekročí s pravdepodobnosťou α. Pre vybrané rozdelenia platí*

- $\Pr(X > u_\alpha) = \alpha$, *kde* $x^*_\alpha = u_\alpha$ *pre* $X \sim N(0,1)$; $u(\alpha)$ *sa často zapisuje ako* u_α;
- $\Pr\left(X > \chi^2_{df}(\alpha)\right) = \alpha$, *kde* $x^*_\alpha = \chi^2_{df}(\alpha)$ *pre* $X \sim \chi^2_{df}$;
- $\Pr(X > t_{df}(\alpha)) = \alpha$, *kde* $x^*_\alpha = t_{df}(\alpha)$ *pre* $X \sim t_{df}$;
- $\Pr(X > F_{df_1,df_2}(\alpha)) = \alpha$, *kde* $x^*_\alpha = F_{df_1,df_2}(\alpha)$ *pre* $X \sim F_{df_1,df_2}$.

Z vyššie uvedeného je zreteľné, že $\alpha \times 100\%$ kritická hodnota je identická s $(1-\alpha) \times 100\%$ kvantilom. Pri kritickej hodnote počítame obsah pod krivkou hustoty príslušného rozdelenia nad týmto bodom a pri kvantile pod týmto bodom.

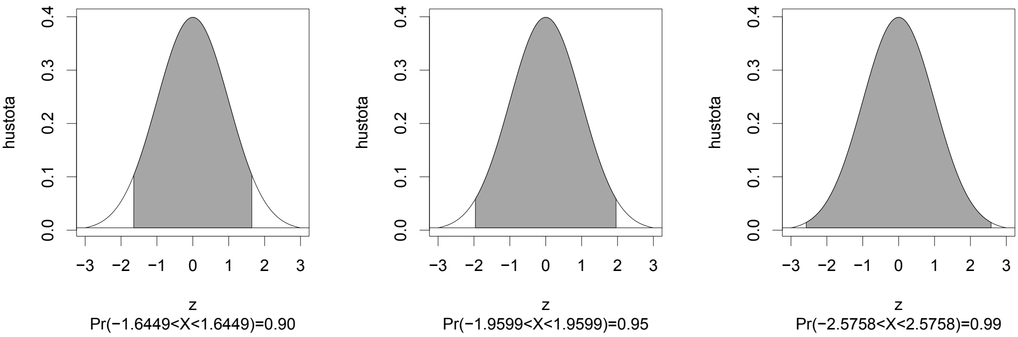
Stupne voľnosti *df* predstavujú množstvo nezávislej informácie, ktoré potrebujeme na charakterizáciu rozdelenia pravdepodobnosti nejakej náhodnej premennej. Používajú sa aj namiesto n ako prevencia pred nadhodnotením odhadu nejakého parametra (napr. rozptylu). Ide najčastejšie o rozsah náhodného výberu n zmenšený o jedna (napr. Studentovo t-rozdelenie, F-rozdelenie), kde jednotka odpočítaná od n znamená jeden voľne viazaný parameter (napr. jedna stredná hodnota, jeden rozptyl a pod.). Pri zložitejších modeloch s viacerými parametrami odpočívame od n počet voľne viazaných parametrov (napr. počet stredných hodnôt) alebo *df* môže tiež predstavovať počet hladín kategoriálnej premennej zmenšený o počet voľne viazaných parametrov.

Príklad 145 (štandardizované normálne rozdelenie) *Vypočítajte kritické hodnoty $u(\alpha)$ rozdelenia $N(0,1)$, kde $\alpha =$ 0.1, 0.05, 0.01, 0.025 a 0.005.*

Riešenie v R (pozri obrázok 4.2 a 4.3)

```
qnorm(1-0.1) # 1.281552
qnorm(1-0.05) # 1.644854
qnorm(1-0.01) # 2.326348
qnorm(1-0.025) # 1.959964
qnorm(1-0.005) # 2.575829
```

Na výpočet pravdepodobnosti pod kvantilom sa používa funkcia `pnorm(Q)`. Na výpočet pravdepodobnosti nad kritickou hodnotou sa používa funkcia `1-pnorm(Q)`. Keďže štandardizované normálne rozdelenie je symetrické okolo nuly, $u(\alpha) = u(1-\alpha)$.



Obr. 4.2: Grafické znázornenie významu pravdepodobnosti pod krivkou rozdelenia medzi dvoma kvantilmi (normálne rozdelenie)

Príklad 146 (Studentove t-rozdelenie) *Vypočítajte kritické hodnoty Studentovho t-rozdelenia so stupňami voľnosti $df = 10$, t.j. $t_{df}(\alpha)$, kde $\alpha = 0.1,\ 0.05,\ 0.01,\ 0.025$ a 0.005.*

Riešenie v R (pozri obrázok 4.4)

```
qt(1-0.1,10) # 1.372184
qt(1-0.05,10) # 1.812461
qt(1-0.01,10) # 2.763769
qt(1-0.025,10) # 2.228139
qt(1-0.005,10) # 3.169273
```

Na výpočet pravdepodobnosti pod kvantilom sa používa funkcia pt(Q,df). Na výpočet pravdepodobnosti nad kritickou hodnotou sa používa funkcia 1-pt(Q,df). Keďže Studentovo t-rozdelenie je symetrické okolo nuly, $t_{df}(\alpha) = t_{df}(1-\alpha)$. Nejaký kvantil t-rozdelenia je približne rovný kvantilu štandardizovaného normálneho rozdelenia až pre veľmi veľké stupne voľnosti (resp. pravdepodobnosti nad kritickými hodnotami sú približne rovnaké). Napr. 1-pnorm(1.644854)≈1-pt(1.644869,100000)= 0.05. Avšak napr. 1-pt(1.644869,100)$\doteq 0.052$.

Príklad 147 (χ^2-rozdelenie) *Vypočítajte kritické hodnoty χ^2-rozdelenia so stupňami voľnosti $df = 10$, t.j. $\chi^2_{df}(\alpha)$, kde $\alpha = 0.1,\ 0.05,\ 0.01,\ 0.025$ a 0.005.*

Riešenie v R (pozri obrázok 4.5)

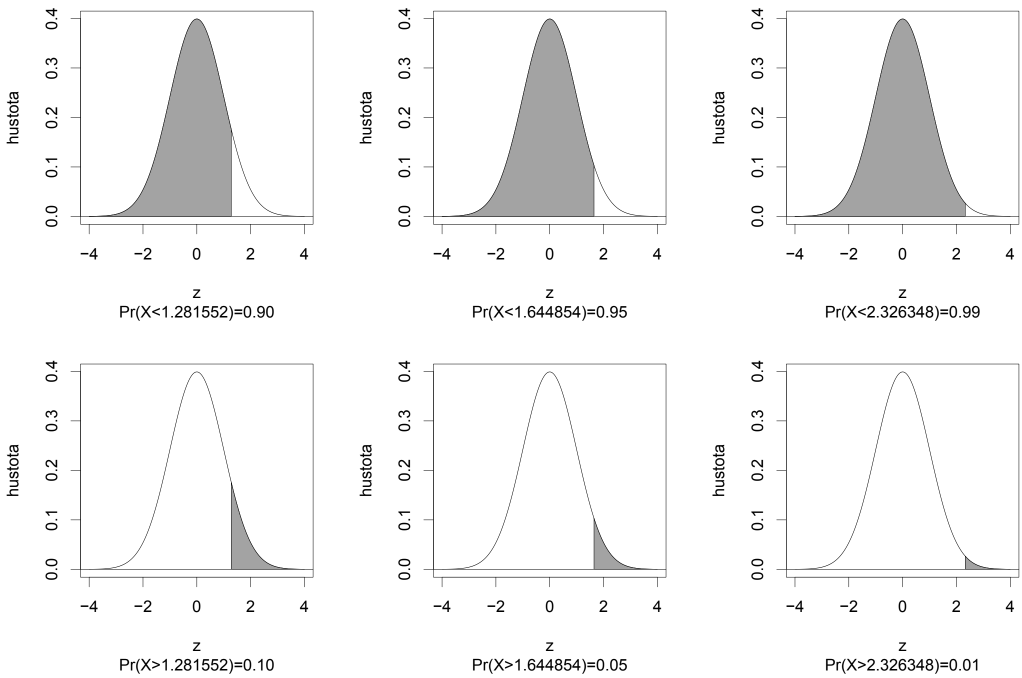
```
qchisq(1-0.1,10) # 15.98718
qchisq(1-0.05,10) # 18.30704
qchisq(1-0.01,10) # 23.20925
qchisq(1-0.025,10) # 20.48318
qchisq(1-0.005,10) # 25.18818
```

Na výpočet pravdepodobnosti pod kvantilom alebo nad kritickou hodnotou sa používa pchisq(Q,df). Keďže χ^2-rozdelenie nie je symetrické, $\chi^2_{df}(\alpha) \neq \chi^2_{df}(1-\alpha)$.

Príklad 148 (F-rozdelenie) *Vypočítajte kritické hodnoty F-rozdelenia so stupňami voľnosti $df_1 = 20$ a $df_2 = 20$, t.j. $F_{df_1,df_2}(\alpha)$, kde $\alpha = 0.1,\ 0.05,\ 0.01,\ 0.025$ a 0.005.*

Riešenie v R (pozri obrázok 4.6)

```
qf(1-0.1,20,20) # 1.793843
qf(1-0.05,20,20) # 2.124155
qf(1-0.01,20,20) # 2.937735
qf(1-0.025,20,20) # 2.464484
qf(1-0.005,20,20) # 3.317786
```



Obr. 4.3: Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou normálneho rozdelenia; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom

Na výpočet pravdepodobnosti pod kvantilom sa používa funkcia `pf(Q,df1,df2)`. Na výpočet pravdepodobnosti nad kritickou hodnotou sa používa funkcia `1-pf(Q,df1,df2)`. Keďže F-rozdelenie nie je symetrické, $F_{df_1,df_2}(\alpha) \neq F_{df_1,df_2}(1-\alpha)$.

4.1 *Asymptotické vlastnosti odhadov

Štatistická inferencia o θ_* sa vykonáva na základe odhadu $\widehat{\theta} = \widehat{\theta}_n$, kde n v dolnom indexe zdôrazňuje pri akom rozsahu náhodného výberu je odhad θ počítaný (index n sa často vynecháva). Podobne štatistická inferencia o $g(\theta_*)$ sa vykonáva na základe odhadu $g(\widehat{\theta}_n)$. **Bodový odhad** parametrickej funkcie $g(\theta)$ je štatistika $T_n = T(X_1, X_2, \ldots, X_n)$ nadobúdajúca hodnoty blízko $g(\theta)$. Nech $T_n = g(\widehat{\theta}_n)$ je nejaký odhad $g(\theta)$, podobne $T_n = \widehat{\theta}_n$ je nejaký odhad θ. Potom môžeme definovať nasledovné typy konvergencií pre T_n (čo platí aj pre $g(\theta) = \theta$):

1. *Konvergencia skoro všade*

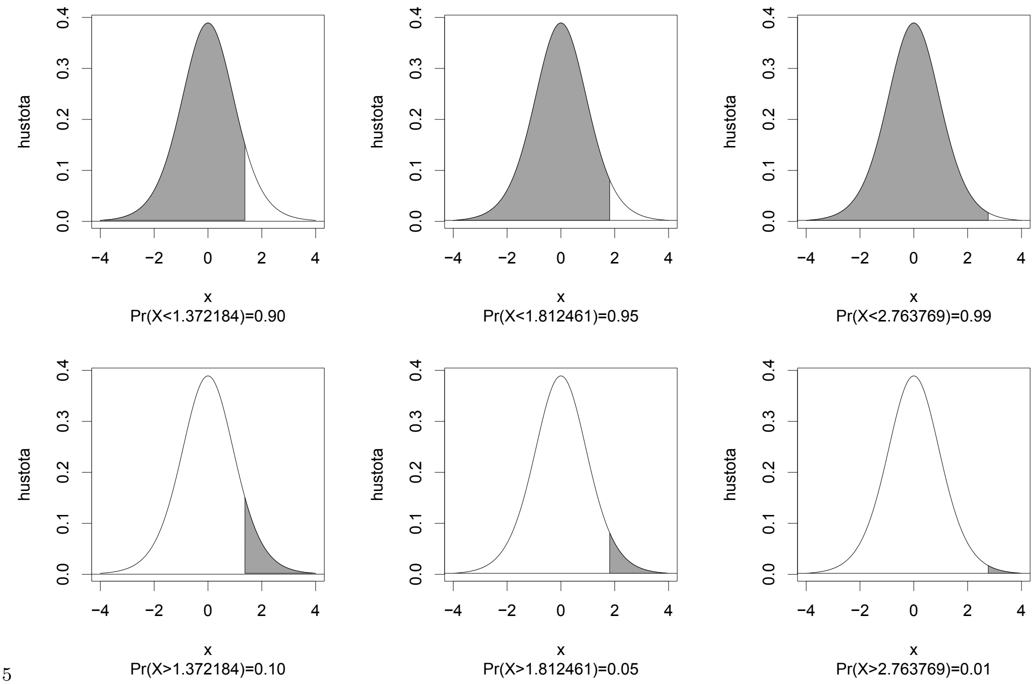
$$\Pr\left(\lim_{n\to\infty} T_n = g(\theta)\right) = 1, \text{ pre } \forall\theta\in\Theta,$$

2. *Konvergencia v kvadratickom strede*

$$T_n \overset{n\to\infty}{\to} g(\theta) \Leftrightarrow E_\theta\left[(T_n - g(\theta))^2\right] \overset{n\to\infty}{\to} 0, \text{ pre } \forall\theta\in\Theta,$$

3. *Konvergencia podľa pravdepodobnosti* (ozn. „$\overset{\mathcal{P}}{\to}$“)

$$\lim_{n\to\infty} \left(\Pr\left(|T_n - g(\theta)| \geq \varepsilon\right)\right) = 0, \text{ pre } \varepsilon > 0, \text{ pre } \forall\theta\in\Theta,$$



Obr. 4.4: Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou t-rozdelenia s $df = 10$; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom

4. *Konvergencia v distribúcii* (ozn. „$\stackrel{\mathcal{D}}{\sim}$")

$$\lim_{n\to\infty} F_n(x) = F_X(x)$$

vo všetkých bodoch x, kde $F_n(x)$ je empirická distribučná funkcia a $F_X(x)$ je spojitá distribučná funkcia.

Pozn.: Z konvergencie skoro všade vyplýva konvergencia v kvadratickom strede, z konvergencie v kvadratickom strede vyplýva konvergencia podľa pravdepodobnosti.

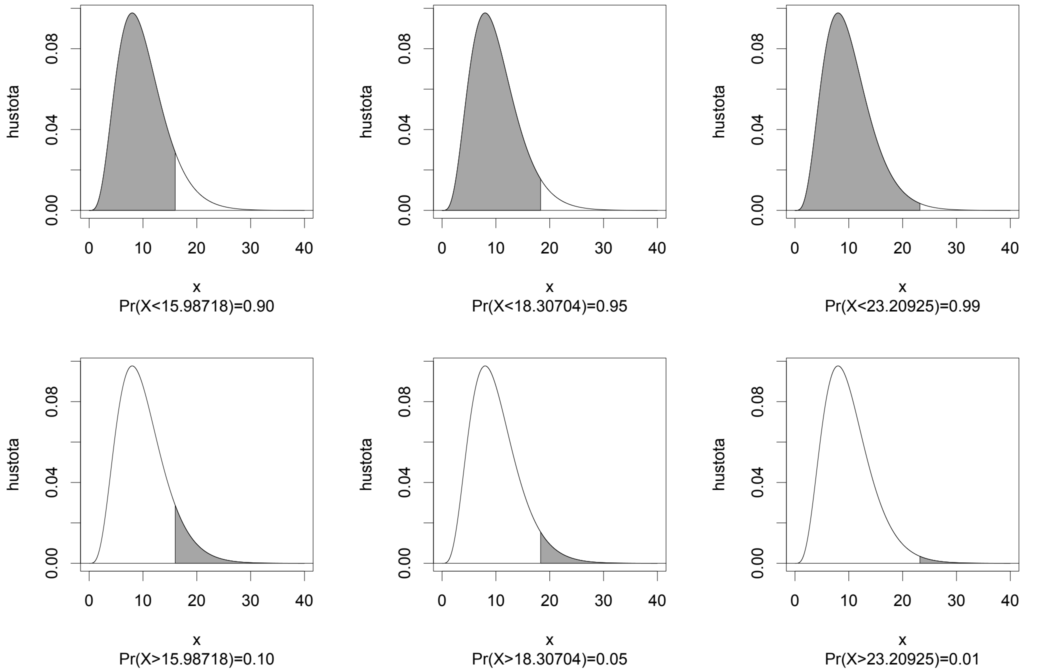
Nech $X_1, X_2, \ldots, X_n$ je náhodný výber z rozdelenia F_*, $g(\theta)$ je parametrická funkcia a T_n, T_1 a T_2 sú štatistiky, $E[T_n]$ je stredná hodnota a $Var[T_n]$ rozptyl štatistiky T_n. Potom

1. hovoríme, že štatistika T_n je **nevychýlený odhad** parametrickej funkcie $g(\theta)$, ak $E[T_n] = g(\theta), \forall\theta \in \Theta$, t.j. odhad T_n nesmie paramatrickú funkciu *systematicky nadhodnocovať* alebo *podhodnocovať* – ak táto podmienka nie je splnená, hovoríme o **vychýlenom odhade**;

2. ak T_1, T_2 sú dva nevychýlené odhady tej istej parametrickej funkcie $g(\theta)$, potom T_1 je **lepší odhad** $g(\theta)$ ako T_2, ak $Var[T_1] < Var[T_2], \forall\theta \in \Theta$, t.j. ak existuje odhad s menším rozptylom, je potrebné ho v štatistickej inferencii použiť namiesto odhadu s väčším rozptylom (to môžeme docieliť optimalizáciou dizajnu experimentu vhodnou voľbou bodov, v ktorých budeme merať);

3. T_n sa nazýva **asymptoticky nevychýlený odhad** $g(\theta)$, ak platí $\lim_{n\to\infty} E[T_n] = g(\theta)$, t.j. s rastúcim rozsahom náhodného výberu n, klesá výchylka odhadu (to môžeme docieliť optimalizáciou dizajnu experimentu vhodnou voľbou n);

4. T_n sa nazýva **konzistentný odhad** $g(\theta)$, ak platí $\lim_{n\to\infty} \Pr(|T_n - g(\theta)| \geq \epsilon) = 0$, t.j. s rastúcim rozsahom náhodného výberu n klesá pravdepodobnosť, že odhad sa bude reali-



Obr. 4.5: Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou χ^2-rozdelenia s $df = 10$; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom

zovať ďaleko od $g(\theta)$ (to môžeme docieliť optimalizáciou dizajnu experimentu vhodnou voľbou n);

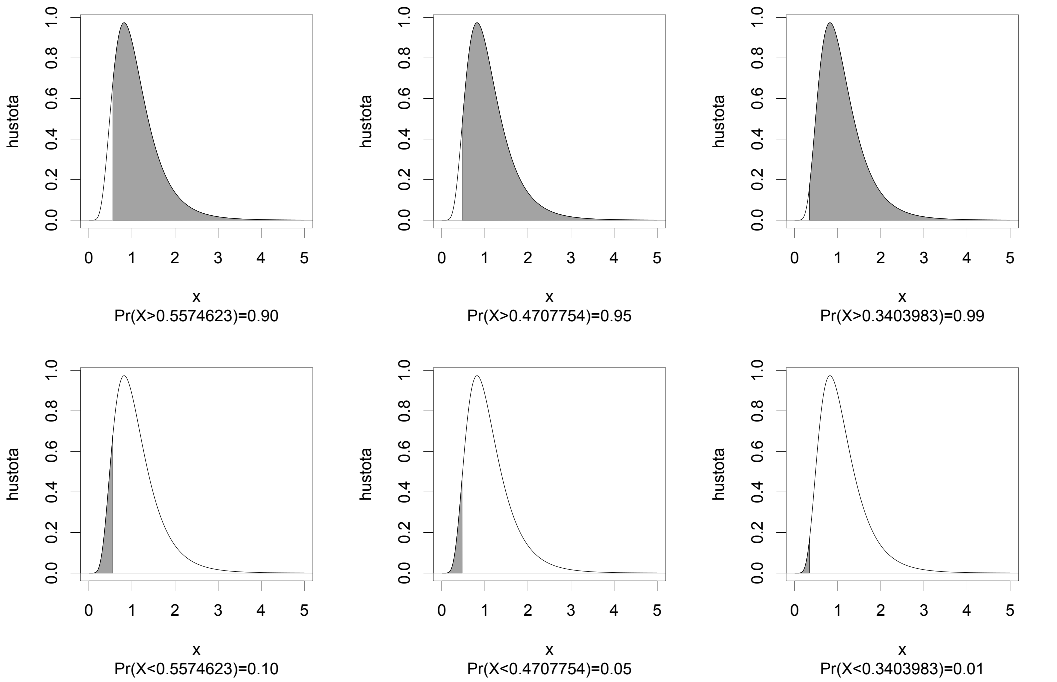
5. T_n sa nazýva **konzistentný** a **asymptoticky efektívny (eficientný, výdatný) odhad** $g(\theta)$, ak platí

   $(T_n - g(\theta)) \overset{\mathcal{D}}{\sim} N(0, C(g(\theta)))$, kde $\mathcal{D}$ znamená „v distribúcii“, $C(g(\theta))$ je Raova-Cramerova spodná hranica definovaná ako $C(g(\theta)) = -\frac{[g'(\theta)]^2}{E_\theta[\frac{\partial^2}{\partial\theta^2} \ln L(\theta|\mathbf{x})]} = \frac{[g'(\theta)]^2}{ni(\theta)}$, kde $i(\theta)$ je Fisherova miera informácie jedného pozorovania, $L(\theta|\mathbf{x})$ je funkcia vierohodnosti (táto vlastnosť sa používa pri testovaní hypotéz ako predpoklad asymptotického rozdelenia testovacej štatistiky za predpokladu normality rozdelenia X);

6. T_n sa nazýva **asymptoticky normálny odhad** $g(\theta)$, ak platí $\frac{T_n - g(\theta)}{\sqrt{C(g(\theta))}} \overset{\mathcal{D}}{\sim} N(0,1)$ (táto vlastnosť sa tiež používa pri testovaní hypotéz ako predpoklad asymptotického rozdelenia testovacej štatistiky za predpokladu normality rozdelenia X).

Z nevychýlenosti odhadu vyplýva jeho asymptotická nevychýlenosť a z asymptotickej nevychýlenosti vyplýva konzistencia, ak $\lim_{n\to\infty} Var[T_n] = 0$. Implikáciou bodu (6) je $T_n \overset{\mathcal{D}}{\sim} N(g(\theta), \frac{[g'(\theta)]^2}{ni(g(\theta))})$, kde $ni(g(\theta)) = \mathcal{I}(g(\theta))$. Ak $g(\theta) = \theta$, potom $T_n \overset{\mathcal{D}}{\sim} N(\theta, \frac{1}{ni(\theta)})$, kde $ni(\theta) = \mathcal{I}(\theta)$. Ak $k > 1$, potom je $\boldsymbol{\theta}$ vektor a môžeme písať

$$\sqrt{n}\left(g(\widehat{\boldsymbol{\theta}}_n) - g(\boldsymbol{\theta})\right) \overset{\mathcal{D}}{\sim} N_k\left(\mathbf{0}, \Delta^T(i(\boldsymbol{\theta}))^{-1}\Delta\right),$$



Obr. 4.6: Grafické znázornenie významu pravdepodobnosti (obsahu) pod krivkou F-rozdelenia s $df_1 = 20$ a $df_2 = 20$; obsah pod (prvý riadok) a nad (druhý riadok) príslušným kvantilom

kde $\Delta = \frac{\partial}{\partial \boldsymbol{\theta}} g(\boldsymbol{\theta})$. Ak $g(\boldsymbol{\theta}) = \boldsymbol{\theta}$, potom

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \overset{\mathcal{D}}{\sim} N_k\left(\mathbf{0}, (i(\boldsymbol{\theta}))^{-1}\right).$$

Veta 6 (vlastnosti odhadu) *Nech náhodná premenná X pochádza z normálneho rozdelenia s parametrami μ a σ^2, $X \sim N(\mu, \sigma^2)$, kde $E[X] = \mu$ je stredná hodnota a $Var[X] = \sigma^2$ je rozptyl náhodnej premennej X. Potom výberový aritmetický priemer $\overline{X}_n = \overline{X}$ je nevychýleným odhadom μ a výberový rozptyl $S^2_{n-1} = S^2$ je nevychýleným odhadom σ^2 ($n-1$ v dolnom indexe S^2_{n-1} znamená počet stupňov voľnosti). Zároveň platí, že $\overline{X}$ a S^2 sú asymptoticky nevychýlené, konzistentné, asymptoticky efektívne a asymptoticky normálne. Preto môžeme písať nasledovné*

$$\sqrt{n}(\overline{X}_n - \mu) \overset{\mathcal{D}}{\sim} N(0, \sigma^2), \text{ čo je ekvivalentné } \overline{X}_n \overset{\mathcal{D}}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right).$$

a

$$\sqrt{n}(S^2_n - \sigma^2) \overset{\mathcal{D}}{\sim} N(0, 2\sigma^4), \text{ kde } S^2_n = \frac{n-1}{n} S^2_{n-1}.$$

Tiež môžeme písať

$$\sqrt{n}(S^2_{n-1} - \sigma^2) \overset{\mathcal{D}}{\sim} \sqrt{n}\sigma^2\left(\frac{\chi^2_{n-1}}{n-1} - 1\right) \text{ a } \sqrt{n}(S^2_n - \sigma^2) \overset{\mathcal{D}}{\sim} \sqrt{n}\sigma^2\left(\frac{\chi^2_{n-1}}{n} - 1\right),$$

čo je ekvivalentné $(n-1)\frac{S^2_{n-1}}{\sigma^2} \overset{\mathcal{D}}{\sim} \chi^2_{n-1}$ a $n\frac{S^2_n}{\sigma^2} \overset{\mathcal{D}}{\sim} \chi^2_{n-1}$. Potom môžeme písať nasledovné

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \overset{\mathcal{D}}{\sim} N_2(\mathbf{0}, \boldsymbol{\Sigma}),$$

kde $\widehat{\boldsymbol{\theta}}_n = (\overline{X}_n, S_n^2)^T$, $\boldsymbol{\theta} = (\mu, \sigma^2)^T$, $\mathbf{0} = (0,0)^T$ *a* $\boldsymbol{\Sigma} = diag(\sigma^2, 2\sigma^4)$. *Na tomto mieste ale treba zdôrazniť, že n použité pri odhadovaní musí byť dostatočne veľké, zvyčajne sa uvádza* $n > 30$. *Realizáciou* $\overline{X}$ *je* $\overline{x}$ *a realizáciou* S^2 *je* s^2.

S_{n-1}^2 je nevychýlený odhad σ^2, S_n^2 (maximálne vierohodný odhad σ^2) je vychýlený odhad σ^2 a bude v priemere σ^2 podceňovať. Avšak priemerná kvadratická chyba (MSE) je pre S_n^2 menšia ako pre S_{n-1}^2. MSE sa ale nemôže používať ako kritérium pre parameter škály (čo je σ^2), pretože MSE penalizuje odhad rovnako pre podcenenie ako aj precenenie odhadu (pre parameter polohy, ako je μ, je MSE vhodným kritériom).

Definícia 35 *Nech dvojrozmerná náhodná premenná* $(X_1, X_2)^T$ *pochádza z dvojrozmerného normálneho rozdelenia s parametrami* $\boldsymbol{\mu}$ *a* $\boldsymbol{\Sigma}$, $(X_1, X_2)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *kde* $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$, $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$, $E[X_1] = \mu_1$ *je stredná hodnota a* $Var[X_1] = \sigma_1^2$ *je rozptyl náhodnej premennej X;* $E[X_2] = \mu_2$ *je stredná hodnota a* $Var[X_2] = \sigma_2^2$ *je rozptyl náhodnej premennej Y. Potom aritmetické priemery* $\overline{X}_j$ $(j = 1, 2)$ *sú nevychýlené odhady* μ_j *a výberové rozptyly* S_j^2 *sú nevychýlenými odhadmi* σ_j^2. *Zároveň platí, že* $\overline{X}_j$ *a* S_j^2 *sú asymptoticky nevychýlené, konzistentné, asymptoticky efektívne a asymptoticky normálne. Ďalej platí, že výberová kovariancia* S_{12} *je nevychýleným odhadom* σ_{12}, *ako aj, že je asymptoticky nevychýlený, konzistentný, asymptoticky efektívny a asymptoticky normálny odhad. Na tomto mieste ale treba zdôrazniť, že n použité pri odhadovaní musí byť dostatočne veľké, zvyčajne sa uvádza* $n > 30$. *Avšak* $E[R_{12}]$ *sa rovná* ρ *len približne (zhoda nastane až pre* $n > 30$*). Naviac píšeme, že* R_{12} *je nevychýleným odhadom* ρ. *Zároveň platí, že* R_{12} *je asymptoticky nevychýlený a konzistentný. Realizáciou* S_{12} *je* s_{12} *a realizáciou* R_{12} *je r.*

Veta 7 (koeficient variácie) *Nech náhodná premenná X pochádza z normálneho rozdelenia s parametrami* μ *a* σ^2, $X \sim N(\mu, \sigma^2)$, *kde* $E[X] = \mu$ *je stredná hodnota a* $Var[X] = \sigma^2$ *je rozptyl náhodnej premennej X. Nech* $g(\boldsymbol{\theta}) = \sigma/\mu$, *kde* $\boldsymbol{\theta} = (\mu, \sigma^2)^T$, $g(\widehat{\boldsymbol{\theta}}_n) = \frac{S_n}{\overline{X}_n}$ *a* $\Delta = \frac{\partial}{\partial \boldsymbol{\theta}} g(\boldsymbol{\theta}) = \left(-\frac{\sigma}{\mu^2}, \frac{1}{2\sigma\mu}\right)^T$. *Potom*

$$\sqrt{n}\left(\frac{S_n}{\overline{X}_n} - \frac{\sigma}{\mu}\right) \overset{\mathcal{D}}{\sim} N_k\left(0, \Delta^T (i(\boldsymbol{\theta}))^{-1} \Delta\right),$$

kde $\Delta^T (i(\boldsymbol{\theta}))^{-1} \Delta = \left(-\frac{\sigma}{\mu^2}, \frac{1}{2\sigma\mu}\right) \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix} \left(-\frac{\sigma}{\mu^2}, \frac{1}{2\sigma\mu}\right)^T = \frac{\sigma^2}{\mu^2}\left(\frac{\sigma^2}{\mu^2} + \frac{1}{2}\right)$.

4.2 Interval spoľahlivosti Waldovho typu

Intervalový odhad parametrickej funkcie $g(\theta)$ je interval (D, H), ktorého hranice sú štatistiky $D(X_1, X_2, \ldots, X_n)$ a $H(X_1, X_2, \ldots, X_n)$. Tento intervalový odhad nazývame $100 \times (1 - \alpha)\,\%$ **interval spoľahlivosti** (IS), kde číslo $1 - \alpha$ je **koeficient spoľahlivosti**. O tomto intervale môžeme *dosť spoľahlivo* prehlásiť, že obsahuje skutočnú (neznámu) hodnotu parametra $g(\theta)$, čo tvrdíme *s primeraným stupňom dôvery* (s dostatočne veľkou pravdepodobnosťou).

Rozlišujeme nasledovné IS: **dvojstranné** (DIS), kde pre DIS sú dolná hranica D a horná hranica H vytvorené tak, aby platilo

$$\Pr\left(D \le g(\theta) \le H\right) = 1 - \alpha, \text{ pre } \forall\theta \in \Theta$$

a **jednostranné** (JIS), kde pre JIS volíme konkrétnu hranicu tak, aby mohlo dôjsť iba k podceneniu alebo preceneniu neznámeho parametra, tj. aby platilo

$$\Pr\left(D_* \le g(\theta)\right) = 1 - \alpha, \text{ pre } \forall\theta \in \Theta,$$

$$\Pr\left(g(\theta) \le H^*\right) = 1 - \alpha, \text{ pre } \forall\theta \in \Theta.$$

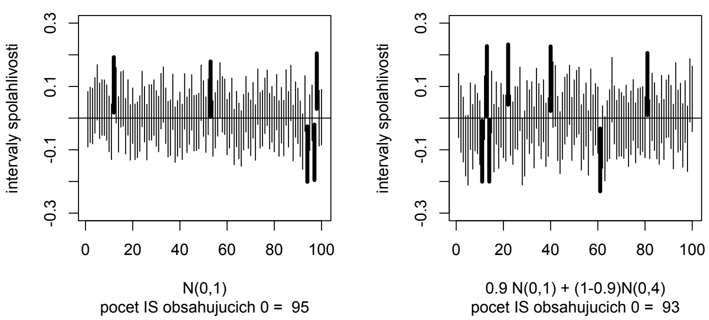
Empirický IS dostaneme použitím kritických hodnôt rozdelenia $g(\widehat{\theta})$, parametra $g(\widehat{\theta})$ a jeho štandardnej chyby $\widehat{SE[g(\widehat{\theta})]}$. Postup pri jeho konštrukcii je nasledovný (postup je pre DIS, podobný je aj pre JIS; pozri napr. Budíková a kol. (2010)):

1. majme štatistiku T_n, ktorá je nevychýleným bodovým odhadom parametrickej funkcie $g(\theta)$;
2. vypočítame **pivotovú štatistiku** (**pivot**) T_{piv}, ktorá je transformáciou štatistiky T_n, je monotónnou funkciou $g(\theta)$, poznáme jej rozdelenie, ktoré je nezávislé na $g(\theta)$; pomocou známeho rozdelenia pivota T_{piv}, nájdeme kritické hodnoty[1] $t_{\alpha/2}$ a $t_{1-\alpha/2}$, potom $\Pr(t_{1-\alpha/2} < T_{\text{piv}} < t_{\alpha/2}) = 1 - \alpha$;
3. nerovnosť $t_{1-\alpha/2} < T_{\text{piv}} < t_{\alpha/2}$ transformujeme ekvivalentnými úpravami na $D < g(\theta) < H$;
4. štatistiky D a H nahradíme ich realizáciami d a h a získame tak $100 \times (1-\alpha)\%$ **empirický IS**, o ktorom prehlásime, že pokrýva $g(\theta)$ s pravdepodobnosťou aspoň $1-\alpha$, t.j. ak mnohonásobne nezávisle získame realizácie $x_1, x_2, \ldots, x_n$ náhodného výberu $X_1, X_2, \ldots, X_n$ z nejakého rozdelenia F_* a pomocou tejto realizácie zostrojíme $100 \times (1-\alpha)\%$ empirické IS pre $g(\theta)$, potom podiel počtu týchto intervalov, ktoré pokrývajú $g(\theta)$ a počtu všetkých zostrojených intervalov (tých, čo pokrývajú a aj tých, čo nepokrývajú $g(\theta)$) bude približne $1-\alpha$.

Príklad 149 *Nech náhodný výber $X_1, X_2, \ldots, X_n$ pochádza z normálneho rozdelenia s parametrami μ a σ^2, t.j. $X \sim N(\mu, \sigma^2)$, potom štatistika $T_n = \overline{X}$, $T_{piv} = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ (kde σ je známe), kritické hodnoty rozdelenia $N(0,1)$ pivota sú rovné $t_{\alpha/2} = u_{\alpha/2}$, $t_{1-\alpha/2} = u_{1-\alpha/2} = -u_{\alpha/2}$, Potom $100 \times (1-\alpha)\%$ empirický IS je (D, H), kde $D = \overline{X} - u_{\alpha/2}\sigma/\sqrt{n}$ a $H = \overline{X} + u_{\alpha/2}\sigma/\sqrt{n}$. Ak dosadíme za $\overline{X}$ odhad $\overline{x}$, dostaneme $100 \times (1-\alpha)\%$ empirický IS, kde $d = \overline{x} - u_{\alpha/2}\sigma/\sqrt{n}$ a $h = \overline{x} + u_{\alpha/2}\sigma/\sqrt{n}$.*

Príklad 150 (MC experiment pre IS) *Nech (a) $X \sim N(0,1)$ a (b) $X \sim [pN(0,1)+(1-p)N(0,4)]$, kde $p = 0.9$, t.j. ide o zmes dvoch normálnych rozdelení $X \sim N(0,1)$ a $X \sim N(0,4)$ v pomere 9:1. Vygenerujte $M = 100$ náhodných výberov s rozsahom $n = 500$ a vypočítajte Waldove $100(1-\alpha)\%$ empirický IS pre μ. Zistite, koľko IS obsahuje strednú hodnotu $\mu = 0$. Toto číslo podelené M predstavuje simulovanú hladinu významnosti α.*

Riešenie v R (pozri obrázok 4.7)



Obr. 4.7: MC experiment pre IS (IS, ktoré neobsahujú $\mu = 0$, sú označené hrubou čiarou); vľavo $X \sim N(0,1)$ a vpravo $X \sim [0.9N(0,1) + (1-0.9)N(0,4)]$

```
M <- 100
n <- 500
var.mix <- priemery.mix <- priemery <- array(0,M)
DH.mix <- DH <- array(0,M)
HH.mix <- HH <- array(0,M)
```

[1] $t_{\alpha/2}$ nie sú kritické hodnoty Studentovho t-rozdelenia, ale kritické hodnoty rozdelenia T_{piv}.

```
prav <- 0.1
i <- 0
while (i < M) {
i <- i + 1
priemery[i] <- mean(rnorm(n))
simul.mix <- rnorm(n,0,1+rbinom(n,1,prav))
priemery.mix[i] <- mean(simul.mix)
var.mix[i] <- var(simul.mix)
DH[i] <- priemery[i] + qnorm(0.025)*sqrt(1/n)
HH[i] <- priemery[i] + qnorm(0.975)*sqrt(1/n)
DH.mix[i] <- priemery.mix[i] + qnorm(0.025)*sqrt(var.mix[i]/n)
HH.mix[i] <- priemery.mix[i] + qnorm(0.975)*sqrt(var.mix[i]/n)
}
"interval.plot" <- function(y1,y2,xlab){
n <- length(y1)
plot(y1,type="n",ylim=c(-.3,.3),xlab=xlab,ylab="intervaly spolahlivosti")
podmienka <- (y1 <=0 & y2 >=0)
segments((1:n)[y1<0&y2>0],y1[y1<0&y2>0],(1:n)[y1<0&y2>0],y2[y1<0&y2>0])
segments((1:n)[y1>0],y1[y1>0],(1:n)[y1>0],y2[y1>0],lwd=4,col="black")
segments((1:n)[y2<0],y1[y2<0],(1:n)[y2<0],y2[y2<0],lwd=4,col="black")
SUM <- sum(podmienka)
title(sub=paste("pocet IS obsahujucich 0 = ",SUM))
abline(h=0)
}
windows(8,4)
par(mfcol=c(1,2))
interval.plot(DH,HH,xlab="N(0,1)")
interval.plot(DH.mix,HH.mix,xlab="0.9 N(0,1) + (1-0.9)N(0,4)")
```

Všeobecne empirický IS pre θ zostrojíme na základe $\widehat{\theta}$ a $\widehat{SE[\widehat{\theta}]}$, ak je $Var[\widehat{\theta}]$ neznáme alebo $SE[\widehat{\theta}]$, ak je $Var[\widehat{\theta}]$ známe (v praxi je často neznáme). **Šírka empirického IS** pre θ (niekedy sa používa aj pojem **dĺžka empirického IS pre** θ) je definovaná ako $d(d,h)$, kde $d(\cdot,\cdot)$ je Euklidovská vzdialenosť. Potom platí

- ak sa spoľahlivosť $1-\alpha$ zvyšuje (hladina významnosti α klesá), šírka IS rastie ($\widehat{SE[\widehat{\theta}]}$ resp. $SE[\widehat{\theta}]$ a n sú fixované, konštatné);
- ak sa spoľahlivosť $1-\alpha$ znižuje (hladina významnosti α rastie), šírka IS klesá ($\widehat{SE[\widehat{\theta}]}$ resp. $SE[\widehat{\theta}]$ a n sú fixované, konštatné);
- ak rozsah n rastie, šírka IS klesá ($\widehat{SE[\widehat{\theta}]}$ resp. $SE[\widehat{\theta}]$ a α sú fixované);
- ak rozsah n klesá, šírka IS rastie ($\widehat{SE[\widehat{\theta}]}$ resp. $SE[\widehat{\theta}]$ a α sú fixované);
- ak $\widehat{SE[\widehat{\theta}]}$ resp. $SE[\widehat{\theta}]$ rastie, tak šírka IS rastie (α a n sú fixované);
- ak $\widehat{SE[\widehat{\theta}]}$ resp. $SE[\widehat{\theta}]$ klesá, tak šírka IS klesá (α a n sú fixované).

Ak sa šírka empirického IS zvyšuje, hovoríme, že neistota o parametri θ rastie. Ak sa šírka empirického IS znižuje, hovoríme, že neistota o θ klesá.

Podobne môžeme uvažovať aj o parametrickej funkcii $g(\theta)$ a jej empirickom IS.

4.3 Testovanie H_0 oproti H_1

Definícia 36 (Obor nezamietania H_0) *Nech X je náhodná premenná z nejakého rozdelenia, ktoré závisí na parametri $\theta \in \Theta$, $g(\theta)$ je parametrická funkcia. Testujme nulovú hypotézu $H_{01}: g(\theta) = g(\theta_0)$ oproti obojstrannej alternatíve $H_{11}: g(\theta) \neq g(\theta_0)$. Nech (D,H) je intervalový odhad parametrickej funkcie $g(\theta)$ so spoľahlivosťou $1-\alpha$. Potom*

$$\mathcal{A}_{IS,1} = \{D,H; g(\theta_0) \in (D,H)\}$$

*je **obor nezamietania testu** H_{01} **oproti** H_{11} **na hladine významnosti** α. Ak máme testovať $H_{02}: g(\theta) \leq g(\theta_0)$ oproti jednostrannej (pravostrannej) alternatíve $H_{12}: g(\theta) > g(\theta_0)$ a ak D_* je dolný odhad pre $g(\theta)$ so spoľahlivosťou $1-\alpha$, potom*

$$\mathcal{A}_{IS,2} = \{D_*; D_* < g(\theta_0)\}$$

*je **obor nezamietania testu** H_{02} **oproti** H_{12} **na hladine významnosti** α. Ak máme testovať $H_{03} : g(\theta) \geq g(\theta_0)$ oproti jednostrannej (ľavostrannej) alternatíve $H_{13} : g(\theta) < g(\theta_0)$ a ak H^* je horný odhad pre $g(\theta)$ so spoľahlivosťou $1-\alpha$, potom*

$$\mathcal{A}_{IS,3} = \{H^*; H^* > g(\theta_0)\}$$

*je **obor nezamietania testu** H_{03} **oproti** H_{13} **na hladine významnosti** α.*

Definícia 37 (Obor zamietania H_0 (kritický obor)) *Nech X je náhodná premenná z nejakého rozdelenia, ktoré závisí na parametri $\theta \in \Theta$, $g(\theta)$ je parametrická funkcia. Testujme nulovú hypotézu $H_{01} : g(\theta) = g(\theta_0)$ oproti obojstrannej alternatíve $H_{11} : g(\theta) \neq g(\theta_0)$. Nech (D, H) je intervalový odhad parametrickej funkcie $g(\theta)$ so spoľahlivosťou $1-\alpha$. Potom*

$$\mathcal{W}_{IS,1} = \{D, H; g(\theta_0) \notin (D, H)\}$$

*je **kritický obor testu** H_{01} **oproti** H_{11} **na hladine významnosti** α. Ak máme testovať $H_{02} : g(\theta) \leq g(\theta_0)$ oproti jednostrannej (pravostrannej) alternatíve $H_{12} : g(\theta) > g(\theta_0)$ a ak D_* je dolný odhad pre $g(\theta)$ so spoľahlivosťou $1-\alpha$, potom*

$$\mathcal{W}_{IS,2} = \{D_*; D_* \geq g(\theta_0)\}$$

*je **kritický obor testu** H_{02} **oproti** H_{12} **na hladine významnosti** α. Ak máme testovať $H_{03} : g(\theta) \geq g(\theta_0)$ oproti jednostrannej (ľavostrannej) alternatíve $H_{13} : g(\theta) < g(\theta_0)$ a ak H^* je horný odhad pre $g(\theta)$ so spoľahlivosťou $1-\alpha$, potom*

$$\mathcal{W}_{IS,3} = \{H^*; H^* \leq g(\theta_0)\}$$

*je **kritický obor testu** H_{03} **oproti** H_{13} **na hladine významnosti** α.*

Definícia 38 (Testovacie kritérium) ***Testovacím kritériom*** *nazveme testovaciu štatistiku $T = T_0 = T_0(X_1, X_2, \ldots, X_n)$, ktorej asymptotické rozdelenie za platnosti H_0 poznáme. Množina hodnôt, ktoré môže T_0 nadobúdať sa delí na dve podmnožiny, a to **obor nezamietania** H_0 (ozn. $\mathcal{A}$) a **obor zamietania** H_0 (kritický obor, ozn. $\mathcal{W}$). Tieto dva obory sú oddelené **kritickými hodnotami** $t_{\alpha/2}$ a $t_{1-\alpha/2}$, resp. t_α a $t_{1-\alpha}$ (podľa typu H_0 a H_1) rozdelenia testovacej štatistiky T_0 (za platnosti H_0).*

Testovanie H_0 oproti H_1 je rozhodovací postup založený na náhodnom výbere $X_1, X_2, \ldots, X_n$, pomocou ktorého zamietame alebo nezamietame H_0. Všeobecne zaužívaný postup je nasledovný:

1. stanovíme H_0 a H_1,
2. zvolíme hladinu významnosti α,
3. zvolíme vhodné testovacie kritérium T_0,
4. zamietneme/nezamietneme H_0 jedným z troch ekvivalentných spôsobov
   - pomocou kritického oboru $\mathcal{W} = \mathcal{W}_T$ (a realizácie $t_0 = t_{\text{obs}}$),
   - pomocou kritického oboru $\mathcal{W}_{\text{IS}}$, t.j. empirického intervalu spoľahlivosti (a $g(\theta_0)$),
   - pomocou p-hodnoty.

Definícia 39 (Testovanie pomocou kritického oboru $\mathcal{W}$) ***Zamietanie*** *H_0. Ak realizácia t_0 testovacej štatistiky T_0 spadá do kritického oboru $\mathcal{W}$ (ekvivalentne nespadá do oboru nezamietania $\mathcal{A}$), zamietame H_0 na hladine významnosti α a znamená to, že máme dostatok dôkazov na zamietnutie H_0. **Nezamietanie** H_0. Ak realizácia t_0 testovacej štatistiky T_0 spadá do oboru nezamietania $\mathcal{A}$ (ekvivalentne nespadá do kritického oboru $\mathcal{W}$), nezamietame H_0 na hladine významnosti α a znamená to, že nemáme dostatok dôkazov na zamietnutie H_0.*

Stanovenie kritického oboru. Ak $t_{\min}$ je najmenšia hodnota testovacieho kritéria T_0 a $t_{\max}$ je najväčšia hodnota testovacieho kritéria T_0, ktoré môže teoreticky nadobúdať, potom

1. **obojstranná alternatíva** – kritický obor $\mathcal{W}_1 = (t_{\min}, t_{1-\alpha/2}) \cup (t_{\alpha/2}, t_{\max})$,
2. **jednostranná (pravostranná) alternatíva** – kritický obor $\mathcal{W}_2 = (t_\alpha, t_{\max})$,
3. **jednostranná (ľavostranná) alternatíva** – kritický obor $\mathcal{W}_3 = (t_{\min}, t_{1-\alpha})$.

Definícia 40 (Testovanie pomocou intervalu spoľahlivosti) ***Zamietanie*** *H_0: Ak IS neobsahuje $g(\theta) = g(\theta_0)$ (za platnosti H_0), zamietame H_0 na hladine významnosti α a znamená to, že máme dostatok dôkazov na zamietnutie H_0.* ***Nezamietanie*** *H_0: Ak IS obsahuje $g(\theta) = g(\theta_0)$ (za platnosti H_0), nezamietame H_0 na hladine významnosti α a znamená to, že nemáme dostatok dôkazov na zamietnutie H_0.*

Použitie $100(1-\alpha)\%$ IS pre $g(\theta)$ je nasledovné

1. **obojstranný IS**, ak ide o obojstrannú hypotézu,
2. **jednostranný (ľavostranný, dolný) IS**, ak ide o jednostrannú (pravostrannú) alternatívu,
3. **jednostranný (pravostranný, horný) IS**, ak ide o jednostrannú (ľavostrannú) alternatívu.

Definícia 41 (Testovanie pomocou pozorovanej hladiny významnosti alebo p-hodnoty) *Minimálnu hladinu významnosti α (pre nejakú vypočítanú testovaciu štatistiku T_0), na základe ktorej bude $H_{02} : g(\theta) \leq g(\theta_0)$ (testovaná oproti $H_{12} : g(\theta) > g(\theta_0)$) zamietnutá, nazývame* ***pozorovaná hladina významnosti*** *alebo* ***p-hodnota****, t.j.*

$$\text{p-hodnota} = \alpha_{obs} = \sup_{\theta \in \Theta_0} \Pr\left(T(X_1, X_2, \ldots, X_n) \geq T(x_1, x_2, \ldots, x_n); \theta\right),$$

čo môžeme písať aj nasledovne

$$\text{p-hodnota} = Pr(\text{nejaká testovacia štatistika sa rovná pozorovanej alebo je väčšia}|H_0 \text{ platí}).$$

Podobne môžeme zadefinovať p-hodnotu aj pre zvyšné dve alternatívy. Pomocou p-hodnoty sa rozhodujeme, či rozdiel hodnoty pozorovanej testovacej štatistiky od jej hodnoty za platnosti H_0 je spôsobený náhodou alebo výberovým rozptylom. Čím je α_{obs} bližšia k nule, tým menšia je pravdepodobnosť, že nejaká testovacia štatistika $T(X_1, X_2, \ldots, X_n)$ má p-hodnotu (za platnosti H_0) rovnú alebo menšiu než pozorovaná. Pravdepodobnosť tejto udalosti je väčšia za platnosti H_1. V tomto zmysle chápeme p-hodnotu ako indikátor hodnovernosti H_0. Ak $\alpha_{\text{obs}} < \alpha = 0.05$, máme dostatok dôkazov na zamietnutie H_0 a hovoríme, že výsledok testu je **štatisticky signifikantný**. Ak $\alpha_{\text{obs}} > \alpha = 0.1$, nemáme dostatok dôkazov na zamietnutie H_0 a hovoríme, že výsledok testu je **štatisticky nesignifikantný**. Hodnoty 0.05 a 0.1 chápeme ako referenčné hodnoty v širšom slova zmysle. Ak sa α_{obs} blíži k hraniciam intervalu $\langle 0.05, 0.1\rangle$, chápeme to ako približovanie sa k jednej z vyššie spomínaných interpretácií. Situácia, keď $\alpha_{\text{obs}} \in \langle 0.05, 0.1)$ sú ťažko interpretovateľné a niekedy sa označujú ako **hraničná štatistická signifikancia**. V aplikáciách vyjadrujeme zamietanie a nezamietanie H_0 nasledovne:

rozsah pre p-hodnotu	hviezdičky signifikancie	zvyčajný slovný popis výsledku testu
$\langle 0, 0.001)$	***	extrémne vysoko štatisticky signifikantný
$\langle 0.001, 0.01)$	**	vysoko štatisticky signifikantný
$\langle 0.01, 0.05)$	*	štatisticky signifikantný
$\langle 0.05, 0.1)$	·	hraničné štatistický signifikantný
$\langle 0.1, 1\rangle$		nesignifikantný

Príklady frekventistických interpretácií p-hodnôt

- p-hodnota < 0.001: odhadnutý efekt sa vyskytne v menej ako jednom prípade z tisíc (šanca výskytu odhadnutého efektu je menej ako $1 : 999$), ak efekt v populácii neexistuje (výskyt takéhoto efektu je vysoko nepravdepodobný, ak efekt v populácii v skutočnosti neexistuje a naopak – výskyt takéhoto efektu je vysoko pravdepodobný, ak efekt v populácii v skutočnosti existuje);

- p-hodnota < 0.01: odhadnutý efekt sa vyskytne v menej ako jednom prípade zo sto (šanca výskytu odhadnutého efektu je menej ako 1 : 99), ak efekt v populácii neexistuje (výskyt takéhoto efektu je veľmi nepravdepodobný, ak efekt v populácii v skutočnosti neexistuje a naopak – výskyt takéhoto efektu je veľmi pravdepodobný, ak efekt v populácii v skutočnosti existuje);
- p-hodnota < 0.05: odhadnutý efekt sa vyskytne v menej ako piatich prípadoch zo sto (šanca výskytu odhadnutého efektu je menej ako 5 : 95 alebo 1 : 19), ak efekt v populácii neexistuje (výskyt takéhoto efektu je dostatočne nepravdepodobný, ak efekt v populácii v skutočnosti neexistuje a naopak – výskyt takéhoto efektu je dostatočne pravdepodobný, ak efekt v populácii v skutočnosti existuje);
- p-hodnota ≥ 0.05: odhadnutý efekt sa vyskytne v piatich alebo viacerých prípadoch zo sto (v 5 alebo viac ako 5 % prípadov);
- p-hodnota $= k, k \in \langle 0.05, 1\rangle$: odhadnutý efekt sa vyskytne v $100 \times k$ prípadoch zo sto (v $100 \times k$ % prípadov).

Spôsob výpočtu p-hodnoty

1. **obojstranná alternatíva** – p-hodnota $= 2\min(\Pr(T_0 \leq t_0|H_0), \Pr(T_0 \geq t_0|H_0))$, čo platí napr. pre normálne a Studentovo rozdelenie testovacej štatistiky (symetrické rozdelenia) a pre χ^2_{df} rozdelenie a F_{df_1,df_2} rozdelenie testovacej štatistiky (nesymetrické rozdelenia), pozri kap. 6 a 7; alebo p-hodnota $= \min(\Pr(T_0 \leq t_0|H_0), \Pr(T_0 \geq t_0|H_0))$, čo platí napr. pre χ^2_{df} rozdelenie a F_{df_1,df_2} rozdelenie testovacej štatistiky (nesymetrické rozdelenie), pozri kap. 5 a 8;
2. **jednostranná (pravostranná) alternatíva** – p-hodnota $= \Pr(T_0 \geq t_0|H_0)$, pozri kap. 6 a 7;
3. **jednostranná (ľavostranná) alternatíva** – p-hodnota $= \Pr(T_0 \leq t_0|H_0)$, pozri kap. 6 a 7.

Konzervatívny a liberálny test a IS

Definícia 42 (Konzervatívny a liberálny test) *Test, ktorého **aktuálna/skutočná hladina významnosti** je menšia ako **nominálna hladina významnosti** α, sa nazýva **konzervatívny** (test by mal teoreticky „rýchlejšie“ zamietať H_0, ale skutočnosť je opačná, t.j. test zamieta „pomalšie“). Test, ktorého aktuálna hladina významnosti je väčšia ako nominálna hladina významnosti α, sa nazýva **liberálny** (test by mal teoreticky „pomalšie“ zamietať H_0, ale skutočnosť je opačná, t.j. test zamieta „rýchlejšie“).*

Definícia 43 (Konzervatívny a liberálny IS) *IS, ktorého **aktuálna/skutočná pravdepodobnosť pokrytia** je väčšia ako **nominálna pravdepodobnosť pokrytia** $1-\alpha$, sa nazýva **konzervatívny** (IS má väčšiu pravdepodobnosť, že obsahuje θ_0, ako by sme teoreticky predpokladali). IS, ktorého aktuálna pravdepodobnosť pokrytia je menšia ako nominálna pravdepodobnosť pokrytia $1-\alpha$, sa nazýva **liberálny** (IS má menšiu pravdepodobnosť, že obsahuje θ_0, ako by sme teoreticky predpokladali).*

Praktické rady

1. Nikdy nehovoríme „H_0 prijímame“, ale „nemáme dostatočné dôkazy na zamietnutie H_0“. Nikdy nehovoríme „H_1 prijímame/neprijímame“, keďže testujeme H_0 a nie H_1.
2. Testovacia štatistika je vytvorená z charakteristík vierohodnostnej funkcie, ktorej tvar je ovplyvnený rozdelením F_*, z ktorého pochádza náhodná premenná X. Čím je rozdelenie F náhodnej premennej X odlišnejšie od skutočného (predpokladaného) rozdelenia F_*, tým viac sa líšia závery z testovania hypotéz od záverov, ktoré sú platné, ak X pochádza z F_*.
3. Voľba nulovej a alternatívnej hypotézy nie je ľubovoľná, ale je ovplyvnená konkrétnou praktickou situáciou.

Príklad 151 (neistota) *Ak X nepochádza z predpokladaného rozdelenia $N(\mu, \sigma^2)$, rozdelenie testovacej štatistiky pre test o μ nie je normálne $N(0,1)$, a teda závery (zamietame/nezamietame H_0) vytvorené použitím asymptotickej normality testovacej štatistiky nebudú platné, resp. neistota záveru sa zvyšuje so zväčšujúcim sa rozdielom rozdelenia X od $N(\mu, \sigma^2)$.*

Príklad 152 (istota vs. neistota) *Ak X pochádza z rozdelenia $N(\mu, \sigma^2)$, môžeme testovať nulové hypotézy o μ pomocou asympotického testu, ktorého testovacia štatistika má normálne rozdelenie. Ak X nepochádza z rozdelenia $N(\mu, \sigma^2)$, ale napr. zo zmesi $(1-p)N(\mu, \sigma_1^2) + pN(\mu, \sigma_2^2)$, kde $\sigma_1 \ll \sigma_2$ (t.j. ide o prítomnosť odľahlých pozorovaní s pravdepodobnosťou p, kde p je nejaké malé číslo), nemôžeme testovať nulové hypotézy o μ pomocou asympotického testu, ktorého testovacia štatistika má normálne rozdelenie.*

Príklad 153 *Majme model normálneho rozdelenia s neznámou strednou hodnotou μ a známym rozptylom $\sigma^2 = 2^2$. Jednovýberovým Z-testom testujte $H_{01}: \mu = 1$ oproti $H_{11}: \mu \neq 1$ na hladine významnosti $\alpha = 0.05$. Strednú hodnotu μ sme odhadli na základe náhodného výberu s rozsahom 30, kde $\overline{x} = 1.3$.*

Riešenie (aj v R)

1. **Testovacia štatistika (testovacie kritérium)** – $T_0 = Z_W = \frac{\overline{X}-\mu_0}{\sigma}\sqrt{n}$ a jej realizácia $t_0 = z_W = \frac{\overline{x}-\mu_0}{\sigma}\sqrt{n} = \frac{1.3-1}{2}\sqrt{30} \doteq 0.822$.

```
aritm.priem <- 1.3
sigma <- 2
n <- 30
alfa <- 0.05
test.stat <- (1.3-1)/2*sqrt(30) # 0.8215838
```

2. **Zamietacia oblasť** – kritická hodnota $u_{1-\alpha/2} \doteq -1.96$ a $u_\alpha \doteq 1.96$, potom kritický obor: $\mathcal{W}_1 = (t_{\min}, t_{1-\alpha/2}) \cup (t_{\alpha/2}, t_{\max}) = (-\infty, -1.96) \cup (1.96, \infty)$.

```
krit.hodn <- qnorm(1-alfa/2) # 1.959964
```

3. $100 \times (1-\alpha)\%$ **empirický DIS** – keďže $\alpha = 0.05$, ide o 95% empirický DIS pre μ_0, kde $(d,h) = (\overline{x} - t_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x} + t_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}) \doteq (1.3 - 1.96\frac{2}{\sqrt{30}}, 1.3 + 1.96\frac{2}{\sqrt{30}}) \doteq (0.584, 2.016)$.

```
IS <- aritm.priem + c(-1,1)*krit.hodn*sigma/sqrt(30) # 0.5843223 2.0156777
```

4. **Štatistický záver** – p-hodnota=$2\Pr(T_0 \geq |t_0|) = 2\Pr(T_0 \geq 0.822) = 0.411$.

```
p.hodn <- 2*(1-pnorm(abs(test.stat))) # 0.4113138
```

   H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí do kritického oboru, (2) $\mu_0 = 1$ patrí DIS a nakoniec (3) p-hodnota je väčšia alebo rovná ako 0.05.

Použitie triády rozdelení pravdepodobnosti (fundamentálne predpoklady testovania hypotéz; príklad testovacej štatistiky Z_W):

1. náhodná premenná X má **normálne rozdelenie** $X \sim N(\mu, \sigma^2)$;
2. náhodná premenná $\overline{X}$ má **normálne rozdelenie** $\overline{X} \sim N(\mu, \sigma^2/n)$;
3. testovacia štatistika Z_W má **štandardizované normálne rozdelenie**, t.j. $Z_W \stackrel{\mathcal{D}}{\sim} N(0,1)$.

Pri iných testoch môže mať náhodná premenná X iné rozdelenie; iné testovacie štatistiky môžu mať mať tiež iné rozdelenie. Rozdelenie X bude špecifikované v kaptiolách 5 až 8 pri každom teste, rovnako aj rozdelenie testovacej štatistiky.

4.4 *Tri typy testovacích štatistík

Zmysluplným kritériom na zistenie odpovede na otázku, či θ dobre popisuje dáta $\mathbf{x}$ sa javí tzv. **relatívna (štandardizovaná) vierohodnosť** $\mathcal{L}(\theta|\mathbf{x}) = L(\theta|\mathbf{x})/L(\widehat{\theta}|\mathbf{x})$. Ignorujúc prípad, keď je $L(\theta|\mathbf{x})$ neohraničená, $\mathcal{L}(\theta|\mathbf{x}) \in \langle 0, 1\rangle$. Nech θ_0 je θ špecifikovaná v hypotéze H_0. Potom $\mathcal{L}(\theta_0|\mathbf{x})$ predstavuje kvantifikáciu zhody medzi dátami a touto hypotézou. Ak $\mathcal{L}(\theta_0|\mathbf{x}) = 1$, potom je zhoda maximálna (za predpokladu, že štatistický model je vhodný). Ak $\mathcal{L}(\theta_0|\mathbf{x}) = 0$, potom sa hypotéza H_0 ukazuje ako úplne nevyhovujúca. V realite však $\mathcal{L}(\theta_0|\mathbf{x})$ leží niekde medzi nulou a jednotkou. Jednou z možností je stanoviť **kritickú hodnotu** r. Potom ak $\mathcal{L}(\theta_0|\mathbf{x}) \le r$, zhoda medzi H_0 a dátami je príliž malá, a teda θ_0 nie je adekvátna a H_0 zamietame. Ak $\mathcal{L}(\theta_0|\mathbf{x}) > r$, potom zhoda medzi H_0 a dátami je dostatočná, a teda θ_0 je adekvátna a H_0 nezamietame. Je dôležité si uvedomiť, že kľúčovou je voľba kritickej hodnoty a možná adaptácia vyššie spomenutého princípu na zložitejšie situácie.

Testujme $H_0 : \theta = \theta_0$ oproti $H_1 : \theta \neq \theta_0$. Potom zovšeobecnením relatívnej vierohodnosti za predpokladu spojitosti $L(\theta|\mathbf{x})$ je **jednoduchý pomer vierohodnosti** rovný

$$\lambda(\mathbf{x}) = \lambda = \frac{L(\theta_0|\mathbf{x})}{\sup_{\theta\in\Theta} L(\theta|\mathbf{x})} = \frac{L(\theta_0|\mathbf{x})}{L(\widehat{\theta}|\mathbf{x})},$$

kde $\lambda(\mathbf{x}) = \mathcal{L}(\theta_0|\mathbf{x})$ je testovacou štatistikou a $L(\theta|\mathbf{x})$ je spojitá pre všetky x.

Testujme $H_0 : \theta \in \Theta_0$ oproti $H_1 : \theta \in \Theta_1$. Potom je **zovšeobecnený pomer vierohodnosti** rovný

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta\in\Theta_0} L(\theta|\mathbf{x})}{\sup_{\theta\in\Theta} L(\theta|\mathbf{x})}.$$

Monotónna transformácia $\lambda(\mathbf{x})$ spolu s korešpondujúcou transformáciou kritických hodnôt nezmení delenie Θ na Θ_0 a Θ_1, a teda nezmení samotný štatistický test. Preto ekvivalentnou testovacou štatistikou bude

$$U_{\mathrm{LR}} = -2\ln\lambda(\mathbf{X}),$$

ktorá sa nazýva **testovacia štatistika pomerom vierohodnosti**. Jej realizácia $u_{\mathrm{LR}} = -2\ln\lambda(\mathbf{x})$. Interpretácia oproti $\lambda(\mathbf{x})$ bude obrátená z dôvodu vynásobenia $2\ln\lambda(\mathbf{x})$ číslom -1. Je zrejmé, že $u_{\mathrm{LR}} \in (0,\infty)$.

Použitím Taylorovho rozvoja $l(\theta_0)$ okolo $\widehat{\theta}$ dostaneme

$$U_{\mathrm{LR}} = -2(l(\theta_0|\mathbf{X}) - l(\widehat{\theta}|\mathbf{X})) \approx -2\left((\theta_0-\widehat{\theta})S(\widehat{\theta}) - \frac{1}{2}(\theta_0-\widehat{\theta})^2\mathcal{I}(\widehat{\theta})\right),$$

kde $S(\widehat{\theta}) = 0$. Dá sa ľahko ukázať, že $\frac{1}{n}\mathcal{I}(\widehat{\theta}) \xrightarrow{\mathcal{P}} i(\theta_0)$; pozri Anděl (2011). Keďže $\widehat{\theta}$ je konzistentný odhad θ_0 za platnosti H_0, potom je konzistentným odhadom θ_0 aj $\tilde{\theta}$, z čoho potom vyplýva nasledovné

$$U_{\mathrm{LR}} \approx n(\theta_0-\widehat{\theta})^2\frac{\mathcal{I}(\theta_0)}{n} \approx n(\theta_0-\widehat{\theta})^2 i(\theta_0) \overset{H_0}{\approx} n(\theta_0-\widehat{\theta})^2 i(\widehat{\theta}) = U_{\mathrm{W}},$$

kde U_{W} sa nazýva **Waldova testovacia štatistika** (aproximácia U_{LR}) a jej realizáciu ozn. u_W. Keďže

$$\sqrt{n}(\widehat{\theta}-\theta_0) \overset{H_0}{\approx} S(\theta_0)/(\sqrt{n}(i(\theta_0))),$$

potom

$$U_{\mathrm{LR}} \approx n(\theta_0-\widehat{\theta})^2 i(\theta_0) \overset{H_0}{\approx} \frac{(S(\theta_0))^2}{n\ i(\theta_0)} = U_{\mathrm{S}},$$

kde U_{S} sa nazýva **skóre testovacia štatistika** (aproximácia U_{LR}) a jej realizáciu ozn. u_S.

Každá z troch vyššie spomenutých testovacích štatistík zachytáva iný aspekt logaritmu funkcie vierohodnosti:

1. U_{LR} meria vhodne normalizovaný rozdiel funkčných hodnôt logaritmu funkcie vierohodnosti v bodoch $\widehat{\theta}$ a θ_0 (t.j. v smere osi y);

2. U_W meria vhodne normalizovaný rozdiel $\widehat{\theta}$ a θ_0 v absolútnej hodnote (v smere osi x);

3. U_S meria vhodne normalizovaný sklon logaritmu funkcie vierohodnosti v bode θ_0.

Príklad 154 *Predpokladajme, že* $X \sim N(\mu, \sigma^2)$, *kde* σ^2 *je známa. Testujme* $H_0 : \theta = \theta_0$ *oproti* $H_1 : \theta \neq \theta_0$, *kde* $\theta = \mu$. *Potom*

1. $U_{LR} = -2(l(\theta_0|\mathbf{X}) - l(\widehat{\theta}|\mathbf{X})) = -\sum_{i=1}^n (X_i - \overline{X})^2/\sigma^2 + \sum_{i=1}^n (X_i - \mu_0)^2/\sigma^2 = n\frac{(\overline{X}-\mu_0)^2}{\sigma^2}$;
2. $U_W = (\overline{X} - \mu_0)^2 \mathcal{I}(\overline{x}) = n\frac{(\overline{X}-\mu_0)^2}{\sigma^2}$;
3. $U_S = \frac{(S(\mu_0))^2}{\mathcal{I}(\mu_0)} = \frac{(n(\overline{X}-\mu_0)/\sigma^2)^2}{n/\sigma^2} = n\frac{(\overline{X}-\mu_0)^2}{\sigma^2}$.

Všetky tri testovacie štatistiky sú rovnaké, t.j. $U_{LR} = U_W = U_S$.

Veta 8 (tri typy testovacích štatistík) *Za podmienok regularity platí:*

1. $U_{LR} = -2(l(\theta_0|\mathbf{X}) - l(\widehat{\theta}|\mathbf{X})) \overset{\mathcal{D}}{\sim} \chi_1^2$;
2. $U_W = (\widehat{\theta} - \theta_0)^2 \mathcal{I}(\widehat{\theta}) \overset{\mathcal{D}}{\sim} \chi_1^2$, *tiež môžeme písať* $U_W^{1/2} = Z_W \overset{\mathcal{D}}{\sim} N(0,1)$;
3. $U_S = \frac{(S(\theta_0))^2}{\mathcal{I}(\theta_0)} \overset{\mathcal{D}}{\sim} \chi_1^2$, *tiež môžeme písať* $U_S^{1/2} = Z_S \overset{\mathcal{D}}{\sim} N(0,1)$.

Veta 9 (tri typy testovacích štatistík) *Ak* $k > 1$, *potom je* $\boldsymbol{\theta}$ *vektor a za podmienok regularity platí:*

1. $U_{LR} = -2(l(\boldsymbol{\theta}_0|\mathbf{X}) - l(\widehat{\boldsymbol{\theta}}|\mathbf{X})) \overset{\mathcal{D}}{\sim} \chi_k^2$;
2. $U_W = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \mathcal{I}(\widehat{\boldsymbol{\theta}})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \overset{\mathcal{D}}{\sim} \chi_k^2$;
3. $U_S = (S(\boldsymbol{\theta}_0))^T (\mathcal{I}(\boldsymbol{\theta}_0))^{-1} S(\boldsymbol{\theta}_0) \overset{\mathcal{D}}{\sim} \chi_k^2$.

Definícia 44 (tri typy testovacích štatistík a IS) *Jednotlivým typom testovacích štatistík prislúchajú nasledovné IS:*

1. *Testovacej štatistike pomerom vierohodnosti* U_{LR} *zodpovedá vierohodnostný IS* $\mathcal{CS}_{1-a}$ *pre* θ, *kde* $\mathcal{CS}_{1-a} = \left\{\theta : U_{LR}(\theta) < \chi_1^2(\alpha)\right\}$, *kde* $U_{LR}(\theta) = -2\ln\frac{L(\theta|\mathbf{x})}{L(\widehat{\theta}|\mathbf{x})}$.
2. *Waldovej testovacej štatistike* U_W *zodpovedá Waldov IS pre* θ, *kde pivot* $T_{piv} = U_W(\theta)$.
3. *Skóre testovacej štatistike* U_S *zodpovedá skóre IS pre* θ *kde pivot* $T_{piv} = U_S(\theta)$.

Definícia 45 (elipsy a elipsoidy spoľahlivosti) $\mathcal{CS}_{1-a}$ *vo všeobecnosti označuje konfidenčnú množinu, ktorá môže zahŕňať nielen IS pre* θ, *ale aj* $100 \times (1-\alpha)\%$ ***elipsy spoľahlivosti*** *pre* $\boldsymbol{\theta}$, *kde* $k = 2$ *a* $100 \times (1-\alpha)\%$ ***elipsoidy spoľahlivosti*** *pre* $\boldsymbol{\theta}$, *kde* $k \geq 3$. *Konfidenčná množina je potom taká množina* θ_0, *ktoré patria do* $\mathcal{CS}_{1-a}$ *(podobne pre* $g(\theta_0)$, $\boldsymbol{\theta}_0$, $g(\boldsymbol{\theta}_0)$*).*

Keď má X normálne rozdelenie, funkcia vierohodnosti má potom parabolický tvar (polynóm druhého stupňa). Pre malé rozsahy výberov n sa môže rozdelenie X veľmi odlišovať od normality a funkcia vierohodnosti je ďaleko od symetrickosti, t.j. symetrického parabolického tvaru. To sa môže stať aj pri výberoch s veľkým rozsahom, kde máme veľa parametrov. V takýchto prípadoch inferencia na základe asymptotickej normality môže dávať neadekvátne výsledky.

Je praktickejšie vypočítať IS pre parametre θ ako testovať hypotézy (Agresti, 2002). Waldov IS bol v minulosti najčastejšie používaný v praxi, pretože je jednoduché ho vypočítať. Častejšie by sa mal však používať vierohodnostný IS, príp. skóre IS, ktorý je preferovaný napr. pre kategoriálne dáta s malým rozsahom.

Použitím U_{LR} testujeme $H_0 : \theta = \theta_0$ oproti $H_1 : \theta \neq \theta_0$, avšak U_{W} a U_{S} umožňuju testovať aj H_0 oproti jednostrannej alternatíve. Ak $n \to \infty$, sú všetky tri štatistiky asymptoticky ekvivalentné (Cox a Hinkley, 1974). Pre výbery s malým rozsahom je test pomerom vierohodnosti viac vhodný ako Waldov test.

Príklad 155 *Predpokladajme, že* $X \sim N(\mu, \sigma^2)$*, kde* σ^2 *je známa. Nech* $\theta = \mu$*. Testujme tri typy hypotéz*

1. $H_{01} : \mu = \mu_0$ *oproti* $H_{11} : \mu \neq \mu_0$*;*
2. $H_{02} : \mu \leq \mu_0$ *oproti* $H_{12} : \mu > \mu_0$*;*
3. $H_{03} : \mu \geq \mu_0$ *oproti* $H_{13} : \mu < \mu_0$*;*

(a) Vypočítajte pravdepodobnosti chýb druhého druhu $\Pr_\mu(CHDD)$ *pri danej alternatíve pre všetky tri typy hypotéz, t.j.* β_{11}*,* β_{12} *a* β_{13}*.*
(b) Vypočítajte silofunkcie pre všetky tri typy hypotéz, t.j. $\beta^*_{11}(\mu)$*,* $\beta^*_{12}(\mu)$ *a* $\beta^*_{13}(\mu)$*.*
(c) Vypočítajte minimálny rozsah náhodného výberu n pri danej α*,* β*,* $\mu - \mu_0$ *a* σ^2 *pre všetky tri typy hypotéz tak, aby* $\mu - \mu_0$ *bolo štatisticky signifikantne odlišné od* 0*.*
(d) Zapíšte kritické obory $\mathcal{W}_1$*,* $\mathcal{W}_2$ *a* $\mathcal{W}_3$*.*
(e) Ako vyzerajú $100(1-\alpha)\%$ *empirické IS pre všetky tri typy hypotéz?*

Riešenie
Predpokladajme, že $X \sim N(\mu, \sigma^2)$, kde μ je neznáma stredná hodnota a $\sigma^2 \in (0, \infty)$ je známa konštanta. $\frac{X-\mu}{\sigma} \overset{\mathcal{D}}{\sim} N(0,1)$ a naviac $\overline{X} \overset{\mathcal{D}}{\sim} N\left(\mu, \sigma^2/n\right)$. Za platnosti nulovej hypotézy $Z_W = \frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} \overset{\mathcal{D}}{\sim} N(0,1)$. Z_W sa nazýva **Waldova testovacia štatistika** a test **jednovýberový Z-test o strednej hodnote** μ. Ak je skutočná hodnota parametra rovná μ, má náhodná veličina $\frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ rozdelenie $N(0,1)$. Distribučnú funkciu normálneho normovaného rozdelenia $N(0,1)$ budeme označovať Φ.

Majme H_{02} oproti H_{12}. Kritický obor je $\mathcal{W}_2$, čo predstavuje najsilnejší test pre danú situáciu. Vyšetríme $\Pr_\mu(\text{CHDD})$ ako funkcie μ, teda $\beta_{12}(\mu) = \Pr_\mu(X \notin \mathcal{W}_2)$, $\mu \in \mathbb{R}^1$. Platí

$$\begin{aligned}
\beta_{12}(\mu) &= \Pr_\mu\left(\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} < u_\alpha\right) = \Pr_\mu\left(\frac{\overline{X}-\mu_0 \pm \mu}{\sigma}\sqrt{n} < u_\alpha\right) \\
&= \Pr_\mu\left(\frac{\overline{X}-\mu}{\sigma}\sqrt{n} < u_\alpha + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) = \Phi\left(u_\alpha + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(u_\alpha - \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right) = \Phi\left(-u_{1-\alpha} - \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right) \\
&= 1 - \Phi\left(u_{1-\alpha} + \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right).
\end{aligned}$$

Silofunkcia bude rovná $\beta^*_{12}(\mu) = 1 - \Phi\left(u_\alpha - \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)$. Z vlastností Φ vyplýva, že funkcia $\beta_{12}(\mu), \mu \in \mathbb{R}^1$ je spojitá, klesajúca v μ s vlastnosťami

$$\beta_{12}(\mu) \begin{cases} \geq 1-\alpha, \text{ ak} & \mu < \mu_0 \\ = 1-\alpha, \text{ ak} & \mu = \mu_0 \\ \leq 1-\alpha, \text{ ak} & \mu > \mu_0 \end{cases},$$

kde $\lim_{\mu\to\infty} \beta_{12}(\mu) = 0$. Ak sa teraz pozrieme na pravdepodobnosť $\beta_{12}(\mu)$ ako funkciu n a $\mu > \mu_0$ fixované, tak je vidieť, že je to funkcia klesajúca v n.

Pokiaľ budeme uvažovať o voľbe rozsahu n, aby test mal hladinu významnosti α a $\Pr_\mu(\text{CHDD})$ bola v pevne zvolenom bode $\mu\,(>\mu_0)$ menšia alebo rovná $\beta \in (0,1)$, teda $\beta_{12}(\mu) \leq \beta$ (β je fixované), dostaneme

$$\Phi\left(u_\alpha + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \leq \beta,$$

čo vedie ku

$$u_\alpha + \frac{\mu_0-\mu}{\sigma}\sqrt{n} \leq x_\beta = -x_{1-\beta} = -u_\beta,$$

kde x_β je kvantil $N(0,1)$ a u_α kritická hodnota $(\Phi(u_\alpha) = 1-\alpha)$, potom

$$n \geq \left(\frac{u_\alpha + u_\beta}{\mu - \mu_0}\right)^2 \sigma^2.$$

Čím väčší rozdiel $\mu - \mu_0$ je pre nás kritický (chceme mať s dostatočnou pravdepodobnosťou zaručené, že ho odhalíme, ak je skutočnosťou), tým menej pozorovaní k tomu potrebujeme.

Majme H_{03} oproti H_{13}. Vyšetríme $\Pr_\mu(\text{CHDD})$ ako funkcie μ, teda $\beta_{13}(\mu) = \Pr_\mu(X \notin \mathcal{W}_3)$, $\mu \in \mathbb{R}^1$. Platí

$$\begin{aligned}
\beta_{13}(\mu) &= \Pr_\mu\left(\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} \geq -u_\alpha\right) = \Pr_\mu\left(\frac{\overline{X}-\mu_0 \pm \mu}{\sigma}\sqrt{n} \geq -u_\alpha\right) \\
&= \Pr_\mu\left(\frac{\overline{X}-\mu}{\sigma}\sqrt{n} \geq -u_\alpha + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) = 1-\Phi\left(-u_\alpha + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= 1-\Phi\left(u_{1-\alpha} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) = \Phi\left(u_\alpha - \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(u_\alpha + \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right).
\end{aligned}$$

Silofunkcia bude rovná $\beta_{13}^*(\mu) = 1-\Phi\left(u_\alpha + \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)$.

Minimálny rozsah n je identický ako v situácii H_{12}.

Majme H_{01} vs. H_{11}. Nech $\beta_{11}(\mu) = \beta_1(\mu) - \beta_2(\mu)$. Platí

$$\begin{aligned}
\beta_1(\mu) &= \Pr_\mu\left(\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} < u_{\alpha/2}\right) = \Pr_\mu\left(\frac{\overline{X}-\mu_0 \pm \mu}{\sigma}\sqrt{n} < u_{\alpha/2}\right) \\
&= \Pr_\mu\left(\frac{\overline{X}-\mu}{\sigma}\sqrt{n} < u_{\alpha/2} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) = \Phi\left(u_{\alpha/2} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(u_{\alpha/2} - \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right),
\end{aligned}$$

$$\begin{aligned}
\beta_2(\mu) &= \Pr_\mu\left(\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} < -u_{\alpha/2}\right) = \Pr_\mu\left(\frac{\overline{X}-\mu_0 \pm \mu}{\sigma}\sqrt{n} < -u_{\alpha/2}\right) \\
&= \Pr_\mu\left(\frac{\overline{X}-\mu}{\sigma}\sqrt{n} < -u_{\alpha/2} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) = \Phi\left(-u_{\alpha/2} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(u_{1-\alpha/2} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) = 1-\Phi\left(u_{\alpha/2} - \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= 1-\Phi\left(u_{\alpha/2} + \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)
\end{aligned}$$

a nakoniec

$$\begin{aligned}
\beta_{11}(\mu) &= \beta_1(\mu) - \beta_2(\mu) \\
&= \Phi\left(u_{\alpha/2} - \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right) + \Phi\left(u_{\alpha/2} + \frac{\mu-\mu_0}{\sigma}\sqrt{n}\right) - 1.
\end{aligned}$$

Silofunkcia bude rovná

$$\begin{aligned}
\beta_{11}^{*}(\mu) &= 1-\beta_{11}(\mu) = 2-\Phi\left(u_{\alpha/2}-\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)-\Phi\left(u_{\alpha/2}+\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)\\
&= 2-\left(1-\Phi\left(u_{1-\alpha/2}+\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)+1-\Phi\left(u_{1-\alpha/2}-\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)\right)\\
&= \Phi\left(u_{1-\alpha/2}+\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)+\Phi\left(u_{1-\alpha/2}-\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right).
\end{aligned}$$

Z uvedeného vyplýva, že pre každé $\delta=\mu-\mu_0$ je silofunkcia

$$\beta_{11}^{*}(\mu_0+\delta)=\beta_{11}^{*}(\mu_0-\delta)$$

a teda silofunkcia je symetrická okolo bodu $\mu=\mu_0$. Pre každé $\alpha\in(0,1)$ a $c=|\mu-\mu_0|/\sigma$ platí

$$\Pr\left(u_{1-\alpha/2}<u<u_{1-\alpha/2}+c\right)>\Pr\left(u_{1-\alpha/2}-c<u<u_{1-\alpha/2}\right)$$

alebo

$$\Phi\left(u_{1-\alpha/2}+c\right)-\Phi\left(u_{1-\alpha/2}\right)>\Phi\left(u_{1-\alpha/2}\right)-\Phi\left(u_{1-\alpha/2}-c\right)$$

a pre každé $\mu\neq\mu_0$

$$\Phi\left(u_{1-\alpha/2}+c\right)+\Phi\left(u_{1-\alpha/2}-c\right)>2\Phi\left(u_{1-\alpha/2}\right)=\alpha$$

Ak chceme vypočítať minimálne n pri daných α, β, σ a μ_0 ako funkciu μ, dostaneme

$$\begin{aligned}
1-\beta &= \beta^{*}\leq\Phi\left(u_{1-\alpha/2}+\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)+\Phi\left(u_{1-\alpha/2}-\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)\\
&= \Phi\left(u_{1-\alpha/2}-\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)+\Phi\left(u_{1-\alpha/2}+\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)
\end{aligned}$$

Ak $\mu>\mu_0$, potom pri rastúcom μ bude rásť aj $\Phi\left(u_{1-\alpha/2}-\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)$ a $\Phi\left(u_{1-\alpha/2}+\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)$ bude klesať (vždy bude $<\alpha/2$). Ak $\mu<\mu_0$, potom pri rastúcom μ bude rásť aj $\Phi\left(u_{1-\alpha/2}+\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)$ a $\Phi\left(u_{1-\alpha/2}-\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)$ bude klesať (vždy bude $<\alpha/2$). Potom môžeme napísať približnú silu $\widetilde{\beta}^{*}$ ako

$$\begin{aligned}
\beta^{*} &= \Phi\left(u_{1-\alpha/2}-\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)+\Phi\left(u_{1-\alpha/2}+\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)\\
&< \Phi\left(u_{1-\alpha/2}+\frac{|\mu_0-\mu|}{\sigma}\sqrt{n}\right)=1-\Phi\left(u_{\alpha/2}-\frac{|\mu_0-\mu|}{\sigma}\sqrt{n}\right)=\widetilde{\beta}^{*}.
\end{aligned}$$

Pozn:

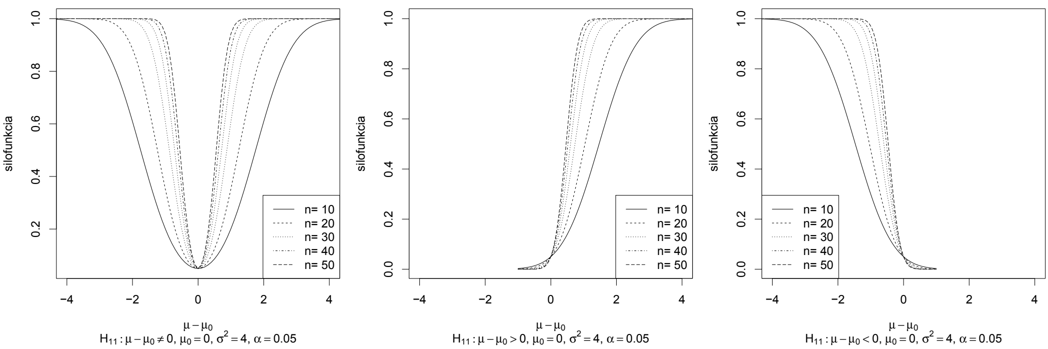
$$\beta_{11}(\mu)\leq\Phi_{\delta,1}\left(u_{\alpha/2}\right)-\Phi_{\delta,1}\left(-u_{\alpha/2}\right),$$

kde $\Phi_{\delta,1}(\cdot)$ je distribučná funkcia $N(\delta,1)$, $\delta=\frac{|\mu_0-\mu|}{\sigma}\sqrt{n}\neq 0$. Potom **minimálny rozsah** n pre rozdiel $|\mu-\mu_0|$ (pri nejakom α a β) je daný nasledovne

$$n\geq\left(\frac{u_{\alpha/2}+u_\beta}{c}\right)^2=\left(\frac{u_{\alpha/2}+u_\beta}{\frac{\mu-\mu_0}{\sigma}}\right)^2=\left(\frac{u_{\alpha/2}+u_\beta}{\mu-\mu_0}\right)^2\sigma^2.$$

Definícia 46 (Kritický obor a silofunkcia Z_W testu o μ) *Kritický obor a silofunkcia sú definované nasledovne (pozri obrázok 4.8):*

H_0	H_1	$\mathcal{W}$	$1-\beta(\mu)$
$\mu=\mu_0$	$\mu\neq\mu_0$	$\mathcal{W}_1=\left\{Z_W;\lvert Z_W\rvert\geq u_{\alpha/2}\right\}$	$1-\Phi\left(u_{\alpha/2}-\frac{\lvert\mu_0-\mu\rvert}{\sigma}\sqrt{n}\right)$
$\mu\leq\mu_0$	$\mu>\mu_0$	$\mathcal{W}_2=\{Z_W;Z_W\geq u_\alpha\}$	$1-\Phi\left(u_\alpha+\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)$
$\mu\geq\mu_0$	$\mu<\mu_0$	$\mathcal{W}_3=\{Z_W;Z_W\leq -u_\alpha\}$	$1-\Phi\left(u_\alpha-\frac{\mu_0-\mu}{\sigma}\sqrt{n}\right)$



Obr. 4.8: Silofunkcie asymptotického testu o μ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 47 (p-hodnora Z_W testu o μ) *Nech Z_W je* ***nejaká testovacia štatistika*** *a z_W je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-hodnota} = \begin{cases} 2Pr(Z_W \geq |z_W||H_{01}),\ ak & H_{11} : \mu \neq \mu_0 \\ Pr(Z_W \geq z_W|H_{02}),\ ak & H_{12} : \mu > \mu_0 \\ Pr(Z_W \leq z_W|H_{03}),\ ak & H_{13} : \mu < \mu_0 \end{cases}.$$

Definícia 48 (Waldove $100(1-\alpha)\%$ empirické IS pre μ) *Waldove $100(1-\alpha)\%$ empirické IS pre μ pre všetky tri typy hypotéz majú nasledovný tvar:*

$$\begin{array}{lll} H_0 & H_1 & \text{hranice } (d,h) \text{ pre } 100(1-\alpha)\% \text{ empirický IS} \\ \mu = \mu_0 & \mu \neq \mu_0 & \mathcal{CS}_{1-\alpha} = \left\{\mu_0 : \mu_0 \in \left(\overline{x} - u_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x} + u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)\right\} \\ \mu \leq \mu_0 & \mu > \mu_0 & \mathcal{CS}_{1-\alpha} = \left\{\mu_0 : \mu_0 \in \left(\overline{x} - u_{\alpha}\frac{\sigma}{\sqrt{n}}, \infty\right)\right\} \\ \mu \geq \mu_0 & \mu < \mu_0 & \mathcal{CS}_{1-\alpha} = \left\{\mu_0 : \mu_0 \in \left(-\infty, \overline{x} + u_{\alpha}\frac{\sigma}{\sqrt{n}}\right)\right\} \end{array}$$

Príklad 156 (plánovanie experimentu pre μ) *Majme náhodnú premennú $X \sim N(\mu, \sigma^2)$, kde μ je neznáma stredná hodnota a $\sigma^2 = 2^2$ je známy rozptyl. Chceme testovať $H_0 : \mu = 1$ oproti $H_1 : \mu = 4$. Kritický obor $\mathcal{W}_X$ je interval $(2, \infty)$, t.j. ak výberová hodnota μ je väčšia ako 2, H_0 zamietame. Ak výberová hodnota μ je menšia alebo rovná 2, H_0 nezamietame. Nájdite pravdepodobnosť chyby prvého druhu α a silu pri vyššie spomenutej alternatíve pre tento experiment.*

Riešenie (aj v R)
$\alpha = \Pr(X > 2|N(1,2)) = 0.31$
$\beta = \Pr(X \leq 2|N(4,2)) = 0.16$

```
alfa.jeden.vyber <- 1-pnorm(2,1,2) # 0.3085375
beta.jeden.vyber <- pnorm(2,4,2) # 0.15865532
```

Príklad 157 (plánovanie experimentu pre μ; test úspešnosti) *Majme náhodnú premennú $X \sim N(\mu, \sigma^2)$, kde μ je neznáma stredná hodnota a $\sigma^2 = 6^2$ je známy rozptyl, ktorá predstavuje skóre úspešnosti pre istú vekovú skupinu. Testujme H_0, že stredná hodnota skóre je rovná 40 oproti H_1, že stredná hodnota skóre nie je rovná 40.*

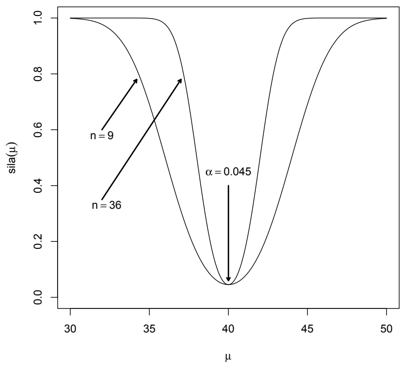
1. *Vypočítajte $\Pr(CHPD)$ pre $n = 9$, ak H_0 zamietame, keď $\overline{x} < 36$ alebo alebo $\overline{x} > 44$.*

2. *Vypočítajte $\Pr(CHPD)$ pre $n = 36$, ak H_0 zamietame, keď $\overline{x} < 38$ alebo alebo $\overline{x} > 42$.*

3. *Nakreslite silofunkciu pre $n = 9$ a $n = 36$ pre hodnoty μ medzi 30 a 50.*

Riešenie (aj v R; pozri obrázok 4.9)
(a) $\Pr(\text{CHPD}) = \Pr(\overline{X} < 36)|N(40, 6^2/9) + \Pr(\overline{X} > 44)|N(40, 6^2/9)$
$1 - \Pr(\text{CHDD}) = \Pr(\overline{X} < 36)|N(\mu, 6^2/9) + \Pr(\overline{X} > 44)|N(\mu, 6^2/9)$
(b) $\Pr(\text{CHPD}) = \Pr(\overline{X} < 38)|N(40, 6^2/36) + \Pr(\overline{X} > 42)|N(40, 6^2/36)$
$1 - \Pr(\text{CHDD}) = \Pr(\overline{X} < 38)|N(\mu, 6^2/36) + \Pr(\overline{X} > 42)|N(\mu, 6^2/36)$

```
alfa.jeden.vyber <- pnorm(36,40,6/sqrt(9))+1-pnorm(44,40,6/sqrt(9)) # 0.04550026
alfa.jeden.vyber <- pnorm(38,40,6/sqrt(36))+1-pnorm(42,40,6/sqrt(36)) # 0.0455003
mu <- seq(30,50,.01)
sila9 <- 1-pnorm(44,mu,6/sqrt(9))+pnorm(36,mu,6/sqrt(9))
sila36 <- 1-pnorm(42,mu,6/sqrt(36))+pnorm(38,mu,6/sqrt(36))
plot(mu, sila9, type="l", ylab=expression(sila(mu)), xlab=expression(mu), ylim=c(0,1))
lines(mu, sila36, type="l")
arrows(32,0.6,34.2,0.78, lwd=2, length=0.05)
arrows(32,0.35,37,0.78, lwd=2, length=0.05)
arrows(40,0.4,40,0.06, lwd=2, length=0.05)
text(32,0.58, expression(n==9))
text(32.3,0.33, expression(n==36))
text(40,0.45, expression(alpha==0.045))
```



Obr. 4.9: Schematický nákres silofunkcií pre $n = 9$ a $n = 36$

Príklad 158 (jednovýberový Z-test o strednej hodnote μ) *Nakreslite presnú silofunkciu pre test $H_0 : \mu = \mu_0$ (oproti $H_1 : \mu \neq \mu_0$, ak σ^2 je známe) do jedného obrázka pre $n = 10, 30, 50, 100, \mu_0 = 0$ a σ=1.*

Príklad 159 (jednovýberový Z-test o strednej hodnote μ) *Porovnajte presnú a približnú silofunkciu $\widetilde{\beta}^*(\mu)$ pre test $H_0 : \mu = \mu_0$ (oproti $H_1 : \mu \neq \mu_0$, ak σ^2 je známe) nakreslením oboch do jedného obrázka pre $n = 20, \mu_0 = 0$ a $\sigma = 1$.*

Príklad 160 (jednovýberový Z-test o strednej hodnote μ) *Odvoďte test pomerom vierohodnosti a $100(1-\alpha)\%$ interval spoľahlivosti pre hypotézu $H_0 : \mu = \mu_0$ oproti $H_0 : \mu \neq \mu_0$, ak σ^2 je známe.*

4.5 Vierohodnostné intervaly spoľahlivosti

Definícia 49 (obojstranný Waldov $100 \times (1-\alpha)\%$ empirický IS pre θ) *Obojstranný Waldov $100 \times (1-\alpha)\%$ empirický IS pre θ má tvar*

$$(d, h) = \left(\widehat{\theta} - t_{\alpha/2}\widehat{SE[\widehat{\theta}]}, \widehat{\theta} + t_{\alpha/2}\widehat{SE[\widehat{\theta}]}\right),$$

kde kritická hodnota $t_{\alpha/2}$ závisí na voľbe $\widehat{\theta}$, pozri kapitole 6, 7 a 8. Definícia 49 nie je všeobecne platná. Výnimky pozri napr. v kap. 6.2 a 7.2. Jednostranné intervaly spoľahlivosti pre θ je možné zadefinovať podobne.

Transformácia $g(\theta)$ nás často zaujíma aj preto, že „vylepšuje" funkciu vierohodnosti v zmysle regularity. Za platnosti regularity je náš stupeň dôvery v maximálne vierohodné odhady $g(\widehat{\theta})$ a $Var[g(\widehat{\theta})]$ väčší. Keďže maximálne vierohodný odhad $\widehat{\theta}$ je invariantný, môžeme pre $\mathcal{I}(g(\widehat{\theta}))$ písať

$$\mathcal{I}(g(\widehat{\theta})) = \mathcal{I}(\widehat{\theta}) \left[\frac{\partial}{\partial \theta} g(\theta)|_{\theta=\widehat{\theta}} \right]^{-2}$$

a pre štandardnú chybu

$$\widehat{SE[g(\widehat{\theta})]} = \widehat{SE(\widehat{\theta})} \left| \frac{\partial}{\partial \theta} g(\theta)|_{\theta=\widehat{\theta}} \right|$$

Definícia 50 (obojstranný Waldov $100 \times (1-\alpha)\%$ empirický IS pre $g(\theta)$) *Obojstranný Waldov $100 \times (1-\alpha)\%$ empirický IS pre $g(\theta)$ bude mať tvar*

$$(d_g, h_g) = \left(g(\widehat{\theta}) - t_{\alpha/2} \widehat{SE[g(\widehat{\theta})]}, g(\widehat{\theta}) + t_{\alpha/2} \widehat{SE[g(\widehat{\theta})]} \right),$$

kde kritická hodnota $t_{\alpha/2}$ závisí na voľbe $g(\widehat{\theta})$, pozri napr. kapitole 6.3, 6.4, 7.3 a 7.4. Definícia 50 nie je všeobecne platná. Jednostranné intervaly spoľahlivosti pre $g(\theta)$ je možné zadefinovať podobne.

Definícia 51 (Vierohodnostný $100 \times (1-\alpha)\%$ empirický IS pre θ) *Vierohodnostný $100 \times (1-\alpha)\%$ empirický IS pre θ má tvar*

$$\mathcal{CS}_{1-\alpha} = \left\{ \theta : -2 \ln \frac{L(\theta|\mathbf{x})}{L(\widehat{\theta}|\mathbf{x})} < \chi_1^2(\alpha) \right\}.$$

Hranice vierohodnostného $100 \times (1-\alpha)\%$ empirického IS pre θ odvodíme nasledovne

$$\Pr\left(\frac{L(\theta|\mathbf{x})}{L(\widehat{\theta}|\mathbf{x})} > c_\alpha \right) = \Pr\left(-2 \ln \frac{L(\theta|\mathbf{x})}{L(\widehat{\theta}|\mathbf{x})} < -2 \ln c_\alpha \right) = 1 - \alpha,$$

kde $c_\alpha = e^{-\frac{1}{2}\chi_1^2(\alpha)}$. Ak $1-\alpha = 0.95$, potom $c_\alpha = 0.1465001 \doteq 0.15$ (15% cut-off štandardizovanej funkcie vierohodnosti $\mathcal{L}(\theta|\mathbf{x}) = L(\theta|\mathbf{x})/L(\widehat{\theta}|\mathbf{x})$), ak $1-\alpha = 0.90$, potom $c_\alpha = 0.2585227 \doteq 0.26$ (26% cut-off) a ak $1-\alpha = 0.99$, potom $c_\alpha = 0.0362452 \doteq 0.04$ (4% cut-off).

Oba IS pre θ sú vytvorené na základe asymptotickej teórie, ale vierohodnostný interval má lepšie asymptotické vlastnosti ako Waldov IS. Waldov IS je správny, ak platí

$$\frac{\widehat{\theta} - \theta}{\widehat{SE[\widehat{\theta}]}} \overset{\mathcal{D}}{\sim} N(0,1).$$

Na základe invariantnosti pomeru vierohodnosti je vierohodnostný IS správny, ak existuje transformácia $g(\cdot)$ (ktorá nemusí byť známa), pre ktorú platí

$$\frac{g(\widehat{\theta}) - g(\theta)}{\widehat{SE[g(\widehat{\theta})]}} \overset{\mathcal{D}}{\sim} N(0,1).$$

Nech (D_g, H_g) je $100(1-\alpha)\%$ IS pre $g(\theta)$, t.j. $D_g < g(\theta) < H_g$. Na základe asymptotickej normality $g(\widehat{\theta})$ je IS (D_g, H_g) vierohodnostný IS pre $g(\theta)$ určený $c_\alpha\%$ cut-off štandardizovanej funkcie vierohodnosti $\mathcal{L}(\theta|\mathbf{x})$. Invariancia pomeru vierohodnosti implikuje tvrdenie, že vierohodnostný IS pre θ určený $c_\alpha\%$ cut-off štandardizovanej funkcie vierohodnosti je rovný

$$\left((g(D_g))^{-1}, (g(H_g))^{-1} \right), \text{ kde } g(D_g))^{-1} < \theta < (g(H_g))^{-1},$$

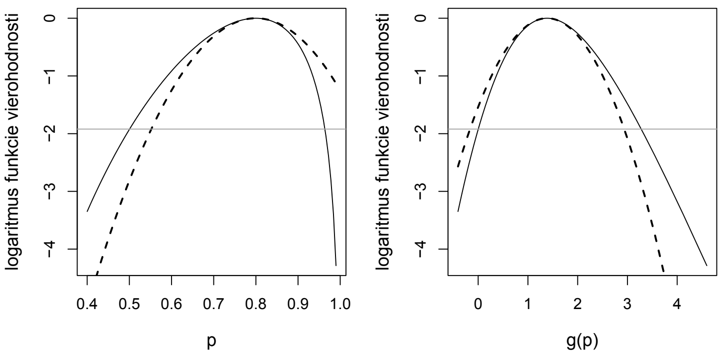
ktorý má koeficient spoľahlivosti rovný $100 \times (1-\alpha)\%$. Toto tvrdenie platí, aj keď $g(\cdot)$ neexistuje a navyše vierohodnostný IS je automaticky IS s najlepšou možnou (normalizujúcou) transformáciou. Najväčším problémom Waldovho IS je, že ak $\widehat{\theta}$ nemá normálne rozdelenie a ak chceme θ transformovať kvôli „vylepšeniu" normality, musíme túto transformáciu poznať. Vierohodnostný IS automaticky aplikuje túto najlepšiu transformáciu. Z toho vyplýva, že aplikovateľnosť vierohodnostného IS je oveľa širšia ako Waldovho IS. Okrem toho je použitie vierohodnostného IS bezpečnejšie vo vzťahu k asymptotickej normalite. Tento fakt zdôrazňuje aj to, že Waldova testovacia štatistika nie je invariantná na transformáciu $g(\cdot)$, kde dôvodom tejto neinvariantnosti je zvyšok v Taylorovom rozvoji. Jeho efekt môže byť o to väčší, o čo nelineárnejšie je $g(\cdot)$ v okolí $\widehat{\theta}$.

Príklad 161 (vylepšená vierohodnosť pomocou $g(\theta)$) *Nakreslite (a) logaritmus funkcie vierohodnosti parametra p binomického rozdelenia $Bin(N, p)$, kde $N = 10$ a $n = 8$, superponovaný jeho kvadratickou aproximáciou. Nakreslite (b) logaritmus funkcie vierohodnosti $g(p) = logit(p) = \ln \frac{p}{1-p}$ (pri rovnakom zadaní N a n ako v (a)) superponovaný jeho kvadratickou aproximáciou. Je funkcia vierohodnosti $g(p)$ regulárnejšia ako funkcia vierohodnosti pre p? (c) Vypočítajte Waldov a vierohodnostný $100 \times (1-\alpha)\%$ empirický IS pre p. (d) Vypočítajte Waldov $100 \times (1-\alpha)\%$ empirický IS pre $g(p)$ z (a) a (b) a transformujte ho späť do originálnej škály. (e) Ukážte, že vierohodnostný IS pre p v škále p je identický s vierohodnostným IS v škále $g(p)$ z (a) a (b) po jeho spätnej transformácii do originálnej škály.*

Riešenie (aj v R, pozri obrázok 4.10**)**
Waldov $100 \times (1-\alpha)\%$ empirický IS pre p:
$\widehat{p} = \frac{8}{10} = 0.8$; $\widehat{SE[\widehat{p}]} = \sqrt{\frac{\widehat{p}(1-\widehat{p})}{N}} = 0.13$.
$(d, h) = \left(\widehat{p} - u_{\alpha/2}\widehat{SE[\widehat{p}]}, \widehat{p} + u_{\alpha/2}\widehat{SE[\widehat{p}]}\right) = (0.55, 1.05)$, kde je horná hranica väčšia ako jedna.
Vierohodnostný $100 \times (1-\alpha)\%$ empirický IS pre p:
$\mathcal{CS}_{1-\alpha} = \left\{p : -2\ln\frac{L(p|\mathbf{x})}{L(\widehat{p}|\mathbf{x})} \leq 3.84\right\}$, kde $(d, h) = (0.50, 0.96)$,
Waldov $100 \times (1-\alpha)\%$ empirický IS pre $g(p)$:
$g(\widehat{p}) = \ln\frac{\widehat{p}}{1-\widehat{p}} = \log\frac{0.8}{0.2} = 1.39$.
$\frac{\partial}{\partial p}g(p) = \frac{1}{p} + \frac{1}{1-p}$; $\widehat{SE[g(\widehat{p})]} = \widehat{SE[\widehat{p}]}\left(\frac{1}{\widehat{p}} + \frac{1}{1-\widehat{p}}\right) = \sqrt{\frac{\widehat{p}(1-\widehat{p})}{N}}\left(\frac{1}{\widehat{p}} + \frac{1}{1-\widehat{p}}\right) = \sqrt{\frac{1}{n} + \frac{1}{N-n}} = 0.79$.
Potom $(d_g, h_g) = (-0.16, 2.94)$, ktorý transformujeme späť do originálnej škály, kde $(d, h) = (0.46, 0.95)$.

```
x <- 8; N <- 10
pcka <- seq(0.4,.99,length=100)
like  <- dbinom(8,10,pcka)
like1 <- like/max(like)
ll    <- log(like1)
vier.is.p <- range(pcka[like1 > cutoff]) # 0.5013131 0.9602020

windows(8,4)
par(mfcol=c(1,2),mar=c(4.5,4.5,1,1))
p.hat <- x/N
i.hat <- N/p.hat/(1-p.hat)
la    <- -i.hat/2*(pcka-p.hat)^2
ra    <- range(ll)
wald.is.p <- p.hat + c(-1,1)*qnorm(0.975)*sqrt(1/i.hat)
wald.is.p # 0.552082 1.047918
cutoff <- exp(-1/2*qchisq(0.95,1)) # 0.1465001

plot(pcka,ll,type="n",xlab="p",ylab="logaritmus funkcie vierohodnosti",
     lwd=2,cex.lab=1.2)
lines(pcka,ll,lwd=1)
lines(pcka,la,lty=2,lwd=2)
abline(h=log(cutoff),col="gray")

gpcka <- log(pcka) - log(1-pcka)
gp.hat <- log(p.hat) - log(1-p.hat)
i.hat <- x*(N-x)/N
lgp <- -i.hat/2*(gpcka-gp.hat)^2
x <- (gp.hat + c(-1,1)*qnorm(0.975)*sqrt(1/i.hat)) # -0.1631932 2.9357819
wald.is.gp <- exp(x)/(1+exp(x))
wald.is.gp # 0.4593 0.9496
```

```
plot(gpcka, ll, type="n", xlab="g(p)", ylab="logaritmus funkcie vierohodnosti",
     lwd=2, cex.lab=1.2)
lines(gpcka, ll, lwd=1)
lines(gpcka, lgp, lty=2, lwd=2)
abline(h=log(cutoff), col="gray")

b <- range(gpcka[like1 > cutoff])
vier.is.gp <- exp(b)/(1+exp(b)) # 0.5013131 0.9602020
```



Obr. 4.10: Funkcia vierohodnosti (plnou čiarou) pre p (vľavo) a $g(p)$ vpravo spolu so superponovanými kvadratickými aproximáciami (čiarkovanou čiarou); horizontálna priamka predstavuje hodnotu *cut-off*

***Výpočet hraníc vierohodnostného intervalu spoľahlivosti**

V príklade 161 sme hranice vierohodnostného empirického IS počítali pre θ patriace nejakému dopredu špecifikovanému intervalu s hranicami θ_D a θ_H, dosadením sekvencie hodnôt $\theta_i \in \langle\theta_D, \theta_H\rangle$, $i = 1, 2, ..., M$ (pre napr. $M = 100$ tak, aby boli vzdialenosti θ_i čo najmenšie), do funkcie $-2\ln\mathcal{L}(\theta_i|\mathbf{x})$. Hranice vierohodnostného $100\times(1-\alpha)\%$ empirického IS pre θ budú potom predstavovať dolnú a hornú hranicu intervalu definovaného pomocou nerovnosti $-2\ln\mathcal{L}(\theta|\mathbf{x}) < \chi^2_1(\alpha)$, ktoré vypočítame pomocou funkcie range(ll), kde ll je funkcia $-2\ln\mathcal{L}(\theta|\mathbf{x})$. Alternatívnou možnosťou je použitie **metódy bisekcie**[2] (Givens a Hoeting, 2005).

Nech pre nejaké $\theta_{01}, \theta_{02} \in \langle\theta_D, \theta_H\rangle$ platí $f(\theta_{01})f(\theta_{02}) < 0$ (t.j. $f(\cdot)$ má na intervale $\langle\theta_{01}, \theta_{02}\rangle$ aspoň jeden koreň), kde $f(\theta) = -2\ln\mathcal{L}(\theta|\mathbf{x}) - \chi^2_1(\alpha) = 0$. Ak má prvá derivácia $f(\cdot)$ konštantné znamienko, potom existuje práve jeden koreň $\theta^* \in \langle\theta_{01}, \theta_{02}\rangle$ funkcie $f(\theta)$. Metóda bisekcie hľadá koreň nasledovným iteračným algoritmom:

1. inicializácia metódy pomocou vhodne zvoleného štartovacieho parametra $\theta^{(0)}$, pre ktorý platí $\theta^{(0)} = (\theta_{01} + \theta_{02})/2$ a $i = 1$,

2. substitúcia hraníc θ_{01} a θ_{02} nasledovne

$$\langle\theta_{i1}, \theta_{i2}\rangle = \begin{cases} \left\langle\theta_{i-1,1}, \theta^{(i-1)}\right\rangle, & \text{ak } f(\theta_{i-1,1})f(\theta^{(i-1)}) < 0 \\ \left\langle\theta^{(i-1)}, \theta_{i-1,2}\right\rangle, & \text{ak } f(\theta_{i-1,1})f(\theta^{(i-1)}) > 0 \end{cases},$$

   ak $f(\theta^{(i-1)}) = 0$, potom algoritmus skončí, ak nie, nasleduje

3. výpočet stredu intervalu $\theta^{(i)} = (\theta_{i1} + \theta_{i2})/2$,

4. iterácia (2) a (3) pomocou *relatívnej aproximačnej chyby* (hovoríme o *relatívnom konvergenčnom kritériu*) pokiaľ nebude platiť

$$\frac{\left|\theta^{(i)} - \theta^{(i-1)}\right|}{\left|\theta^{(i-1)}\right|} < \epsilon,$$

[2] Ak sú metóda bisekcie a jej modifikácie aplikované na skóre funkciu $S(\theta)$, kde $f(\theta) = l(\theta|\mathbf{x})$, hľadáme maximum logaritmu funkcie vierohodnosti, a preto sa značenie všade posunie o jednu deriváciu smerom doprava. Ak sa v metódach z kapitoly 2.4 *Maximalizácia funkcie vierohodnosti posunie značenie o jednu deriváciu smerom doľava, je možné tieto metódy použiť aj na hľadanie hraníc vierohodnostného IS .

kde ϵ je vhodne zvolené malé číslo (prahová hodnota), pomocou *absolútnej aproximačnej chyby* (hovoríme o *absolútnom konvergenčnom kritériu*) pokiaľ nebude platiť

$$\left|\theta^{(i)} - \theta^{(i-1)}\right| < \epsilon,$$

alebo pomocou

$$\left|f(\theta^{(i)})\right| < \epsilon.$$

Metóda bisekcie konverguje lineárne (rád metódy je rovný jednej), t.j. chyba sa zmenšuje v každom kroku dvakrát. V i-tom kroku bude menšia alebo rovná ako $(\theta_{02} - \theta_{01})/2^{i+1}$. Pretože je konvergencia bisekčnej metódy pomalá (vyžaduje veľké množstvo iterácií na dosiahnutie stanovenej presnosti), bolo vytvorených niekoľko jej modifikácií, ktoré sa dajú zahrnúť pod anglický pojem *bracketing methods*, t.j. metód ktoré ohraničujú koreň vnútri sekvencie vnorených intervalov zmenšujúcej sa dĺžky. Z nich je najčastejšie používanou je **Brentova metóda** (Brent (1973); niekedy nazývaná aj **Brent-Dekkerova metóda**) pomenovaná podľa Richarda Peirce Brenta (1949−). Kombinuje metódu bisekcie s inverznou lineárnou alebo kvadratickou interpoláciou. Ak je použitá inverzná lineárna interpolácia, potom ide o metódu ekvivalentnú **metóde sečníc**, ktorá modifikuje krok (3) nasledovne

$$\theta^{(i)} = \begin{cases} \theta^{(i-1)} - \frac{\theta^{(i-1)} - \theta^{(i-2)}}{f(\theta^{(i-1)}) - f(\theta^{(i-2)})} f(\theta^{(i-1)}), & \text{ak } f(\theta^{(i-1)}) \neq f(\theta^{(i-2)}) \\ (\theta_{i1} + \theta_{i2})/2, & \text{inak} \end{cases},$$

kde aproximácia prvej derivácie $f'(\theta^{(i-1)}) \approx \frac{f(\theta^{(i-1)}) - f(\theta^{(i-2)})}{\theta^{(i-1)} - \theta^{(i-2)}}$. Metóda sečníc má jednoduchú geometrickú interpretáciu – bod $\theta^{(i)}$ je priesečník sečnice vedenej bodmi $[\theta^{(i-1)}, f(\theta^{(i-1)})]$ a $[\theta^{(i-2)}, f(\theta^{(i-2)})]$ s x-ovou osou. Ak je $f(\theta)$ dvakrát diferencovateľná, funkcia $f(\theta)$ má jednoduchý koreň ($f'(\theta) \neq 0$ pre každé $\theta \in \langle\theta_D, \theta_H\rangle$), potom metóda konverguje, pokiaľ zvolíme θ_{01}, θ_{02} dostatočne blízko hľadaným hraniciam IS (rád metódy je rovný $(1 + \sqrt{5})/2 \doteq 1.618$).

Brentova metóda modifikuje metódu sečníc na zrýchlenie konvergencie pridaním ďalších podmienok v kroku (3). Je implementovaná v R vo funkcii uniroot(f, interval,tol,...), ktorá hľadá korene funkcie f v danom intervale špecifikovanom v argumente interval. Konvergencia metódy môže byť kontrolovaná argumentom tol, čo je prahová hodnota ϵ. Pri hľadaní hraníc vierohodnostného $100 \times (1 - \alpha)\%$ IS je potrebné funkciu uniroot() použiť dvakrát nasledovne:

1. pre dolnú hranicu IS definujeme štartovací interval ako $\left\langle \theta_D, \widehat{\theta} \right\rangle$,
2. pre hornú hranicu IS definujeme štartovací interval ako $\left\langle \widehat{\theta}, \theta_H \right\rangle$.

Vo výstupoch funkcie uniroot() bude pre (1) odhad dolnej hranice IS $\widehat{\theta}_D$ a pre (2) odhad hornej hranice IS $\widehat{\theta}_H$ (root), hodnota funkcie f v odhadnutých hraniciach, pre (1) $f(\widehat{\theta}_D)$ a pre (2) $f(\widehat{\theta}_H)$, ozn. froot a odhadnutá výchylka od skutočného koreňa estim.perc. Ak má funkcia $f(\theta)$ viac ako jeden argument, argument θ musí byť ako prvý.

Príklad 162 (vylepšená vierohodnosť pomocou $g(\theta)$; pokrač.) *Vypočítajte vierohodnostný 95% empirický IS pre p pomocou Brentovej metódy pre dáta z príkladu 161. Porovnajte tento IS s IS vypočítaným v príklade 161.*

Riešenie (aj v R)
Vierohodnostný 95% empirický IS pre p:
$\mathcal{CS}_{1-\alpha} = \left\{p : -2\ln\frac{L(p|\mathbf{x})}{L(\widehat{p}|\mathbf{x})} \leq 3.84\right\}$, kde $(d, h) = (0.50, 0.96)$.
Hranice IS vypočítané oboma metódami sa zhodujú na dve desatinné miesta.

```
x <- 8; N <- 10
p.hat <- x/N
"ll" <- function(p,p.hat) {
-2*log(dbinom(8,10,p)/dbinom(8,10,p.hat))-qchisq(0.95,1)
 }
vier.is.p.brent <- c(uniroot(ll,c(0,p.hat),p.hat=x/N)$root,
                     uniroot(ll,c(p.hat,1),p.hat=x/N)$root)
vier.is.p.brent # 0.5005669 0.9636336
```

5 Testy dobrej zhody

V štatistickej inferencii často predpokladáme, že dáta pochádzajú z normálneho rozdelenia $N(\mu, \sigma^2)$. Na identifikáciu rozdelenia, z ktorého dáta pochádzajú, potrebujeme **testy dobrej zhody**, kde nulovou hypotézou je tvrdenie o (kumulatívnej) distribučnej funkcii. Ak sú pre nulovú hypotézu všetky potrebné parametre rozdelenia špecifikované, hovoríme o **jednoduchej nulovej hypotéze**. Ak nulová hypotéze presne nešpecifikuje všetky potrebné parametre rozdelenia, hovoríme o **zloženej nulovej hypotéze**. V nasledujúcich odstavcoch prezentujeme dva typy testov – χ^2 **test dobrej zhody** (primárne vytvorený pre diskrétne dáta) a **Kolmogrov-Smirnovov test dobrej zhody** (vytvorený pre spojité dáta).

5.1 χ^2 test dobrej zhody

Majme jednoduchý náhodný výber s rozsahom n s neznámou populačnou distribučnou funkciou $F(x)$. Nulovú hypotézu špecifikujeme ako „$F(x)$ má nejaké známe rozdelenie $F_0(x)$ pre všetky x" alebo „$F(x)$ je rovné nejakej známej distribučnej funkcii $F_0(x)$ pre všetky x", t.j. $H_0 : F(x) = F_0(x)$ pre všetky x. Alternatívnu hypotézu špecifikujeme ako „$F(x)$ nie je rovné nejakej známej distribučnej funkcii $F_0(x)$ pre nejaké x", t.j. $H_1 : F(x) \neq F_0(x)$ pre nejaké x.

χ^2 test dobrej zhody je založený na normalizovanej testovacej štatistike, ktorá hodnotí vertikálne rozdiely medzi pozorovanými a očakávanými hodnotami (ak H_0 platí) v k disjunktných kategóriách. Dáta je teda potrebné rozdeliť do týchto kategórií na základe vopred určenej schémy a vypočítať početnosti v nich. Tieto nazývame **pozorované početnosti**. Ak H_0 úplne špecifikuje populáciu, z ktorej pochádza náhodný výber, môžeme vypočítať pravdepodobnosť, že náhodné pozorovanie padne do jednej z k kategórií, čomu hovoríme **očakávané pravdepodobnosti**. Pokiaľ tieto pravepodobnosti vynásobíme rozsahom n, dostaneme **očakávané početnosti** pre každú kategóriu za platnosti H_0.

Ak H_0 platí, rozdiel medzi pozorovanými a očakávanými početnosťami bude malý. χ^2 **testovacia štatistika** na testovanie H_0 oproti H_1, ak je H_0 úplne špecifikovaná, má tvar

$$\chi^2 = \sum_{j=1}^{k} \frac{(\text{pozorované}_j - \text{očakávané}_j)^2}{\text{očakávané}_j} = \sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j} \overset{\mathcal{D}}{\sim} \chi^2_{df},$$

kde O_j sú pozorované početnosti, E_j očakávané početnosti v každej z k kategórii. Stupne voľnosti $df = k - 1$, ak je H_0 jednoduchá. Test sa nazýva **jednovýberový** χ^2**-test dobrej zhody**. Ak je však H_0 zložená, musíme odpočítať jednotku za každý odhadovaný parameter, t.j. ak je počet odhadovaných parametrov k_p, potom $df = k - k_p - 1$, kde $k_p < k - 1$. Veľké hodnoty realizácie χ^2, ozn. χ^2_{obs}, budú indikovať nekonzistenciu dát s H_0 (nedostatok dôkazov na jej zamietnutie).

Príklad 163 (χ^2-test dobrej zhody) *Majme dáta* Grades *z knižnice* PASWR*, ktoré reprezentujú SAT skóre (n = 200) náhodne vybranej vzorky študentov z jednej univerzity v USA. Otestujte na hladine významnosti $\alpha = 0.05$, či majú dáta normálne rozdelenie. Použite intervaly $\langle \mu - 3\sigma, \mu - 2\sigma \rangle$, $(\mu - 2\sigma, \mu - \sigma \rangle$, $(\mu - \sigma, \mu \rangle$, $(\mu, \mu + \sigma \rangle$, $(\mu + \sigma, \mu + 2\sigma \rangle$ a $(\mu + 2\sigma, \mu + 3\sigma \rangle$. Nakreslite histogram použitím vyššie spomenutých intervalov a superponujte ho s očakávanými hodnotami SAT skóre v každej kategórii, keď $F_0(x) \sim N(\mu, \sigma^2)$.*

Riešenie (aj v R; pozri tabuľku 5.1)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: SAT skóre majú normálne rozdelenie (ich distribučná funkcia pochádza z normálneho rozdelenia) oproti H_1: SAT skóre nemajú normálne rozdelenie (ich distribučná funkcia nepochádza z normálneho rozdelenia).

- **matematická formulácia** – testujeme $H_0 : F(x) = F_0(x) \sim N(\mu, \sigma^2)$ oproti $H_1 : F(x) \neq F_0(x)$ pre nejaké x.

2. **Testovacia štatistika** – Stredná hodnota μ a rozptyl σ^2 sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetický priemer $\overline{x} = 1134.65$ a rozptyl $s^2 = 21201.89$ (smerodajná odchýlka $s = 145.61$).

```
library(PASWR) # nacitanie kniznice
attach(Grades)
priemer <- mean(sat) # 1134.6500
smerodch <- sd(sat) # 145.6087
n <- length(sat) # 200
```

Intervaly zo zadania sú $\langle 697.8240, 843.4326\rangle$, $(843.4326, 989.0413\rangle$, $(989.0413, 1134.6500\rangle$, $(1134.6500, 1280.2587\rangle$, $(1280.2587, 1425.8674\rangle$ a $(1425.8674, 1571.4760\rangle$. Dáta budú rozdelené na kategórie nasledovne

```
bin <- seq(priemer-3*smerodch, priemer+3*smerodch, smerodch) # kategorie
table(cut(sat, breaks=bin))
poz.poc <- hist(sat, breaks=bin, plot=F)$counts # pozorovane pocetnosti
ocak.prav <- c(pnorm(-2), pnorm(-1:2) - pnorm(-2:1),1-pnorm(2))
ocak.poc <- n*ocak.prav # ocakavane pocetnosti
VYSL <- cbind(ocak.prav, ocak.poc, poz.poc)
```

Tabuľka 5.1: Očakávané pravdepodobnosti a početnosti a pozorované početnosti pre SAT skóre

interval	$f_0(x)$	$nf_0(x)$	n_i
$\langle\mu - 3\sigma, \mu - 2\sigma\rangle$	0.02275	4.55003	4
$(\mu - 2\sigma, \mu - \sigma\rangle$	0.13591	27.18102	27
$(\mu - \sigma, \mu\rangle$	0.34135	68.26895	65
$(\mu, \mu + \sigma\rangle$	0.34135	68.26895	80
$(\mu + \sigma, \mu + 2\sigma\rangle$	0.13591	27.18102	21
$(\mu + 2\sigma, \mu + 3\sigma\rangle$	0.02275	4.55003	3

Pozorovaná testovacia štatistika je potom rovná $\chi^2_{\text{obs}} = 4.17$.

```
chisq.obs <- sum((poz.poc-ocak.poc)^2/ocak.poc) # 4.173654
```

3. **Zamietacia oblasť** – H_0 bude zamietnutá, ak $\chi^2_{\text{obs}} > \chi^2_\alpha(df)$, kde $df = k - 1 = 6 - 1 = 5$, ak by sme neodhadovali žiaden parameter rozdelenia. My však odhadujeme dva, preto $df = 6-2-1 = 3$, $\chi^2_{1-\alpha}(df) = \chi^2_{0.05}(3)$ je kritická hodnota.

4. **Štatistický záver** – p-hodnota=0.24.

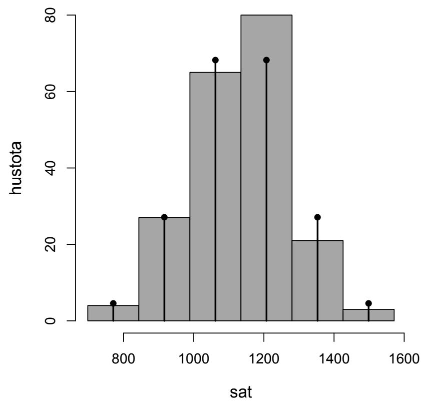
```
p.hodn <- 1-pchisq(chisq.obs,5) # 0.5246948 [ak df=5]
p.hodn <- 1-pchisq(chisq.obs,3) # 0.2433129
```

Kedže p-hodnota nie je menšia ako hladina významnosti $\alpha = 0.05$ (ekvivalentne χ^2_{obs} nepatrí do zamietacej oblasti), H_0 nezamietame na hladine významnosti α. Je možné použiť aj funkciu chisq.test(x=poz.poc,p=ocak.prav).

5. **Slovný záver** – Nemáme dostatok dôkazov na zamietnutie nulovej hypotézy o tom, že skutočnou distribučnou funkciou SAT skóre je distribučná funkcia normálneho rozdelenia.

Histogram superponovaný s očakávanými hodnotami SAT skóre v každej kategórii nakreslíme nasledovne (pozri obrázok 5.1)

```
windows(5,5)
par(mar=c(4.5,4.5,1,1))
hist(sat, breaks=bin, col="gray", ylab="hustota", freq=TRUE, main="", cex.lab=1.2)
x <- bin[2:7] - smerodch/2 # centrovanie do stredu intervalov
lines(x, ocak.poc, type="h", lwd=2)
points(x, ocak.poc, pch=16)
```



Obr. 5.1: Histogram superponovaný s očakávanými hodnotami SAT skóre

Príklad 164 (χ^2-test dobrej zhody; pokrač.) *Zopakujte výpočet z predchádzajúceho príkladu na intervaloch definovaných pomocou hraníc:*
(a) kvartilové hranice – $x_{min}, \widetilde{x}_{0.25}, \widetilde{x}_{0.50}, \widetilde{x}_{0.75}, x_{max}$*;*
(b) decilové hranice – $x_{min}, \widetilde{x}_{0.1}, \widetilde{x}_{0.2}, \ldots, \widetilde{x}_{0.8}, \widetilde{x}_{0.9}, x_{max}$*.*
Nakreslite histogram použitím vyššie spomenutých intervalov a superponujte ho s očakávanými hodnotami SAT skóre v každej kategórii, keď $F_0(x) \sim N(\mu, \sigma^2)$*. Porovnajte výsledky s výsledkami predchádzajúceho príkladu.*

Príklad 164 vysvetľuje efekt zmeny intervalov na výsledok testu dobrej zhody s normálnym rozdelením.

Príklad 165 (χ^2-test dobrej zhody) *Johann Gregor Mendel vo svojich pokusoch s krížením rastlín hrachu (Pisum sativum) študoval dedičnosť siedmych rôznych znakov. V každom z pokusov, pri sledovaní jedného znaku, získal po krížení dvoch čistých línií (t.j. dominantného homozygota AA s recesívnym homozygotom aa) generáciu, v ktorej mali všetky rastliny rovnaký fenotyp (t.j. heterozygoti Aa). Po ich samooplodnení (čo je prirodzený spôsob rozmnožovania hrachu) získal ďalšiu generáciu, v ktorej sa vyskytovali sledované znaky v dvoch formách, a to zakaždým v pomere veľmi blízkom 3:1. Jedným zo znakov, ktoré študoval, bola farba semien. Po krížení* 258 *hybridov získal celkove 8023 semien, z ktorých* 6022 *bolo žltých a* 2001 *zelených (Matalová, 2008). Otestujte platnosť fenotypového štiepneho pomeru* $3:1$ *na hladine významnosti* $\alpha = 0.05$*.*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – H_0 : Fenotypový štiepny pomer je rovný $3:1$ oproti H_1 : fenotypový štiepny pomer nie je rovný $3:1$.
   - **matematická formulácia** – $H_0 : \mathbf{p} = \mathbf{p}_0$, kde $\mathbf{p} = (p_1, p_2)$, $\mathbf{p}_0 = (p_{01}, p_{02})$, $p_{01} = 0.75$ a $p_{02} = 0.25$, oproti $H_1 : \mathbf{p} \neq \mathbf{p}_0$.
2. **Testovacia štatistika** – $N = 6022 + 2001 = 8023$, očakávané početnosti $8023 \times 0.75 = 6017.25$ a $8023 \times 0.25 = 2005.75$.

```
poz.poc <- c(6022,2001)
N <- sum(poz)
ocak.prav <- c(0.75,0.25)
ocak.poc <- N*ocak.prav
```

   Pozorovaná testovacia štatistika je potom rovná $\chi^2_{\text{obs}} \doteq 0.015$.

```r
chisq.obs <- sum((poz.poc-ocak.poc)^2/ocak.poc) # 0.01499855
```

3. **Zamietacia oblasť** – H_0 bude zamietnutá, ak $\chi^2_{\text{obs}} > \chi^2_{df}(\alpha) \doteq 3.84$, kde $\chi^2_{df}(\alpha)$ je kritická hodnota χ^2 rozdelenia s $df = 2 - 1 = 1$ a $\alpha = 0.05$.

```r
chisq.krit <- qchisq(1-0.05,1) # 3.841459
```

4. **Štatistický záver** – p-hodnota = 0.9.

```r
p.hodn <- 1-pchisq(chisq.obs, df=1) # 0.902528
```

Kedže p-hodnota nie je menšia ako hladina významnosti $\alpha = 0.05$ (ekvivalentne χ^2_{obs} nepatrí zamietacej oblasti), H_0 nezamietame na hladine významnosti α. Je možné použiť aj funkciu chisq.test() nasledovne

```r
VYSL <- chisq.test(x=c(6022,2001),p=c(0.75,0.25))
VYSL <- c(VYSL$stat,VYSL$param,VYSL$p.val)
names(VYSL) <- c("test-stat","df","p-hodn")
VYSL
#test-stat        df      p-hodn
#    0.0150    1.0000     0.9025
```

5. **Slovný záver** – Nemáme dostatok dôkazov na zamietnutie nulovej hypotézy o tom, že skutočný fenotypový pomer je 3 : 1.

Testovať H_0 oproti H_1 môžeme aj MC testom, kde odhadneme p-hodnotu simulačne pomocou $M = 100000$ simulácií z $Bin(N, p)$, kde $N = 8023, p = 0.75$.

```r
"chisq.stat" <- function(poz.poc,ocak.prav) {
                N <- sum(poz.poc)
                ocak <- N*ocak.prav
                stat <- sum((poz.poc-ocak.poc)^2/ocak.poc)
                return(stat)
}
M <- 100000
ocak.prav <- c(0.75,0.25)
chisq.obs <- chisq.stat(poz.poc,ocak.prav) # 0.01499855
chisq.sim <- numeric(M)
for (i in 1:M){
x1 <- rbinom(1,N,0.75)
x2 <- N - x1
chisq.sim[i] <- chisq.stat(c(x1,x2),ocak.prav)
}
p.hodn <- mean(chisq.sim>chisq.obs) # 0.89787
```

P-hodnota vypočítaná na základe MC simulácie je veľmi podobná vypočítanej asymptoticky (na dve desatinné miesta sú obe identické), keďže máme veľký počet opakovaní, ako aj veľké početnosti.

Príklad 166 (χ^2-test dobrej zhody; početnosti úmrtí) *Otestujte zhodu početností X Pruských armádnych jednotiek, v ktorých nastalo n úmrtí zapríčinených kopnutím koňom za rok (pozri príklad 78, tabuľka 2.7) s Poissonovým rozdelením s parametrom λ, t.j. $X \sim Poiss(\lambda)$ na hladine významnosti $\alpha = 0.05$.*

Príklad 167 (χ^2-test dobrej zhody; početnosti chlapcov) *Otestujte zhodu početností rodín X s n chlapcami (pozri príklad 81, tabuľku 2.8 a 2.9) s binomickým rozdelením s parametrami N a π, t.j. $X \sim Bin(N, \pi)$ na hladine významnosti $\alpha = 0.05$.*

Príklad 168 (χ^2-test dobrej zhody; úrazy robotníkov) *Otestujte zhodu početností robotníkov X s n úrazmi v továrni (pozri príklad 82, tabuľka 2.10 a 2.11; príklad 109, tabuľka 2.14)*
(a) s Poissonovým rozdelením s parametrom λ, t.j. $X \sim Poiss(\lambda)$ a
(b) s negatívne binomickým rozdelním s parametrami α a π, t.j. $Negbinom(\alpha, \pi)$ na hladine významnosti $\alpha = 0.05$.

Príklad 169 (χ^2-test dobrej zhody; početnosti mužov) *Nech* $\Pr(muž) = 0.515$ *znamená pravdepodobnosť výskytu mužov v populácii (pozri príklad 78). Nech sa početnosti mužov* X *v populácii správajú podľa modelu* $Bin(N,p)$*, nazvime ich (v tomto príklade) pozorované početnosti. Aproximujme* X *normálnym rozdelením* $N(Np, Npq)$ *a prislúchajúce početnosti nazvime teoretické početnosti (počítame ich v* $n = 0, 1, \ldots, N$*). Otestujte zhodu pozorovaných a teoretických početností na hladine významnosti* $\alpha = 0.05$*. Nakreslite (1) hustotu normálneho rozdelenia superponovanú pravdepodobnosťou funkciou binomického rozdelenia a (2) distribučnú funkciu normálneho rozdelenia superponovanú distribučnou funkciou binomického rozdelenia. Rozsahy* N *volíme nasledovne (a)* $N = 20$*, (b)* $N = 50$ *a (c)* $N = 500$*.*

Príklad 170 (χ^2-test dobrej zhody; fetálna aktivita) *Nech* X *predstavuje početnosti päťsekundových intervalov (z 240) v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom, v ktorých sa plod ovce* n*-krát pohol (pozri tabuľku 5.2; Leroux a Puterman (1992)). Vypočítajte očakávané početnosti za predpokladu, že* $X \sim Poiss(\lambda)$*. Otestujte zhodu pozorovaných a teoretických početností na hladine významnosti* $\alpha = 0.05$*.*

Tabuľka 5.2: Pozorované početnosti m_n päťsekundových intervalov v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom, v ktorých sa plod ovce n-krát pohol

n	0	1	2	3	4	5	6	7
pozorované m_n	182	41	12	2	2	0	0	1

Riešenie (aj v R; pozri tabuľku 5.3 a 5.4)
$\widehat{\lambda} = \frac{\sum_n n m_n}{\sum_n m_n} = \frac{86}{240} = 0.358$. Potom budú čakávané početnosti nasledovné Vzhľadom ku nízkym

Tabuľka 5.3: Očakávané početnosti m_n päťsekundových intervalov v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom (zaokrúhlené na nula desatinných miest), v ktorých sa plod ovce n-krát pohol (Poissonovo rozdelenie)

n	0	1	2	3	4	5	6	7
m_n	182	41	12	2	2	0	0	1
očakávané m_n	168	60	11	1	0	0	0	0

početnostiam pre $n > 2$, m_n nim zodpovedajúce zahrnieme do jednej skupiny. Potom $\chi^2_{\text{obs}} \doteq 16.57$ a p-hodnota $\doteq 0.0025$ ($df = k - k_p - 1 = 2$, kde $k = 4$ a $k_p = 1$).

```
n <- 0:7
m.n <- c(182,41,12,2,2,0,0,1)
lambda.hat <- sum(n*m.n)/sum(m.n)
ocak.poc <- dpois(n,lambda.hat)*sum(m.n)
m.n.new <- c(m.n[1:3],sum(m.n[4:8]))
ocak.poc.new <- c(ocak.poc[1:3],sum(ocak.poc[4:8]))
chi.obs <- sum((m.n.new-ocak.poc.new)^2/ocak.poc.new) # 16.56523
1-pchisq(chi.obs,df=2) # 0.0002528748
```

Ak máme v pozorovaných početnostiach príliž veľa núl (ide o *overdisersion*, tak ako v príklade 170), vhodným alternatívnym modelom bude **ZIP model** (z angličtiny *zero-inflated Poisson model*; Lambert 1992), kde

$$\Pr(X = x) = pI(x = 0) + (1-p)f(x, \lambda) = \begin{cases} p + (1-p)f(0,\lambda), & x = 0 \\ (1-p)f(n,\lambda), & x = 1, 2, \ldots \end{cases},$$

kde $p \in \langle 0, 1\rangle$ a $x = n$. Potom bude **funkcia vierohodnosti pre ZIP model** definovaná nasledovne

$$L((\lambda, p)^T | \mathbf{x}) = (p + (1-p)f(0,\lambda))^{m_0} \prod_{I(n>0)} (1-p)^{m_n} f(n,\lambda).$$

Logaritmus funkcie vierohodnosti pre ZIP model je rovný

$$l((\lambda,p)^T|\mathbf{x}) = m_0 \log\left(p + (1-p)f(0,\lambda)\right) + \sum_{I(n>0)} m_n \log((1-p)f(x,\lambda)).$$

Príklad 171 (χ^2-test dobrej zhody; fetálna aktivita) *Majme dáta z príkladu 170. Vypočítajte očakávané početnosti za predpokladu, že $X \sim ZIP(\lambda, n)$. Otestujte zhodu pozorovaných a teoretických početností na hladine významnosti $\alpha = 0.05$.*

Riešenie (aj v R)
$\widehat{\lambda} = 0.847$ a $\widehat{p} = 0.577$.

Tabuľka 5.4: Očakávané početnosti m_n päťsekundových intervalov v posledných 2/3 ťarchavosti zaznamenaných ultrazvukom (zaokrúhlené na nula desatinných miest), v ktorých sa plod ovce n-krát pohol (ZIP rozdelenie)

n	0	1	2	3	4	5	6	7
m_n	182	41	12	2	2	0	0	1
očakávané m_n	182.00	36.86	15.61	4.41	0.93	0.16	0.02	0.00

Vzhľadom ku nízkym početnostiam pre $n > 2$, nim zodpovedajúce m_n zahrnieme do jednej skupiny. Potom $\chi^2_{\text{obs}} \doteq 1.35$ a p-hodnota $\doteq 0.245$ ($df = k - k_p - 1 = 1$, kde $k = 4$ a $k_p = 2$).

```
"ll.zip" <- function(theta){
n <- 0:7
m.n <- c(182,41,12,2,2,0,0,1)
m0 <- m.n[1]
if((theta[2]+(1-theta[2]))==1)
{ll <- m0*log(theta[2]+(1-theta[2])*dpois(0,theta[1]))+
        sum(m.n[2:8]*log((1-theta[2])*dpois(n[2:8],theta[1])))}
return(ll)
}
OPTtheta <- optim(c(0.358,0.5),ll.zip,control=list(fnscale=-1))
theta.hat <- OPTtheta$par # 0.8473558 0.5770943
lambda.hat <- theta.hat[1]
p.hat <- theta.hat[2]
ocak.poc <- c(p.hat+(1-p.hat)*dpois(0,lambda.hat),
              (1-p.hat)*dpois(1:7,lambda.hat))*sum(m.n)
round(ocak.poc,2)
#[1] 182.00  36.86  15.61   4.41   0.93   0.16   0.02   0.00
m.n.new <- c(m.n[1:3],sum(m.n[4:8]))
ocak.poc.new <- c(ocak.poc[1:3],sum(ocak.poc[4:8]))
chi.obs <- sum((m.n.new-ocak.poc.new)^2/ocak.poc.new) # 1.353315
1-pchisq(chi.obs,df=1) # 0.2446995
```

V príkladoch 170 a 171 bola použitá $\widehat{\lambda}$ vypočítaná zo všetkých m_n, ale v χ^2 teste dobrej zhody boli použité m_n spojením viacerých $n > 2$. Chernoff a Lehmann (1954) ukázali, že asymptotické rozdelenie χ^2 štatistiky vytvorenej na základe $\widehat{\lambda}$ vypočítanej zo všetkých m_n konverguje ku sume χ^2 štatistiky s $df = k - k_p - 1$ a váženej sume k_p štatistík χ^2 s $df = 1$. Z toho vyplýva, že použitie $\chi^2_{k-k_p-1}$ je *liberálne*, t.j. aktuálna (skutočná) hladina významnosti presiahne nominálnu hladinu významnosti. *Konzervatívnym* prístupom je použitie χ^2_{k-1}. Odhad $\widehat{\lambda}$ vypočítaný použitím zoskupených m_n je robustnejší na odľahlé pozorovania ako odhad $\widehat{\lambda}$ vypočítaný použitím pozorovaných m_n (Cressie a Read, 1984).

Pre dáta z príkladu 170 potom konzervatívny prístup vedie k použitiu χ^2_3, kde $\chi^2_{\text{obs}} \doteq 16.57$ a p-hodnota $\doteq 0.00087$ (porovnaj s liberálnou p-hodnotou $\doteq 0.00025$). Pre dáta z príkladu 171 potom konzervatívny prístup tiež vedie k použitiu χ^2_3, kde $\chi^2_{\text{obs}} \doteq 1.35$ a p-hodnota $\doteq 0.717$ (porovnaj s liberálnou p-hodnotou $\doteq 0.245$).

Správny prístup predstavuje použitie $\widehat{\lambda}$ vypočítanej zo zoskupených m_n a následné použitie χ^2 testu dobrej zhody tiež zo zoskupených m_n. Pre príklad 170 bude mať jadro funkcie vierohodnosti nasledovný tvar

$$L(\lambda|\mathbf{x}) = \left(\frac{\lambda^0 e^{-\lambda}}{0!}\right)^{182} \left(\frac{\lambda^1 e^{-\lambda}}{1!}\right)^{41} \left(\frac{\lambda^2 e^{-\lambda}}{2!}\right)^{12} \left(\sum_{n=3}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!}\right)^5.$$

Potom $\widehat{\lambda} \doteq 0.335$, $\chi^2_{\text{obs}} \doteq 20.237$ a p-hodnota $\doteq 0.00004$ ($df = 2$). Pre príklad 171 bude mať jadro funkcie vierohodnosti nasledovný tvar

$$L((\lambda, p)^T|\mathbf{x}) = (p + (1-p)f(0,\lambda))^{182}\,(1-p)^{41} f(1,\lambda)(1-p)^{12} f(2,\lambda)(1-p)^5 \prod_{n=3}^{\infty} f(n,\lambda).$$

Potom $\widehat{\lambda} \doteq 0.954$, $\widehat{p} \doteq 0.607$, $\chi^2_{\text{obs}} \doteq 2.415$ a p-hodnota $\doteq 0.1202$ ($df = 1$).

5.2 Kolmogorov-Smirnovov test dobrej zhody

Detaily Kolmogorov-Smirnovovho testu dobrej zhody sa odlišujú od χ^2 testu dobrej zhody, avšak princíp výpočtu vertikálnych vzdialeností je zachovaný. Teraz ale používame všetkých n pozorovaní a počítame vertikálny rozdiel medzi (kumulatívnou) distribučnou funkciou $F_0(x)$, kde máme všetky parametre rozdelenia špecifikované a empirickou (kumulatívnou) distribučnou funkciou $\widehat{F}_n(x)$ pre všetky x.

Za platnosti H_0 bude vzdialenosť medzi $F_0(x)$ a $\widehat{F}_n(x)$ malá pre všetky x. **Jednovýberová Kolmogorov-Smirnovova testovacia štatistika** je definovaná nasledovne

$$D_n = \sup_{\forall x} |\widehat{F}_n(x) - F_0(x)|.$$

D_n nezávisí na $F_0(x)$, ak je $F(x)$ spojitá. Test sa nazýva **jednovýberový Kolmogorov-Smirnovov test dobrej zhody**. Je potrebné zdôrazniť, že D_n môžeme použiť, len ak je hypotéza jednoduchá. Empirická distribučná funkcia je definovaná nasledovne

$$\widehat{F}_n(x) = \sum_{i=1}^{n} I(x_i \leq x)/n$$

alebo alternatívne ako

$$\widehat{F}_n(x) = \begin{cases} 0, & \text{ak } x < x_{(1)}, \\ \frac{i}{n}, & \text{ak } x_{(i)} \leq x < x_{(i+1)}, \\ 1, & \text{ak } x \geq x_{(n)}. \end{cases}$$

Ak sa nevyskytujú v pozorovaniach zhody,

$$D_n = \max_{\forall i} M_i,$$

kde

$$M_i = \max\left\{ |\widehat{F}_n(x_i) - F_0(x_i)|, |F_0(x_i) - \widehat{F}_n(x_{i-1})| \right\}.$$

Keďže $\widehat{F}_n(x_i) = \frac{i}{n}$ a $\widehat{F}_n(x_{i-1}) = \frac{i-1}{n}$, môžeme písať

$$M_i = \max\left\{ D_i^+ = |\frac{i}{n} - F_0(x_i)|, D_i^- = |F_0(x_i) - \frac{i-1}{n}| \right\}.$$

Nulovú a alternatívnu hypotézu definujeme nasledovne

$$H_0 : F(x) = F_0(x) \text{ pre všetky } x \text{ oproti } H_1 : F(x) \neq F_0(x) \text{ pre nejaké } x.$$

H_0 zamietame, keď $D_n > D_n(\alpha)$, kde $D_n(\alpha)$ je kritická hodnota.

Príklad 172 (Kolmogorov-Smirnovov test dobrej zhody) *Majme výšky $n = 12$ náhodne vybraných 10-ročných dievčat $\mathbf{x} = (131, 132, 135, 141, 141, 141, 141, 142, 143, 146, 146, 151)^T$. Otestujte na hladine významnosti $\alpha = 0.05$, či majú dáta normálne rozdelenie, kde $F_0(x) \sim N(\mu, \sigma^2)$.*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: Výška 10-ročných dievčat má normálne rozdelenie (jej distribučná funkcia pochádza z normálneho rozdelenia) oproti H_1: výška 10-ročných dievčat nemá normálne rozdelenie (jej distribučná funkcia nepochádza z normálneho rozdelenia).
   - **matematická formulácia** – testujeme $H_0 : F(x) = F_0(x) \sim N(\mu, \sigma^2)$ oproti $H_1 : F(x) \neq F_0(x)$ pre nejaké x.

2. **Testovacia štatistika** – Stredná hodnota μ a rozptyl σ^2 sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetický priemer $\overline{x} = 140.83$ a rozptyl $s^2 = 33.79$ (smerodajná odchýlka $s = 5.81$)

```
vyska <- c(131,132,135,141,141,141,141,142,143,146,146,151)
priemer <- mean(vyska) # 140.8333
smerodch <- sd(vyska) # 5.812734
```

   Najprv vypočítame očakávané hodnoty $F_0(x_{(i)})$, potom zoradíme výšky podľa veľkosti, vypočítame pozorované hodnoty $\widehat{F}_n(x_{(i)}) = \frac{i}{n}$, $\widehat{F}_n(x_{(i-1)}) = \frac{i-1}{n}$ ako aj D_i^+, D_i^-, M_i a nakoniec D_n. Keďže $F_0(x) \sim N(140.83, 5.81)$, potom $F_0(x_{(i)}) = \Pr(X \leq x_{(i)}) = \Pr(\frac{Y-140.83}{5.81} \leq \frac{x_{(i)}-140.83}{5.81}) = \Pr(Z \leq \frac{x_{(i)}-140.83}{5.81})$.

```
vyska <- sort(vyska)
n <- length(vyska)
FoX <- pnorm(vyska,mean=priemer,sd=smerodch) # teoreticke pravdepodobnosti
FnX <- seq(1:n)/n # pozorovane pravdepodobnosti
Fn1X <- (seq(1:n)-1)/n
D.plus <- abs(FnX-FoX)
D.minus <- abs(FoX-Fn1X)
VYSL <- cbind(x,FnX,Fn1X,FoX,D.plus,D.minus)
Mi <-apply(VYSL[,c(5,6)],1,max)
VYSL <- cbind(VYSL,Mi)
```

   Pozn.: Funkciu ecdf() (príkaz v podobe Fn ¡- ecfd(vyska); FnX ¡- Fn(vyska)) nie je možné použiť, pretože pri zhodách je posunutá $\widehat{F}_n(x_{i-1})$ vypočítaná z $\widehat{F}_n(x_i)$ nesprávna.

   Pozorovaná testovacia štatistika je potom rovná $D_n \doteq 0.26$.

```
Dn <- max(Mi) # 0.2614372
```

3. **Zamietacia oblasť** – H_0 bude zamietnutá, ak $D_n > D_n(\alpha)$, kde $D_n(\alpha)$ je kritická hodnota.

4. **Štatistický záver** – p-hodnota $\doteq 0.385$.

```
p.hodn <- ks.test(vyska,y="pnorm",mean=priemer,sd=smerodch)$p.val # 0.384988
```

   Kedže p-hodnota nie je menšia ako hladina významnosti $\alpha = 0.05$ (ekvivalentne D_n nepatrí zamietacej oblasti), H_0 nezamietame na hladine významnosti α.

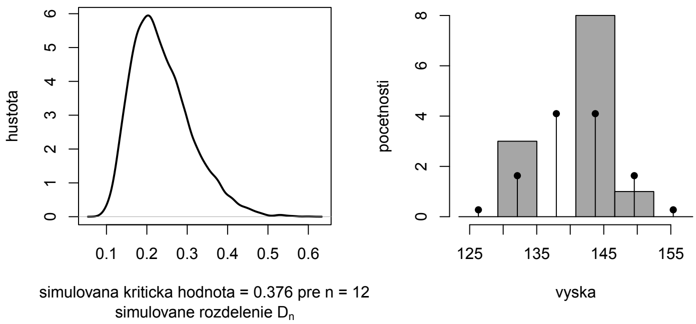
   Pozn.: Ak by sme vykonali MC simuláciu kritických hodnôt štatistiky D_n pre $n = 12$ pre dáta z normálneho rozdelenia ($M = 10000$), dalo by sa ukázať, že pri $n = 12$ H_0 nezamietame (pozri obrázok 5.2 vľavo).

```
"ks.mc" <- function (n=10,M=10000,alpha=0.05) {
    Dn <- replicate(M,ks.test(rnorm(n),pnorm)$statistic)
    kh <- quantile(Dn,1-alpha)
    plot(density(Dn),col="black",lwd=2,ylab="hustota",main="",
        xlab = paste("simulovana kriticka hodnota =",round(kh,3),"pre n =", n))
    title(sub=list(expression(paste("simulovane rozdelenie ",D[n]))))
    return(kh)
}
par(mfcol=c(1,2))
ks.mc(n=12,M=10000,alpha=0.05) # 0.376
```

5. **Slovný záver** – Nemáme dostatok dôkazov na zamietnutie nulovej hypotézy o tom, že skutočnou distribučnou funkciou výšky 10-ročných dievčat je distribučná funkcia normálneho rozdelenia.

Histogram superponovaný s očakávanými početnosťami výšok nakreslíme nasledovne (pozri obrázok 5.2 vpravo)

```
bin <- seq(priemer-3*smerodch, priemer+3*smerodch, smerodch) # kategorie
hist(vyska, breaks=bin, col="gray", ylab="pocetnosti", freq=TRUE, main="")
x <- bin[2:7]-smerodch/2 # centrovanien do stredu intervalov
ocak.prav <- c(pnorm(-2), pnorm(-1:2)-pnorm(-2:1), 1-pnorm(2))
ocak <- 12*ocak.prav # ocakavane pocetnosti
lines(x, ocak, type="h")
points(x, ocak, pch=16)
```



Obr. 5.2: Simulované rozdelenie D_n (vľavo) a histogram superponovaný s očakávanými početnosťami výšok (vpravo)

Ak je hypotéza zložená, Kolmogorov-Smirnovov test je veľmi konzervatívny. Avšak D_n môžeme použiť na výpočet, len ak odhadneme parametre príslušného rozdelenia, kde $\widehat{F}_0(x)$ substituujeme za $F_0(x)$. Potom však nastávajú problémy s rozdelením D_n. Problém rieši modifikácia Kolmogorovho-Smirnovovho testu, kedy sa tento test nazýva **Lillieforsov test dobrej zhody**, použitím MC simulácií, kde kritické hodnoty označíme $D_n^{(l)}(\alpha)$. Simulované hodnoty D_n pozri na obrázku 5.3.

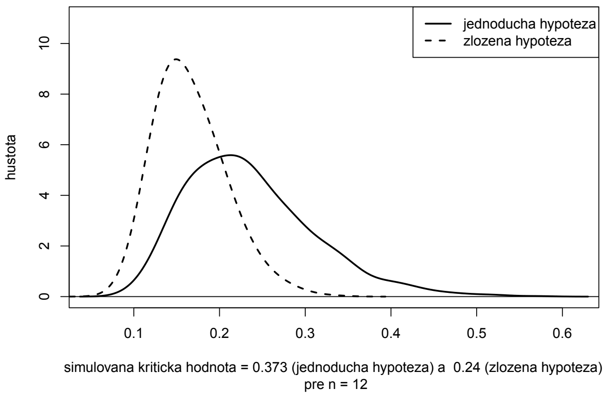
```
"ks.l.mc" <- function (n=10,M=1000,alpha=0.05)
{
    Dn <- c()
    DnL <- c()
    for (i in 1:M) {
        x <- rnorm(n)
        mu <- mean(x)
        sig <- sd(x)
        Dn[i] <- ks.test(x,pnorm)$statistic
        DnL[i] <- ks.test(x,pnorm,mean=mu,sd=sig)$statistic
    }
    ys <- range(density(DnL)$y)
    xs <- range(density(Dn)$x)
    cv <- quantile(Dn,1-alpha)
    cvL <- quantile(DnL,1-alpha)
    plot(density(Dn,bw=0.02),col="black",lwd=2,ylim=ys,
        xlim=xs,main="",xlab="",ylab="hustota",
        sub=paste("simulovana kriticka hodnota =",round(cv,3),
        "(jednoducha hypoteza) a ", round(cvL,3),
        "(zlozena hypoteza)\n pre n =",n))
    lines(density(DnL,bw=0.02),col="black",lwd=2,lty=2)
    legend("topright",legend=c("jednoducha hypoteza",
        "zlozena hypoteza"),lty=c(1,2),lwd=2)
    #box()
    abline(h=0)
}
ks.l.mc(n=12,M=1000)
```

Z histogramu je zreteľné, že rozdiely medzi očakávanými a odhadnutými početnosťami sú značné z dôvodu štruktúry dát a ich rozsahu, čo je v rozpore so záverom Kolmogorov-Smirnovovho testu, ktorý H_0 nezamietol. Avšak použitím Lillieforsovho testu normality H_0 zamietame, kde $D_n > D_n^{(l)}(\alpha)$.

Pri testovaní sa často používa **Dallal-Wilkinsonova aproximácia p-hodnoty** v podobe

$$\widehat{\text{p-hodnota}} = \exp(-7.01256 D_n^2 (n+2.78019) + 2.99587 D_n \sqrt{n+2.78019} - 0.122119 + \frac{0.974598}{\sqrt{n}} + \frac{1.67997}{n})$$

pre $n \in (5, 100)$ a p-hodnotu ≤ 0.1. Ak je $n \geq 100$, potom D_n vo vyššie uvedenom vzorci nahradíme $D_m = D_n(\frac{m}{100})^{0.49}$, kde m je skutočný rozsah a n substituujeme číslom 100. Ak p-hodnota > 0.1,



Obr. 5.3: Simulované hustoty rozdelenia D_n pre jednoduchú a zloženú hypotézu

potom D_n nahradíme $D_{\mathrm{mod}} = D_n(\sqrt{n} - 0.01 + 0.85\sqrt{n})$. Podľa veľkosti D_{mod} vypočítame p-hodnotu nasledovne (Dallal a Wilkinson, 1986; Stephens, 1974):

- ak $D_{\mathrm{mod}} \leq 0.302$, $\widehat{\text{p-hodnota}} = 1$,
- ak $D_{\mathrm{mod}} \leq 0.5$,

  $$\widehat{\text{p-hodnota}} = 2.76773 - 19.828315 D_{\mathrm{mod}} + 80.709644 D_{\mathrm{mod}}^2 - 138.55152 D_{\mathrm{mod}}^3 + 81.218052 D_{\mathrm{mod}}^4,$$

- ak $D_{\mathrm{mod}} \leq 0.9$,

  $$\widehat{\text{p-hodnota}} = -4.901232 + 40.662806 D_{\mathrm{mod}} - 97.490286 D_{\mathrm{mod}}^2 + 94.029866 D_{\mathrm{mod}}^3 - 32.355711 D_{\mathrm{mod}}^4,$$

- ak $D_{\mathrm{mod}} \leq 1.31$,

  $$\widehat{\text{p-hodnota}} = 6.198765 - 19.558097 D_{\mathrm{mod}} + 23.186922 D_{\mathrm{mod}}^2 - 12.234627 D_{\mathrm{mod}}^3 + 2.423045 D_{\mathrm{mod}}^4.$$

Na Lillieforsov test normality môžeme použiť aj funkciu

```
library(nortest)
lillie.test(vyska)$stat # 0.2614372
lillie.test(vyska)$p.val # 0.02296345
```

Nulovú hypotézu o tom, že skutočnou distribučnou funkciou výšky 10-ročných dievčat je distribučná funkcia normálneho rozdelenia, zamietame.

Výhody Kolmogorovho-Smirnovovho testu. Kolmogorov-Smirnovov test má dve hlavné výhody oproti χ^2 testu

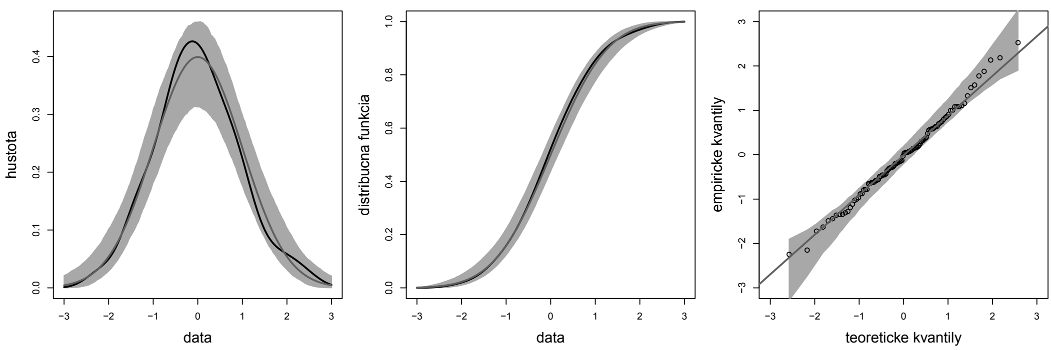
1. je použiteľný aj pre malé rozsahy výberov, kedy je validita χ^2 testu dobrej zhody otázna,
2. jeho sila je väčšia ako sila χ^2 testu dobrej zhody pre akýkoľvek rozsah výberu.

Ak však ide o zloženú hypotézu, kde je potrebné odhadnúť aj parametre rozdelenia, Kolmogorov-Smirnov test v jeho klasickej podobe nie je možné použiť. Výsledky by v tomto prípade boli konzervatívne v zmysle pravdepodobnosti chyby prvého druhu, ktorá by bola v skutočnosti menšia, t.j. pozorovaná hladina významnosti by bola tiež menšia, čo by znamenalo „pomalšie“ zamietanie za použitia $D_n(\alpha)$ namiesto $D_n^{(l)}(\alpha)$. Ak odhadujeme strednú hodnotu a rozptyl, použijeme Lillieforsov test normality. Približný vzťah medzi kritickými hodnotami $D_n^{(l)}(\alpha)$ tohoto testu a kritickými hodnotami $D_n(\alpha)$

môžeme vyjadriť ako $D_n^{(l)}(\alpha) \approx \frac{2}{3} D_n(\alpha)$. Pre veľké hodnoty n hodnoty $D_n^{(l)}(\alpha)$ klesajú tak, ako klesá $\frac{1}{\sqrt{n}}$.

Príklad 173 (Lillieforsov test normality) *Pomocou Lillieforsovho testu otestujte normalitu premenných:*
(a) výška postavy žien (body.H; v mm; dáta: anova-head.txt),
(b) stranový rozdiel vertikálneho priemeru diafýzy kľúčnej kosti (simd.R a simd.L; v mm) na pravej a ľavej strane tela (dáta: paired-means-clavicle2.txt),
(c) najväčšia výška mozgovne u mužov (skull.pH; v mm; dáta: one-sample-correlation-skull-mf.txt) a
(d) morfologická výška tváre u žien (face.H; v mm; dáta: one-sample-correlation-skull-mf.txt).

Príklad 174 (pásy normality) *Na základe vygenerovaných pseudonáhodných čísel $X \sim N(\mu, \sigma^2)$, $n = 100, \mu = 0, \sigma^2 = 1$, kde $M = 1000$, odhadnite (a) hustotu m-tej realizácie pomocou funkcie density(); ponechajte argument n=512 a nastavte from=-3 a to=3; (b) distribučnú funkciu m-tej realizácie pomocou (a) a funkcie cumsum() a (c) empirické kvantily m-tej realizácie pomocou funkcie qqnorm(). Vygenerované čísla $x_{mi}, m = 1, 2, \ldots, 1000$ a $i = 1, 2, \ldots, 100$, uložte po riadkoch do matice $\mathbf{X}$, ktoré budú mať rozmery 1000×100. Odhadnuté hustoty a distribučné funkcie uložte po riadkoch do matíc $\mathbf{H}$ a $\mathbf{D}$, ktoré bude mať rozmery 1000×512 a empirické kvantily do matice $\mathbf{K}$, ktorá bude mať rozmery 1000×100. Pre každú z matíc $\mathbf{H}$, $\mathbf{D}$ a $\mathbf{K}$ vypočítajte $\widetilde{x}_{0.05}$ a $\widetilde{x}_{0.95}$ po stĺpcoch a zobrazte ich ako pásy pomocou funkcie polygon(). Do obrázkov vkreslite (a) teoretickú hustotu, (b) teoretickú distribučnú funkciu a (c) kvantilovú priamku (pomocou funkcie qqline()) červenou farbou. Obrázky usporiadajte ako trojicu vedľa seba. Dáta, ktorých normalitu chceme graficky testovať, budú (1) $X \sim N(0, 1)$, $n = 100$, (2) $X \sim [pN(0, 1) + (1-p)N(0, 4)]$, $n = 100$ a $p = 0.95$ a (3) $X \sim [pN(0, 1) + (1-p)N(0, 4)]$, $n = 100$ a $p = 0.9$. Zobrazte (1), (2) a (3) oddelene do grafov (a), (b) a (c). Okomentujte. Riešenie (1) pozri na obrázku 5.4.*



Obr. 5.4: Pásy spoľahlivosti normálneho rozdelenia – pre hustotu (vľavo), distribučnú funkciu (uprostred) a kvantilovú priamku (vpravo)

6 Testovanie hypotéz o jednom parametri

V tejto kapitole sa budeme venovať testovaniu hypotéz o strednej hodnote μ a rozptyle σ^2 za predpokladu normality X, t.j. $X \sim N(\mu, \sigma^2)$; o korelačnom koeficiente ρ premenných X a Y za predpokladu dvojrozmernej normality $(X, Y)^T$, t.j. $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a nakoniec o pravdepodobnosti p za predpokladu $X \sim Bin(N, p)$ a $X \sim Poiss(\lambda)$.

6.1 Asymptotické testy o strednej hodnote

Jednovýberový test hypotézy o strednej hodnote normálneho rozdelenia sa v biologickej antropológii používa najčastejšie v situáciách, keď máme k dispozícii namerané hodnoty nejakej spojitej premennej (napríklad výšky postavy) a chceme zistiť, či sa stredná hodnota výšky postavy našej populácie líši alebo nelíši od referenčnej (napr. reprezentatívneho, celoštátneho súboru) či publikovanej (napr. z iného mesta, iného obdobia alebo inej generácie) strednej hodnoty výšky postavy vzorky, z ktorej nemáme k dispozícii primárne dáta (individuálne hodnoty všetkých prípadov), ale iba sekundárne dáta ako štatistické charakteristiky/parametre (aritmetický priemer, počet jedincov a smerodajnú odchýlku). Keďže číselne sa stredné hodnoty (pri dostatočnej presnosti, počte desatinných miest) akýchkoľvek dvoch populácií vždy líšia (t.j. jedna je vždy menšia a druhá väčšia), je potrebné zistiť, či je rozdiel štatisticky významný (signifikantný; v minulosti sa používalo označenie štatisticky závažný). Potrebujeme rozhodnúť, či má rozdiel takú váhu, že je opodstatnené hľadať vysvetlenie príčin tohto rozdielu (interpretáciu výsledkov) alebo je rozdiel iba malý, daný náhodnými vplyvmi a nemá zmysel sa ním zaoberať (resp. má zmysel zaoberať sa jeho neexistenciou). Nie je preto správne porovnávať strednú hodnotu s referenčnou hodnotou iba číselne, príp. graficky (i keď sa to, žiaľ, v prehľadových štúdiách a diskusiách antropologických štúdií často deje), ale je nevyhnutné rozdiel vždy štatisticky testovať. Až potom sa môžeme pokúsiť o vysvetlenie rozdielu (v prípade zamietnutia nulovej hypotézy o zhode strednej hodnoty s referenčnou strednou hodnotou) alebo, naopak, o vysvetlenie zhody či veľkej podobnosti oboch vzoriek (v prípade nezamietnutia nulovej hypotézy). Na to však spravidla nestačia hodnoty rozdielu alebo výsledky samotného testu, potrebujeme aj ďalšie informácie o populácii, vzorke a spôsobe jej vzniku. Príčinou rozdielu v sledovanom znaku môžu byť genetické (veľkosť populácie, špecifické evolučné adaptácie a i.), environmentálne (fyzikálne a geografické faktory prostredia, záťaž, strava, a i.) či socio-kultúrne (behaviorálne, kultúrne, sociálne a ekonomické) rozdiely medzi geografickými oblasťami, populáciami, generáciami, sociálnymi vrstvami atď. Štandardne sa vo väčšine znakov vo všetkých ľudských populáciách vyskytujú rozdiely vekové (ontogenetické rastové a vývojové zmeny v priebehu dospievania a involučné funkčné a morfologické zmeny v priebehu starnutia) a rozdiely pohlavné (sexuálny dimorfizmus), ktoré je nevyhnutné vždy zohľadňovať, t.j. nie je možné porovnávať vzorky z tej istej alebo odlišných populácii, u ktorých nepoznáme vekové a pohlavné zloženie (počet mužov a žien a ich vek).

Pomocou jednovýberového testu strednej hodnoty možno testovať rozdiely spojitých premenných (rozmery živého človeka, rozmery skeletu), ktoré majú známe rozdelenie. Nevýhodou využitia jednovýberového testu a porovnávania s referenčnou či publikovanou hodnotou je ale skutočnosť, že testujeme rozdiely medzi dátami zmeranými inými pozorovateľmi, inými meradlami a stanovené väčšinou nedostatočne spresneným a nekontrolovaným spôsobom. Presnosť meradiel, spôsob merania a odlišné okolnosti pri zbere dát v oboch vzorkách môžu byť príčinou časti rozdielov v stredných hodnotách. Preto je vhodnejšie použiť dvojvýberové testy, kde obe porovnávané vzorky merala jedna osoba rovnakým meradlom a najlepšie vo vhodnom znáhodnenom poradí z hľadiska príslušnosti meraných jednotiek k porovnávaným skupinám/vzorkám.

Nech $X \sim N(\mu, \sigma^2)$, kde σ^2 je neznáma. Majme dvojice hypotéz $H_{01} : \mu = \mu_0$ oproti $H_{11} : \mu \neq \mu_0$, $H_{02} : \mu \leq \mu_0$ oproti $H_{12} : \mu > \mu_0$, $H_{03} : \mu \geq \mu_0$ oproti $H_{13} : \mu < \mu_0$, ktoré chceme testovať.

Ak H_0 platí, potom

$$T_W = \frac{\overline{X} - \mu_0}{S}\sqrt{n} \overset{\mathcal{D}}{\sim} t_{n-1},$$

kde $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ a rozdelenie t_{df}, kde $df = n-1$, sa nazýva **centrálne t-rozdelenie** s $n-1$ stupňami voľnosti. T_W sa nazýva **Waldova testovacia štatistika** (alebo **t-štatistika**) a test **jednovýberový Studentov t-test o strednej hodnote μ.**

Ak H_0 neplatí, táto situácia vedie k **necentrálnemu t-rozdeleniu** s df stupňami voľnosti a parametrom necentrality λ, ozn. $t_{df,\lambda}$, kde

$$T_{W,\lambda} = \frac{Z_W + \lambda}{\sqrt{V/df}} = \frac{\sqrt{n}\left(\overline{X} - \mu\right)/\sigma + \sqrt{n}\left(\mu - \mu_0\right)/\sigma}{S/\sigma} \overset{\mathcal{D}}{\sim} t_{df,\lambda},$$

kde $df = n-1$, $Z_W = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \overset{\mathcal{D}}{\sim} N(0,1)$, parameter necentrality $\lambda = (\delta/\sigma)\sqrt{n}$, $\delta = \mu - \mu_0$ je **minimálne detegovateľná vzdialenosť** medzi μ a μ_0, $V = \frac{(n-1)S^2}{\sigma^2} \overset{\mathcal{D}}{\sim} \chi^2_{df}$ a je nezávislé na Z_W. Označme kumulatívnu distribučnú funkciu $T_{W,\lambda}$ ako $G_{df,\lambda}(t) = \Pr\left(T_{W,\lambda} \le t\right)$. Ak $\lambda = 0$, potom sa necentrálne t-rozdelenie redukuje na centrálne (Studentove) t-rozdelenie.

Ak H_{02} neplatí, potom silofunkcia

$$1 - \beta(\mu,\sigma) = \Pr_{\mu,\sigma}\left(\frac{\overline{X} - \mu_0}{S}\sqrt{n} \ge t_{n-1}(\alpha)\right) = 1 - G_{n-1,\lambda}\left(t_{n-1}(\alpha)\right),$$

kde argumenty (μ,σ) indikujú, že pravdepodobnosť je počítaná za predpokladu $X \sim N(\mu,\sigma^2), \mu > \mu_0$.
Ak H_{03} neplatí, potom silofunkcia

$$1 - \beta(\mu,\sigma) = \Pr_{\mu,\sigma}\left(\frac{\overline{X} - \mu_0}{S}\sqrt{n} \le -t_{n-1}(\alpha)\right) = G_{n-1,\lambda}\left(t_{n-1}(\alpha)\right),$$

kde argumenty (μ,σ) indikujú, že pravdepodobnosť je počítaná za predpokladu $X \sim N(\mu,\sigma^2), \mu < \mu_0$.
Ak H_{01} neplatí, potom silofunkcia

$$\begin{aligned} 1 - \beta(\mu,\sigma) &= \Pr_{\mu,\sigma}\left(\frac{\overline{X} - \mu_0}{S}\sqrt{n} \le -t_{n-1}(\alpha/2) \vee \frac{\overline{X} - \mu_0}{S}\sqrt{n} \ge t_{n-1}(\alpha/2)\right) \\ &= 1 - G_{n-1,\lambda}\left(t_{n-1}(\alpha/2)\right) + G_{n-1,\lambda}\left(-t_{n-1}(\alpha/2)\right), \end{aligned}$$

kde argumenty (μ,σ) indikujú, že pravdepodobnosť je počítaná za predpokladu $X \sim N(\mu,\sigma^2), \mu \ne \mu_0$.

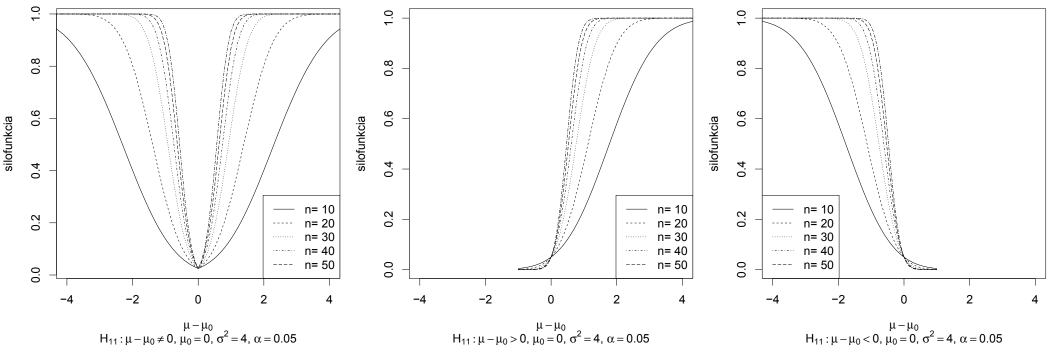
Definícia 52 (Kritický obor a silofunkcia T_W testu o μ) *Kritický obor a silofunkcia sú definované (za platnosti $t^2_{n-1}(\alpha/2) \approx F_{1,n-1}(\alpha)$ a $t^2_{n-1,\lambda}(\alpha/2) \approx F_{1,n-1,\lambda^2}(\alpha)$) nasledovne (pozri obrázok 6.1):*

H_0	H_1	$\mathcal{W}$	$1-\beta(\mu)$
$\mu = \mu_0$	$\mu \ne \mu_0$	$\mathcal{W}_1 = \{T_W; \lvert T_W\rvert \ge t_{n-1}(\alpha/2)\}$	$\Pr\left(F_{1,n-1,\lambda^2} \ge F_{1,n-1}(\alpha)\right)$
$\mu \le \mu_0$	$\mu > \mu_0$	$\mathcal{W}_2 = \{T_W; T_W \ge t_{n-1}(\alpha)\}$	$\Pr\left(t_{n-1,\lambda} \ge t_{n-1}(\alpha)\right)$
$\mu \ge \mu_0$	$\mu < \mu_0$	$\mathcal{W}_3 = \{T_W; T_W \le -t_{n-1}(\alpha)\}$	$\Pr\left(t_{n-1,\lambda} \le -t_{n-1}(\alpha)\right)$

Definícia 53 (p-hodnota T_W testu o μ) *Nech T_W je* ***nejaká testovacia štatistika*** *a $t_W = \frac{\overline{x}-\mu_0}{s}\sqrt{n}$ je jej realizácia (****pozorovaná, vypočítaná testovacia štatistika****), potom p-hodnotu počítame nasledovne:*

$$p\text{-hodnota} = \begin{cases} 2Pr(T_W \ge |t_W| | H_{01}), \ ak & H_{11}: \mu \ne \mu_0 \\ Pr(T_W \ge t_W | H_{02}), \ ak & H_{12}: \mu > \mu_0 \\ Pr(T_W \le t_W | H_{03}), \ ak & H_{13}: \mu < \mu_0 \end{cases}$$

Definícia 54 (Waldove $100 \times (1-\alpha)\%$ empirické IS pre μ) *Waldove $100\times(1-\alpha)\%$ empirické IS pre μ pre všetky tri typy hypotéz majú nasledovný tvar:*



Obr. 6.1: Silofunkcie asymptotického testu o μ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

H_0	H_1	*hranice* (d, h) *pre* $100(1-\alpha)\%$ *empirický IS*
$\mu = \mu_0$	$\mu \neq \mu_0$	$\mathcal{CS}_{1-\alpha} = \left\{\mu_0 : \mu_0 \in \left(\overline{x} - t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}, \overline{x} + t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}\right)\right\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\mathcal{CS}_{1-\alpha} = \left\{\mu_0 : \mu_0 \in \left(\overline{x} - t_{n-1}(\alpha)\frac{s}{\sqrt{n}}, \infty\right)\right\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\mathcal{CS}_{1-\alpha} = \left\{\mu_0 : \mu_0 \in \left(-\infty, \overline{x} + t_{n-1}(\alpha)\frac{s}{\sqrt{n}}\right)\right\}$

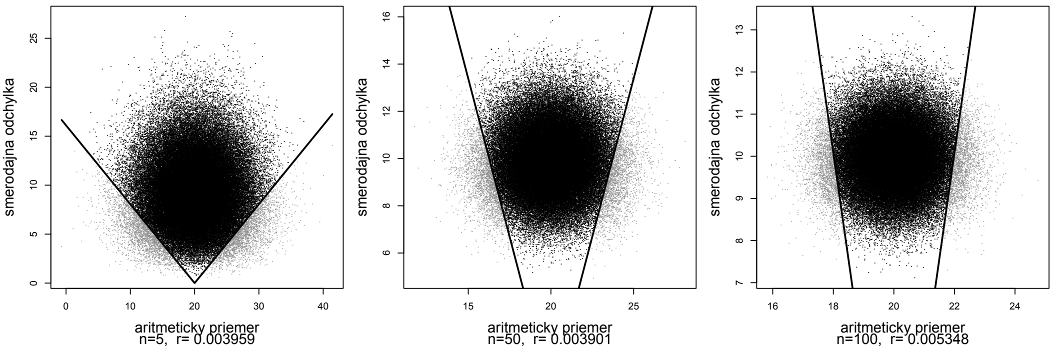
Príklad 175 (nezávislosť μ a σ^2; pravdepodobnosť pokrytia) *Nech $X \sim N(\mu, \sigma^2)$, kde $\mu = 20$ a $\sigma^2 = 100$. Vypočítajte Pearsonov korelačný koeficient $r_{\overline{X},S}$ pomocou simulačnej štúdie. Nakreslite sivou farbou rozptylový graf $(\overline{x}_m, s_m)$, kde $m = 1, 2, \ldots, M$, kde $M = 100000$. Dokreslite do grafu čiernou farbou také body, pre ktoré platí $t_{W,m} = \left|\frac{\overline{x}_m - \mu}{s_m}\sqrt{n}\right| < t_{n-1}(\alpha/2)$, ako aj hranice, ktoré definujú také body $(\overline{x}_m, s_m)$, pre ktoré $t_{W,m} = t_{n-1}(\alpha/2)$. Vypočítajte pravdepodobnosť pokrytia 95% DIS pre μ ako podiel $\sum_m I(t_{W,m} < t_{n-1}(\alpha/2))/M$. Zvoľte (a) $n = 5$, (b) $n = 50$ a (c) $n = 100$.*

Riešenie v R (pozri obrázok 6.2)

```
M <- 100000
n <- 5
mu <- 20
sigma <- 10
x <- rnorm(M*n,mu,sigma)
DATA <- matrix(x,nrow=M)
mu.hat <- rowMeans(DATA)
sd.hat <- sqrt(rowSums((DATA-mu.hat)^2)/(n-1))
cor(mu.hat,sd.hat) # 0.002212632

windows(12,4)
par(mfcol=c(1,3),mar=c(5,4.2,1,1))
plot(mu.hat, sd.hat,xlab="aritmeticky priemer",ylab="smerodajna odchylka",
     col="darkgrey",pch=16,cex=0.2,cex.lab=1.5,
     sub=paste("n=5, ","r=",round(cor(mu.hat,sd.hat),6)),cex.sub=1.5)
## pravdepodobnost pokrytia
t.krit <- qt(0.975,n-1)
tW.obs <- sqrt(n)*(mu.hat-mu)/sd.hat
pokrytie <- (1:M)[abs(tW.obs)<t.krit]
points(mu.hat[pokrytie],sd.hat[pokrytie],pch=16,cex=0.2)
mu.hat.i <- seq(min(mu.hat),max(mu.hat),length=1000)
sd.hat.i <- abs(sqrt(n)*(mu.hat.i-mu)/t.krit)
lines(mu.hat.i,sd.hat.i,lwd=2)
p.pokrytie <- sum(abs(tW.obs)<t.krit)/M # 0.94856
```

Ak sú $\overline{X}$ a S nezávislé, potom $\rho_{\overline{X},S} = 0$ a $r_{\overline{X},S} \approx 0$. Na základe výsledkov simulačnej štúdie môžeme konštatovať, že $r_{\overline{X},S} \approx 0$ pre $n = 5, 50$ a 100 (pozri obrázok 47). Pravdepodobnosť pokrytia je pre $n = 5$ rovná 0.94856, pre $n = 50$ je rovná 0.95021 a pre pre $n = 50$ je rovná 0.95015. Z toho vyplýva, že aktuálne hodnoty pravdepodobnosti pokrytia (získané zo simulačnej štúdie) sú pre $X \sim N(\mu, \sigma^2)$ dostatočne blízke nominálnej hodnote $1 - \alpha = 0.95$.

Príklad 176 (nezávislosť μ a σ^2; pravdepodobnosť pokrytia) *Nech $X \sim [pN(\mu, \sigma_1^2) + (1-p)N(\mu, \sigma_2^2)]$, kde $p = 0.9$, $\mu = 20$, $\sigma_1^2 = 100$ a $\sigma_2^2 = 400$. Vypočítajte Pearsonov korelačný koeficient $r_{\overline{X},S}$*



Obr. 6.2: Rozptylový graf $\overline{x}_i, s_i$, $i = 1, 2, \ldots, M, M = 100000$ pre $n = 5$ (vľavo), $n = 50$ (v strede) a $n = 100$ (vpravo)

pomocou simulačnej štúdie. Nakreslite sivou farbou rozptylový graf $(\overline{x}_m, s_m)$, *kde* $m = 1, 2, \ldots, M$, *kde* $M = 100000$. *Dokreslite do grafu čiernou farbou také body, pre ktoré platí* $t_{W,m} = \left|\frac{\overline{x}_m - \mu}{s_m}\sqrt{n}\right| < t_{n-1}(\alpha/2)$, *ako aj hranice, ktoré definujú také body* $(\overline{x}_m, s_m)$, *pre ktoré* $t_{W,m} = t_{n-1}(\alpha/2)$. *Vypočítajte pravdepodobnosť pokrytia 95% DIS pre* μ *ako podiel* $\sum_m I\left(t_{W,m} < t_{n-1}(\alpha/2)\right)/M$. *Zvoľte (a)* $n = 5$, *(b)* $n = 50$ *a (c)* $n = 100$.

Jednovýberový Studentov t-test o strednej hodnote μ v R (funkcia **t.test()**).
Argumenty (vstupy) funkcie:

1. vektor dát x;
2. alternatíva alternative="two.sided" je prednastavená, ďalšie voľby sú "greater", "less";
3. stredná hodnota za platnosti nulovej hypotézy mu, t.j. μ_0;
4. spoľahlivosť conf.level=p, prednastavené je p =.95.

Výstupy funkcie:

1. názov použitého testu method;
2. testovacia štatistika statistic;
3. počet stupňov voľnosti df parameter;
4. p-hodnota p.value;
5. alternatívna hypotéza alternative hypothesis;
6. interval spoľahlivosti pre μ conf.int;
7. bodový odhad (aritmetický priemer $\overline{x}$) sample estimates.

Príklad 177 (necentrálne t-rozdelenie) *Nakreslite distribučnú funkciu necentrálneho t-rozdelenia* $t_{n-1,\lambda}$, *kde* $\delta = \mu - \mu_0$ *a* $\lambda = \delta/(\sigma/\sqrt{n})$. *Použite* $\mu_0 = 0$, $\delta = 1$, $\sigma = 1.4$ *a* $n = 26$. *Vypočítajte pravdepodobnosť nad kvantilom* $x_{0.975}$ *pod krivkou hustoty tohoto rozdelenia.*

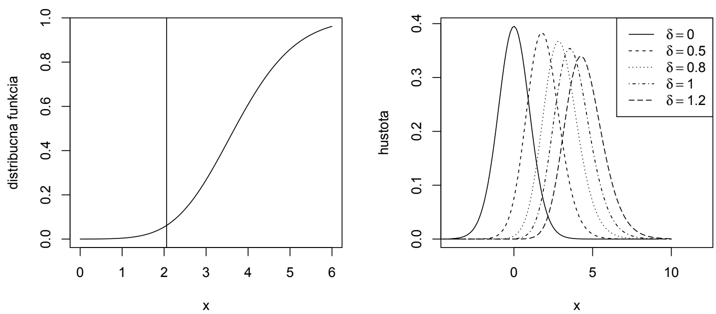
Riešenie v R (pozri obrázok 6.3 vľavo)

```
delta <- 1
sigma <- 1.4
n <- 26
lambda <- delta/sigma*sqrt(n)
curve(pt(x,25,ncp=lambda),from=0,to=6,xlab="x",ylab="distribucna funkcia")
abline(v=qt(.975,25)) # 2.06
1-pt(qt(.975,25),25,ncp=lambda) # 0.9381038
```

Príklad 178 (necentrálne t-rozdelenie) *Nakreslite hustoty jedného centrálneho a štyroch necentrálnych t-rozdelení* $t_{n-1,\lambda}$ *(*$\delta = \mu - \mu_0$ *a* $\lambda = \delta/(\sigma/\sqrt{n})$*) do jedneho obrázka tak, aby boli odlíšiteľné farbou alebo typom čiary. Použite* $\mu_0 = 0$, $\delta = 0, 0.5, 0.8, 1$ *a* 1.2, $\sigma = 1.4$ *a* $n = 26$.

Riešenie v R (pozri obrázok 6.3 vpravo)

```
delta <- c(0,0.5,0.8,1,1.2)
sigma <- 1.4
n <- 26
lambda <- delta/sigma*sqrt(n)
curve(dt(x,n,ncp=delta[1]),from=-4,to=9,xlab="x",ylab="hustota",xlim=c(-4,12))
curve(dt(x,n,ncp=delta[2]),add=T,from=-5,to=10,lty=2)
curve(dt(x,n,ncp=delta[3]),add=T,from=-5,to=10,lty=3)
curve(dt(x,n,ncp=delta[4]),add=T,from=-5,to=10,lty=4)
curve(dt(x,n,ncp=delta[5]),add=T,from=-5,to=10,lty=5)
legend("topright",c(expression(paste(delta==0,sep="")),
        expression(paste(delta==0.5,sep="")),expression(paste(delta==0.8,sep="")),
        expression(paste(delta==1,sep="")),expression(paste(delta==1.2,sep=""))),
        lty=1:5)
```



Obr. 6.3: Distribučná funkcia (vľavo) a hustoty (vpravo) necentrálneho t-rozdelenia pri rôznych parametroch necentrality λ vyjadreného pomocou δ

Príklad 179 *Nech* $X \sim N(\mu, \sigma^2)$, *kde odhady* $\overline{x} = 4$ *a* $s^2 = 2.89^2$. *Rozsah náhodného výberu* $n = 25$.

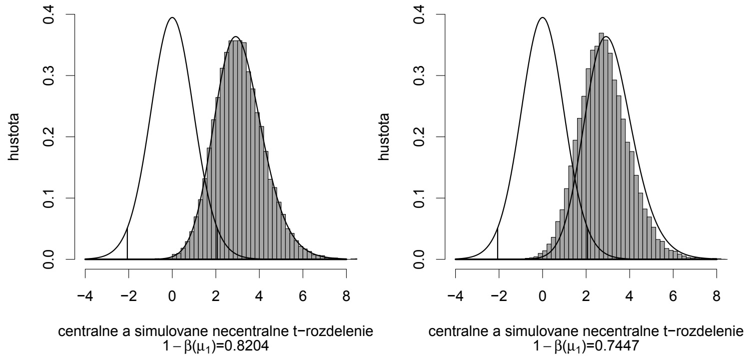
(a) Testujte $H_0 : \mu = 2.5$ *oproti* $\mu \neq 2.5$ *na hladine významnosti* $\alpha = 0.05$.

(b) Vypočítajte silu $1 - \beta$ *pre* $\mu_0 = 2.5$ *a* $\mu_1 = 4$ *(*μ_1 *predstavuje hodnotu* μ *za platnosti* H_1*) za predpokladu, že* $\sigma = 2.5$.

(c) Použite R na simuláciu hustoty rozdelenia $t_{n-1,\lambda}$ *testovacích štatistík* $t^{(m)}_{W,\lambda} = \frac{\overline{x}_m - \mu_0}{s_m}\sqrt{n}$ *(necentrálne t-rozdelenie s* $n-1$ *stupňami voľnosti a parametrom necentrality* λ*), kde* $n = 25$, $\lambda = 3$, $m = 1, 2, \ldots, M$, *pri* $M = 20000$ *opakovaniach. Na základe tohoto rozdelenia výpočítajte silu testu pre* $\mu_0 = 2.5$ *a* $\mu_1 = 4$ *(pozri obrázok 45). (1)* $X \sim N(4, 2.5^2)$ *a (2)* $X \sim [pN(4, 2.5^2) + (1-p)N(4, 4.5^2)]$, *kde* $p = 0.9$.

Riešenie (aj v R)
(a) $t_W = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \doteq \frac{4-2.5}{2.89/\sqrt{25}} \doteq 2.595$, p-hodnota=$2\times \Pr(|T_W| \geq t_{\text{obs}} | T_W \sim t_{24}) \doteq 0.016$;
(b) $\lambda = \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} \doteq \frac{4-2.5}{2.5/\sqrt{25}} \doteq 3.0$;
Za platnosti alternatívy $t_W = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$.
$1 - \beta(\mu_1 = 4) = \Pr(\text{zamietni } H_0|H_1) = \Pr(T_W < t_{n-1}(1-\alpha/2) \cup T_W > t_{n-1}(\alpha/2)|T \sim t_{n-1,\lambda}) = \Pr(T_{W,3} < t_{24}(0.975) \cup T_{W,3} > t_{24}(0.025)|T_{W,3} \sim t_{24,3}) \doteq 0.82$.
Platí nasledovné: $\Pr(T_{W,3} < t_{24}(0.975) \cup T_{W,3} > t_{24}(0.025)|T_{W,3} \sim t_{24,3}) = \Pr(F_{1,24,3^2} > F_{1,24}(0.05))$, kde $F_{1,24}(0.05) = (t_{24}(0.025))^2$.

```
cv <- qt(0.975,24) # 2.063899
p.hodn <- 2*(1-pt(2.595156,24)) # 0.01587727
sila.t.test <- pt(qt(0.025,24),24,3)+(1-pt(qt(0.975,24),24,3)) # 0.8207219
sila.t.test <- 1-pf(qt(0.975,24)^2,1,24,3^2) # 0.8207219
```



Obr. 6.4: Hustota centrálneho a necentrálneho t-rozdelenia; vľavo – hustota necentrálneho t-rozdelenia je superponovaná histogramom simulácií pre $X \sim N(4, 2.5^2)$ a vpravo – pre $X \sim [pN(4, 2.5^2) + (1-p)N(4, 4.5^2)]$, kde $p = 0.9$

Simuláciu silofunkcie (1) pozri na obrázku 6.4 vľavo a (2) na obrázku 6.4 vpravo.

Pravdepodobnosť empirickej CHPD pre MC experiment je pravdepodobnosť p signifikantných testovacích štatistík medzi ich M opakovaniami, ak H_0 platí. Potom $SE(p) = \sqrt{p(1-p)/M}$ je menšia alebo rovná $0.5/\sqrt{M}$.

Príklad 180 (pravdepodobnosť empirickej CHPD t-testu) *Nech $X \sim N(\mu, \sigma^2)$, kde $\mu = 500$ a $\sigma^2 = 100$. Testujte $H_0 : \mu = 500$ oproti $H_1 : \mu > 500$, ak $\alpha = 0.05$, σ je neznáme. Použite R na simuláciu empirickej* Pr*(CHPD), kde počet simulácií je $M = 10000$ a rozsah náhodného výberu je $n = 20$ pre jednovýberový Studentov t-test o strednej hodnote μ. Použite funkciu* `t.test(x, alternative = "greater", mu = mu0)` *a pre každú testovaciu štatistiku $t_m, m = 1, 2, \ldots, M$ vypočítajte p-hodnotu a jej štandardnú chybu za platnosti H_0. Ide o zistenie relatívnej početnosti p zamietnutých H_0 na hladine významnosti $\alpha = 0.05$ medzi M testami, kde $p = \Pr(CHPD) = \frac{\sum_{m=1}^{M} I(H_0 \ zamietame)}{M}$.*

Riešenie v R

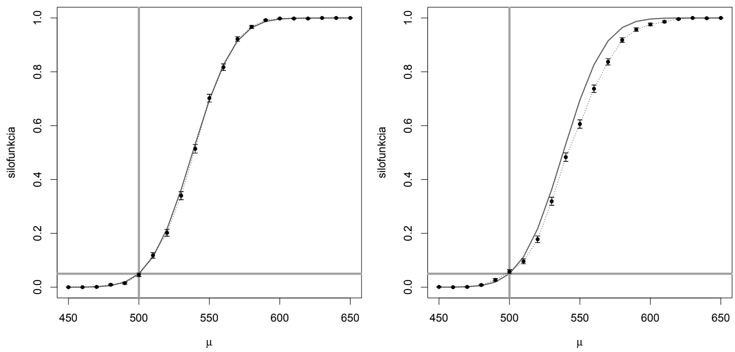
```
n <- 20
alfa <- .05
mu0 <- 500
sigma <- 100
M <- 10000
p.hodnoty <- numeric(M)
for (j in 1:M) {
      x <- rnorm(n,mu0,sigma)
      ttest <- t.test(x,alternative="greater",mu=mu0)
      p.hodnoty[j] <- ttest$p.value
      }
p.hat <- mean(p.hodnoty<alfa)
se.hat <- sqrt(p.hat*(1-p.hat)/M)
print(c(p.hat,se.hat)) # 0.052100000 0.002222287
```

Odhadnutá pravdepodobnosť chyby prvého druhu bude oscilovať okolo **nominálnej hladiny významnosti** $\alpha = 0.05$, pretože všetky náhodné výbery boli generované za platnosti H_0 za predpokladu modelu normálneho rozdelenia.

Ak nie je možné vypočítať silofunkciu testu H_0 oproti fixovanej obojstrannej H_1 analyticky, je možné túto silu odhadnúť pomocou MC metód. Treba si ale uvedomiť, že hoci sila $1-\beta$ je definovaná na celom parametrickom priestore Θ, α je definovaná na podpriestore Θ_0.

Príklad 181 (empirická silofunkcia t-testu) *Nech (a) X pochádza z normálneho rozdelenia, $X \sim N(\mu_1, 100^2)$, a (b) X pochádza zo zmesi dvoch normálnych rozdelení, $X \sim [pN(\mu_1, 100^2) + (1-p)N(\mu_1, 200^2)]$, kde $p = 0.9$. Rozsah náhodného výberu $n = 20$. Použite R na simuláciu empirickej silofunkcie pre jednovýberový Studentov t-test. Testujeme $H_0 : \mu = 500$ oproti $H_1 : \mu \neq 500$, kde $\mu_1 = 450, 460, \ldots, 640, 650$ (ide o obojstrannú alternatívu). Použite funkciu* `t.test(x,mu=500)`*, na výpočet každej testovacej štatistiky $t_m, m = 1, 2, \ldots, M$, kde $M = 10000$, vypočítajte p-hodnotu* *korešpondujúcu t_m a porovnajte ju s hladinou významnosti $\alpha = 0.05$. Tak získate empirickú silofunkciu $1-\widehat{\beta}(\mu_1)$ pri danej alternatíve. Do grafu zakreslite $1-\widehat{\beta}(\mu_1)$ pri danej alternatíve ako aj ich štandardné chyby $SE[\widehat{1-\widehat{\beta}(\mu_1)}] = \sqrt{\frac{(1-\widehat{\beta}(\mu_1))\widehat{\beta}(\mu_1)}{M}}$ v podobe chybovej úsečky $1-\widehat{\beta}(\mu_1) \pm SE[\widehat{1-\widehat{\beta}(\mu_1)}]$. Do grafu vkreslite aj teoretickú silofunkciu $1-\beta(\mu_1), \mu_1 \in \langle 450, 650\rangle$ (použite funkciu* `power.t.test()`*).*

Riešenie (pozri obrázok 6.5)



Obr. 6.5: Empirická (krivka s chybovými úsečkami vo vybraných bodoch) a teoretická (hladká krivka) silofunkcia t-testu; simulácie; vľavo – $X \sim N(500, 100^2)$ a vpravo – $X \sim [pN(500, 100^2) + (1-p)N(500, 200^2)]$, kde $p = 0.9$

Test pomerom vierohodnosti pre μ.

Majme H_{01} oproti H_{11}. Nech $X \sim N(\mu, \sigma^2)$, kde $\boldsymbol{\theta} = (\mu, \sigma^2)^T$. Funkcia vierohodnosti je rovná

$$L(\boldsymbol{\theta}|\mathbf{x}) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right).$$

MLE $\boldsymbol{\theta}$ je rovný

$$\widehat{\boldsymbol{\theta}} = (\overline{x}, \widehat{\sigma}^2)^T = \left(\frac{1}{n}\sum_{i-1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2\right)^T,$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \mu \neq \mu_0\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\mu_0, \widehat{\sigma}_0^2)^T$ t.j. $\Theta_0 = \{\boldsymbol{\theta} : \mu = \mu_0\}$. Výpočet $\max(l(\boldsymbol{\theta}|\mathbf{x}))|\boldsymbol{\theta} \in \Theta_0)$ sa redukuje na výpočet MLE $\widehat{\sigma}_0^2$ parametra σ^2 za predpokladu $\mu = \mu_0$ je známe, kde ďalej použijeme $l(\boldsymbol{\theta}|\mathbf{x})$ v bode $\boldsymbol{\theta}_0$. Derivácia $l(\boldsymbol{\theta}|\mathbf{x})$ podľa σ^2 je rovná

$$\frac{\partial}{\partial\sigma^2} l(\boldsymbol{\theta}|\mathbf{x}) = \frac{1}{2}\left(\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i-\mu_0)^2 - \frac{n}{\sigma^2}\right) = 0,$$

odkiaľ

$$\widehat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2.$$

Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{x})) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x}) - l(\boldsymbol{\theta}_0|\mathbf{x}) \\ &= \left(-\frac{n}{2}(\ln(2\pi) + \ln\widehat{\sigma}^2 + 1)\right) - \left(-\frac{n}{2}(\ln(2\pi) + \ln\widehat{\sigma}_0^2 + 1)\right) \\ &= \frac{n}{2}\ln\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}^2} \end{aligned}$$

a **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln(\lambda(\mathbf{X})) \stackrel{\mathcal{D}}{\sim} \chi^2_1$, kde H_0 bude zamietnutá pre veľké hodnoty podielu $\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2}$. Dá sa ukázať, že $\widehat{\sigma}^2_0 = \widehat{\sigma}^2 + (\overline{x} - \mu_0)^2$ a potom $\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2} = 1 + \frac{(\overline{x}-\mu_0)^2}{\widehat{\sigma}^2}$. Keďže $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2 = \frac{n}{n-1}\widehat{\sigma}^2$, podiel $\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2}$ je rastúcou funkciou $|t_W|$ a potom

$$u_{\mathrm{LR}} = -2\ln(\lambda(\mathbf{x})) = n\ln\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2} = n\ln\left(1 + \frac{t^2_W}{n-1}\right).$$

Vierohodnostný $100 \times (1-\alpha)\%$ **empirický DIS pre** μ bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{\mu_0 : U_{\mathrm{LR}} < \chi^2_1(\alpha)\right\}.$$

Majme H_{02} oproti H_{12}. Nech $\widehat{\mu}_0 = \overline{x}$ ak $\overline{x} \leq \mu_0$ a $\widehat{\mu}_0 = \mu_0$ inak. Potom $u_{\mathrm{LR}} = 0$ ak $t_W \leq 0$ a $u_{\mathrm{LR}} = \frac{n}{2}\ln(1 + \frac{t^2_W}{n-1})$ pre $t_W > 0$.

Majme H_{03} oproti H_{13}. Nech $\widehat{\mu}_0 = \overline{x}$ ak $\overline{x} \geq \mu_0$ a $\widehat{\mu}_0 = \mu_0$ inak. Potom $u_{\mathrm{LR}} = 0$ ak $t_W \geq 0$ a $u_{\mathrm{LR}} = \frac{n}{2}\ln(1 + \frac{t^2_W}{n-1})$ pre $t_W < 0$.

Príklad 182 (test o strednej hodnote μ) *Majme dáta* one-sample-mean-skull-mf.txt *a premennú dĺžka lebky* skull.L *v mm starovekej egyptskej mužskej populácie, o ktorej predpokladáme, že má normálne rozdelenie $N(\mu, \sigma^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty dĺžky lebky tejto populácie so strednou hodnotou dĺžky lebky novovekej egyptskej mužskej populácie 177.568 mm na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre strednú hodnotu dĺžky lebky, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku T_W, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2). Výsledky Waldovho testu skontrolujte pomocou funkcie* t.test()*.*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: stredná hodnota dĺžky lebky starovekej egyptskej mužskej populácie je zhodná so strednou hodnotou dĺžky lebky novovekej egyptskej mužskej populácie oproti H_1: stredná hodnota dĺžky lebky starovekej egyptskej mužskej populácie nie je zhodná so strednou hodnotou dĺžky lebky novovekej egyptskej mužskej populácie.
   - **matematická formulácia** – $H_0 : \mu = \mu_0$ oproti $H_1 : \mu \neq \mu_0$, kde $\mu_0 = 177.568$.
2. **Testovacia štatistika** – Stredná hodnota μ a rozptyl σ^2 sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetický priemer $\overline{x} \doteq 182.037$ a rozptyl $s^2 \doteq 40.582$ (smerodajná odchýlka $s \doteq 6.370$).

```
DATA <- read.table("one-sample-mean-skull-mf.txt",header=TRUE)
attach(DATA)
x <- na.omit(skull.L)
n <- length(x) # 217
priemer <- mean(x) # 182.0369
rozptyl <- var(x) # 40.58197
smerodch <- sd(x) # 6.370398
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $t_W = \frac{\overline{x}-\mu_0}{s}\sqrt{n} = \frac{182.037-177.568}{6.370}\sqrt{217} \doteq 10.334$.
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} = n\ln\left(1 + \frac{t^2_W}{n-1}\right) \doteq 87.172$.

```
mu0 <- 177.568
tW.obs <- (priemer-mu0)/smerodch*sqrt(n) # 10.33382
sigma.sq.hat <- (n-1)/n*var(x) # 40.58197
sigma0.sq.hat <- sum((x-mu0)^2)/n # 60.36572
uLR.obs <- n*log(sigma0.sq.hat/sigma.sq.hat) # 87.17249
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $t_{n-1}(1-\alpha/2) = t_{216}(1-0.025) \doteq -1.971$ a $t_{n-1}(\alpha/2) = t_{216}(0.025) \doteq 1.971$;
   kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-1}(1-\alpha/2)) \cup (t_{n-1}(\alpha/2), t_{\max}) = (-\infty, -1.971) \cup (1.971, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi^2_1(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi^2_1(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
t.krit.d <- qt(0.025,df=n-1) # -1.971007
t.krit.h <- qt(0.975,df=n-1) # 1.971007
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100 \times (1-\alpha)\%$ empirický DIS = 95% empirický DIS pre μ:
   $(d, h) = (\overline{x} - t_{n-1}(1-\alpha/2)\frac{s}{\sqrt{n}}, \overline{x} + t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}) \doteq (182.037 - 1.971\frac{6.370}{\sqrt{30}}, 182.037 + 1.971\frac{6.370}{\sqrt{30}}) = (181.185, 182.890)$.
   Vierohodnostný 95% empirický DIS pre μ:
   $\mathcal{CS}_{0.95} = \{\mu_0 : u_{\mathrm{LR}}(\mu_0) < \chi^2_1(\alpha)\} \doteq (181.188, 182.886)$.

```
IS.W <- priemer+c(-1,1)*t.krit.h*smerodch/sqrt(n) # 181.1845 182.8892
min.mu <- priemer-1*sd(x)
max.mu <- priemer+1*sd(x)
mu.0i <- seq(min.mu,max.mu,by = 0.0001)
tW.obs.i <- (priemer-mu.0i)/sd(x)*sqrt(n)
uW.obs.i <- tW.obs.i^2
uLR.i <- n*log(1+uW.obs.i/(n-1))
IS.LR <- range(mu.0i[which(uLR.i<lr.krit.hodn)]) # 181.1876 182.8862
```

5. **Štatistický záver** –
   Waldov test: p-hodnota $\doteq 2\Pr(T_W \geq |10.334||H_0) < 0.0001$.
   Test pomerom vierohodnosti: p-hodnota $\doteq \Pr(U_{\mathrm{LR}} \geq 87.172|H_0) < 0.0001$.

```
p.hodn.W <- 2*(1-pt(abs(tW.obs),df=n-1)) # < 0.0001
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # < 0.0001
```

   H_0 na hladine významnosti zamietame, pretože (1) testovacia štatistika patrí kritickému oboru, (2) $\mu_0 = 177.568$ nepatrí DIS a nakoniec (3) p-hodnota je menšia ako 0.05.

6. **Slovný záver** – Zamietame nulovú hypotézu o tom, že stredná hodnota dĺžky lebky starovekej egyptskej mužskej populácie je zhodná so strednou hodnotou dĺžky lebky novovekej egyptskej mužskej populácie.

7. **Antropologický slovný záver** – V starovekej egyptskej populácii je hodnota dĺžky lebky štatisticky významne vyššia než v populácii novovekej. V uvedených populáciách by tak mohlo ísť o prejav sekulárneho trendu v zmysle skracovania dĺžky lebky, keďže ide o zmenu určitého parametra za dlhé časové obdobie. Vyvodzovanie záverov však vyžaduje testovanie rozdielov aj v šírkových rozmeroch, keďže brachycefalizácia predstavuje relatívne skracovanie lebky a ukazuje sa, že za zmeny v tvare lebky zodpovedajú predovšetkým zmeny v jej šírke.

Riešenie pomocou funkcie t.test():

```
W.test <- t.test(x,mu=mu0)
W.test$estimate # 182.0369
W.test$conf.int # 181.1845 182.8892
W.test$stat # 10.33382
W.test$p.val # 1.344108e-20
```

Príklad 183 (vierohodnostný DIS pre μ) *Majme dáta* one-sample-mean-skull-mf.txt *a premennú dĺžka lebky* skull.L *v mm starovekej egyptskej mužskej populácie, o ktorej predpokladáme, že má normálne rozdelenie $N(\mu, \sigma^2)$. Vypočítajte vierohodnostný 95% empirický DIS pre strednú hodnotu dĺžky lebky μ pomocou 15% cut-off relatívnej (štandardizovanej) funkcie vierohodnosti $\mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) = L(\boldsymbol{\theta}|\mathbf{x})/L(\widehat{\boldsymbol{\theta}}|\mathbf{x})$ a porovnajte ho s vierohodnostným 95% empirický DIS pre μ.*

Príklad 184 (test o strednej hodnote μ) *Majme dáta* one-sample-mean-skull-mf.txt *a premennú šírku lebky* skull.B *v mm starovekej egyptskej mužskej populácie, o ktorej predpokladáme, že má normálne rozdelenie $N(\mu, \sigma^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty šírky lebky tejto populácie so strednou hodnotou šírky lebky novovekej egyptskej mužskej populácie* 136.402 *mm na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre strednú hodnotu šírky lebky, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku T_W, (2) testovaciu štatistiku pomeru vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2).*

Párový prípad. Predpokladajme, že $X_j \sim N(\mu_j, \sigma_j^2)$, kde σ_j^2 nie sú známe a $j = 1, 2$. Potom $X_d = X_1 - X_2 \sim N(\mu_d, \sigma_d^2)$, kde $\mu_d = \mu_1 - \mu_2$, $\sigma_d = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$ a ρ je korelačný koeficient medzi premennými X_1 a X_2. Tiež platí $\overline{X}_d = \overline{X}_1 - \overline{X}_2 \sim N(\mu_d, \sigma_d^2/n)$. Majme dvojice hypotéz $H_{01} : \mu_d = \mu_0$ oproti $H_{11} : \mu_d \neq \mu_0$, $H_{02} : \mu_d \leq \mu_0$ oproti $H_{12} : \mu_d > \mu_0$, $H_{03} : \mu_d \geq \mu_0$ oproti $H_{13} : \mu_d < \mu_0$, ktoré chceme testovať. Na ich testovanie použijeme

1. **Waldovu testovaciu štatistiku** $t_W^{(d)}$, ktorá je ekvivalentná t_W, kde $\overline{x}$ nahradíme $\overline{x}_d$ a s nahradíme $s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_{d,i} - \overline{x}_d)^2}$; prislúchajúci test nazývame **párový Studentov t-test o strednej hodnote μ_d**;

2. **testovaciu štatistiku pomerom vierohodnosti** $u_{\text{LR}}^{(d)}$, ktorá je ekvivalentná u_{LR}, kde $\widehat{\sigma}_0^2$ nahradíme $\widehat{\sigma}_{d,0}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{d,i} - \mu_d)^2$ a $\widehat{\sigma}^2$ nahradíme $\widehat{\sigma}_d^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{d,i} - \overline{x}_d)^2$.

Párový prípad vzniká v praxi v troch základných situáciách:

1. jeden výskumník vykoná dve opakované merania jednej premennej v čase (sledovanou premennou je rozdiel merania v prvom a druhom časovom bode); merania slúžia na zistenie **intraindividuálnej chyby**;

2. dvaja výskumníci vykonajú každý jedno meranie jednej premennej (sledovanou premennou je rozdiel meraní prvého a druhého výskumníka); merania slúžia na zistenie **interindividuálnej chyby**;

3. jeden výskumník vykoná jedno meranie premennej na pravej a jedno na ľavej strane tela (sledovanou premennou je rozdiel meraní na pravej a ľavej strane); merania slúžia na zistenie **(a)symetrie**.

Párový Studentov t-test o strednej hodnote μ_d v R (funkcia **t.test()**).
Argumenty (vstupy) funkcie:

1. vektor dát x;
2. vektor dát y;
3. alternatíva alternative="two.sided" je prednastavená, ďalšie voľby sú "greater", "less";
4. stredná hodnota za platnosti nulovej hypotézy mu, t.j. μ_0;
5. párový prípad paired = TRUE (prednastavený je jednovýberový prípad paired = FALSE);
6. spoľahlivosť conf.level=p, prednastavené je p =.95.

Výstupy funkcie:

1. názov použitého testu method;
2. testovacia štatistika statistic;
3. počet stupňov voľnosti df parameter;
4. p-hodnota p.value;

5. alternatívna hypotéza alternative hypothesis;
6. interval spoľahlivosti pre μ_d conf.int;
7. bodový odhad rozdielu stredných hodnôt (aritmetický priemer $\overline{x}_d$) sample estimates.

Príklad 185 (test o strednej hodnote μ_d) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v podobe jej prvého (*simd.1*) a druhého (*simd.2*) merania v mm na pravej strane (*side *je rovná R). Zaujíma nás rozdiel prvého a druhého merania (**intraindividuálna chyba**) na pravej strane. O tomto rozdiele X_d predpokladáme, že má normálne rozdelenie $N(\mu_d, \sigma_d^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre strednú hodnotu rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku $T_W^{(d)}$, (2) testovaciu štatistiku pomeru vierohodnosti $U_{LR}^{(d)}$ a DIS prislúchajúce (1) a (2).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: stredná hodnota rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane je rovná nule oproti H_1: stredná hodnota rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane nie je rovná nule.
   - **matematická formulácia** – $H_0 : \mu_d = \mu_0$ oproti $H_1 : \mu_d \neq \mu_0$, kde $\mu_0 = 0$.

2. **Testovacia štatistika** – Stredná hodnota μ_d a rozptyl σ_d^2 sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetický priemer $\overline{x}_d \doteq 0.011$ a rozptyl $s_d^2 \doteq 0.040$ (smerodajná odchýlka $s_d \doteq 0.200$).

```
DATA <- read.table("paired-means-clavicle.txt",header=TRUE)
attach(DATA)
xd <- simd.1[side=="R"]-simd.2[side=="R"]
n <- length(xd) # 40
priemer <- mean(xd) # 0.01075
rozptyl <- var(xd) # 0.04015583
smerodch <- sd(xd) # 0.2003892
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $t_W^{(d)} = \frac{\overline{x}_d - \mu_0}{s_d}\sqrt{n} \doteq \frac{0.011-0}{0.200}\sqrt{40} \doteq 0.339$.
   Testovacia štatistika pomerom vierohodnosti $u_{\text{LR}}^{(d)} = n \ln\left(1 + \frac{(t_W^{(d)})^2}{n-1}\right) \doteq 0.118$.

```
mu0 <- 0
tdW.obs <- (priemer-mu0)/smerodch*sqrt(n) # 0.3392846
sigmad.sq.hat <- (n-1)/n*var(xd) # 0.03915194
sigmad0.sq.hat <- sum((xd-mu0)^2)/n # 0.0392675
udLR.obs <- n*log(sigmad0.sq.hat/sigmad.sq.hat) # 0.1178918
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $t_{n-1}(1-\alpha/2) = t_{39}(1-0.025) \doteq -2.023$ a $t_{n-1}(\alpha/2) = t_{39}(0.025) \doteq 2.023$;
   kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-1}(1-\alpha/2)) \cup (t_{n-1}(\alpha/2), t_{\max}) = (-\infty, -2.023) \cup (2.023, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi_1^2(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi_1^2(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
t.krit.d <- qt(0.025,df=n-1) # -2.022691
t.krit.h <- qt(0.975,df=n-1) # 2.022691
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
Waldov $100 \times (1-\alpha)\%$ empirický DIS=95% empirický DIS pre μ_d:
$(d,h) = (\overline{x}_d - t_{n-1}(1-\alpha/2)\frac{s_d}{\sqrt{n}}, \overline{x}_d + t_{n-1}(\alpha/2)\frac{s_d}{\sqrt{n}}) \doteq (-0.053, 0.075)$.
Vierohodnostný 95% empirický DIS pre μ_d:
$\mathcal{CS}_{0.95} = \left\{\mu_0 : u_{\mathrm{LR}}^{(d)}(\mu_0) < \chi_1^2(\alpha)\right\} \doteq (-0.052, 0.074)$.

```
IS.dW <- priemer+c(-1,1)*t.krit.h*smerodch/sqrt(n) # -0.05333758   0.07483758
min.mu <- priemer-1*sd(xd)
max.mu <- priemer+1*sd(xd)
mu.0i <- seq(min.mu,max.mu,by=0.0001)
tdW.obs.i <- (priemer-mu.0i)/sd(xd)*sqrt(n)
udW.obs.i <- tdW.obs.i^2
udLR.i <- n*log(1+udW.obs.i/(n-1))
IS.dLR <- range(mu.0i[which(udLR.i<lr.krit.hodn)]) # -0.0520392   0.0735608
```

5. **Štatistický záver** –
Waldov test: p-hodnota $\doteq 2\mathrm{Pr}(T_W^{(d)} \geq |0.339||H_0) \doteq 0.736$.
Test pomerom vierohodnosti: p-hodnota $\doteq \mathrm{Pr}(U_{\mathrm{LR}} \geq 0.118|H_0) \doteq 0.731$.

```
p.hodn.dW <- 2*(1-pt(abs(tdW.obs),df=n-1)) # 0.7362156
p.hodn.dLR <- 1-pchisq(udLR.obs,df=1) # 0.7313324
```

H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\mu_0 = 0$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že stredná hodnota rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane je rovná nule.

7. **Antropologický slovný záver** – Rozdiely medzi opakovanými meraniami sú dostatočne nízke, intraindividuálna chyba akceptovateľná a predpoklady ďalších analýz splnené.

Riešenie pomocou funkcie t.test():

```
Wd.test <- t.test(simd.1[side=="R"],simd.2[side=="R"],mu=0,paired=TRUE)
Wd.test$estimate # 0.01075
Wd.test$conf.int # -0.05333758   0.07483758
Wd.test$stat # 0.3392846
Wd.test$p.val #  0.7362156
```

Príklad 186 (test o strednej hodnote μ_d) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v podobe jej prvého (*simd.1*) a druhého (*simd.2*) merania v mm na ľavej strane (*side *je rovná* L*). Zaujíma nás rozdiel prvého a druhého merania (**intraindividuálna chyba**) na ľavej strane. O tomto rozdiele X_d predpokladáme, že má normálne rozdelenie $N(\mu_d, \sigma_d^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej strane na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre strednú hodnotu rozdielu prvého a druhého merania vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej strane, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku $T_W^{(d)}$, (2) testovaciu štatistiku pomeru vierohodnosti $U_{LR}^{(d)}$ a DIS prislúchajúce (1) a (2).*

Príklad 187 (test o strednej hodnote μ_d) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v mm (priemer prvého a druhého merania* simd.1 *a* simd.2*) na ľavej (*side *sa rovná* L*) a pravej (*side *sa rovná* R*). Tieto merania vykonal prvý výskumník. Druhý výskumník zmeral vertikálny priemer v strede dĺžky tela kľúčnej kosti jedenkrát (*simd*). Zaujíma nás rozdiel meraní oboch výskumníkov (**interindividuálna chyba**) na pravej strane. O tomto rozdiele X_d predpokladáme, že má normálne rozdelenie $N(\mu_d, \sigma_d^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti u prvého a druhého výskumníka* *na pravej strane na hladine významnosti* $\alpha = 0.05$. *(b) Vypočítajte* $100 \times (1-\alpha)\%$ *empirický DIS pre strednú hodnotu rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej a pravej strane, kde koeficient spoľahlivosti* $1-\alpha = 0.95$. *Použite (1) Waldovu testovaciu štatistiku* $T_W^{(d)}$, *(2) testovaciu štatistiku pomeru vierohodnosti* $U_{LR}^{(d)}$ *a DIS prislúchajúce (1) a (2).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: stredná hodnota rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti u prvého a druhého výskumníka na pravej strane je rovná nule oproti H_1: stredná hodnota rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti u prvého a druhého výskumníka na pravej strane nie je rovná nule.
   - **matematická formulácia** – $H_0: \mu_d = \mu_0$ oproti $H_1: \mu_d \neq \mu_0$, kde $\mu_0 = 0$.

2. **Testovacia štatistika** – Stredná hodnota μ_d a rozptyl σ_d^2 sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetický priemer $\overline{x}_d \doteq -0.457$ a rozptyl $s_d^2 \doteq 0.709$ (smerodajná ochýlka $s_d \doteq 0.842$).

```r
DATA <- read.table("paired-means-clavicle.txt",header=TRUE)
attach(DATA)
simd.xbar.R <- (simd.1[side=="R"]+simd.2[side=="R"])/2
xd <- na.omit(simd.xbar.R-simd[side=="R"])
n <- length(xd) # 39
priemer <- mean(xd) # -0.4569231
rozptyl <- var(xd) # 0.7085442
smerodch <- sd(xd) # 0.8417507
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $t_W^{(d)} = \frac{\overline{x}_d-\mu_0}{s_d}\sqrt{n} \doteq \frac{-0.457-0}{0.842}\sqrt{39} \doteq -0.390$.
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}}^{(d)} = n\ln\left(1+\frac{(t_W^{(d)})^2}{n-1}\right) \doteq 10.305$.

```r
mu0 <- 0
tdW.obs <- (priemer-mu0)/smerodch*sqrt(n) # -3.389939
sigmad.sq.hat <- (n-1)/n*var(xd) # 0.6903764
sigmad0.sq.hat <- sum((xd-mu0)^2)/n # 0.899155
udLR.obs <- n*log(sigmad0.sq.hat/sigmad.sq.hat) # 10.30452
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $t_{n-1}(1-\alpha/2) = t_{38}(1-0.025) \doteq -2.024$ a $t_{n-1}(\alpha/2) = t_{38}(0.025) \doteq 2.024$;
   kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-1}(1-\alpha/2)) \cup (t_{n-1}(\alpha/2), t_{\max}) = (-\infty, -2.024) \cup (2.024, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi_1^2(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi_1^2(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```r
t.krit.d <- qt(0.025,df=n-1) # -2.024394
t.krit.h <- qt(0.975,df=n-1) # 2.024394
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100 \times (1-\alpha)\%$ empirický DIS = 95% empirický DIS pre μ_d:
   $(d,h) = (\overline{x}_d - t_{n-1}(1-\alpha/2)\frac{s_d}{\sqrt{n}}, \overline{x}_d + t_{n-1}(\alpha/2)\frac{s_d}{\sqrt{n}}) \doteq (-0.730, -0.184)$.
   Vierohodnostný 95% empirický DIS pre μ_d:
   $\mathcal{CS}_{0.95} = \left\{\mu_0 : u_{\mathrm{LR}}^{(d)}(\mu_0) < \chi_1^2(\alpha)\right\} \doteq (-0.724, -0.190)$.

```r
IS.dW <- priemer+c(-1,1)*t.krit.h*smerodch/sqrt(n) # -0.7297871 -0.1840591
min.mu <- priemer-1*sd(xd)
max.mu <- priemer+1*sd(xd)
mu.0i <- seq(min.mu,max.mu,by = 0.0001)
tdW.obs.i <- (priemer-mu.0i)/sd(xd)*sqrt(n)
udW.obs.i <- tdW.obs.i^2
udLR.i <- n*log(1+udW.obs.i/(n-1))
IS.dLR <- range(mu.0i[which(udLR.i<lr.krit.hodn)]) # -0.7241738 -0.1896738
```

5. **Štatistický záver** –
   Waldov test: p-hodnota $\doteq 2\Pr(T_W^{(d)} \geq |-0.390||H_0) \doteq 0.002$.
   Test pomerom vierohodnosti: p-hodnota $\doteq \Pr(U_{\mathrm{LR}}^{(d)} \geq 10.305|H_0) \doteq 0.001$.

```
p.hodn.dW <- 2*(1-pt(abs(tdW.obs),df=n-1)) # 0.001642062
p.hodn.dLR <- 1-pchisq(udLR.obs,df=1) # 0.001327044
```

   H_0 na hladine významnosti zamietame, pretože (1) testovacia štatistika patrí kritickému oboru, (2) $\mu_0 = 0$ nepatrí DIS a nakoniec (3) p-hodnota je menšia ako 0.05.

6. **Slovný záver** – Zamietame nulovú hypotézu o tom, že stredná hodnota rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti u prvého a druhého výskumníka na pravej strane je rovná nule.

7. **Antropologický záver** – Hodnoty meraní prvého výskumníka sú v priemere o 0.46 mm vyššie než hodnoty druhého výskumníka. Dôvodom tohto štatisticky významného rozdielu môže byť odlišnosť v kalibrácii merania, keďže obaja výskumníci merali iným prístrojom. Iným možným vysvetlením systematického rozdielu je odlišné chápanie definície rozmeru a vertikály na kosti.

Riešenie pomocou funkcie `t.test()`:

```
Wd.test <- t.test(simd.xbar.R,simd[side=="R"],mu=0,paired=TRUE)
Wd.test$estimate # -0.4569231
Wd.test$conf.int # -0.7297871 -0.1840591
Wd.test$stat # -3.389939
Wd.test$p.val # 0.001642062
```

Príklad 188 (test o strednej hodnote μ_d) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v mm (priemer prvého a druhého merania* simd.1 *a* simd.2*) na ľavej (*side *sa rovná* L*) a pravej (*side *sa rovná* R*). Tieto merania vykonal prvý výskumník. Druhý výskumník zmeral vertikálny priemer v strede dĺžky tela kľúčnej kosti jedenkrát (*simd*). Zaujíma nás rozdiel meraní oboch výskumníkov (**interindividuálna chyba**) na ľavej strane. O tomto rozdiele X_d predpokladáme, že má normálne rozdelenie $N(\mu_d, \sigma_d^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti u prvého a druhého výskumníka na ľavej strane na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre strednú hodnotu rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej a pravej strane, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku $T_W^{(d)}$, (2) testovaciu štatistiku pomeru vierohodnosti $U_{LR}^{(d)}$ a DIS prislúchajúce (1) a (2).*

Príklad 189 (test o strednej hodnote μ_d) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v mm (priemer prvého a druhého merania* simd.1 *a* simd.2*) na ľavej (*side *sa rovná* L*) a pravej (*side *sa rovná* R*). Zaujíma nás rozdiel priemerných meraní na pravej a ľavej strane (**(a)symetria**). O tomto rozdiele X_d predpokladáme, že má normálne rozdelenie $N(\mu_d, \sigma_d^2)$. (a) Otestujte hypotézu o zhode strednej hodnoty rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej a ľavej strane na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre strednú hodnotu rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej a pravej strane, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku $T_W^{(d)}$, (2) testovaciu štatistiku pomeru vierohodnosti $U_{LR}^{(d)}$ a DIS prislúchajúce (1) a (2).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: stredná hodnota rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej a pravej strane je rovná nule oproti H_1: stredná hodnota rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej a pravej strane nie je rovná nule.

- **matematická formulácia** – $H_0 : \mu_d = \mu_0$ oproti $H_1 : \mu_d \neq \mu_0$, kde $\mu_0 = 0$.

2. **Testovacia štatistika** – Stredná hodnota μ_d a rozptyl σ_d^2 sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetický priemer $\overline{x}_d \doteq -0.029$ a rozptyl $s_d^2 \doteq 0.611$ (smerodajná odchýlka $s_d \doteq 0.782$).

```
DATA <- read.table("paired-means-clavicle.txt",header=TRUE)
attach(DATA)
simd.xbar.R <- (simd.1[side=="R"]+simd.2[side=="R"])/2
simd.xbar.L <- (simd.1[side=="L"]+simd.2[side=="L"])/2
xd <- simd.xbar.R-simd.xbar.L
n <- length(xd) # 40
priemer <- mean(xd) # -0.029
rozptyl <- var(xd) # 0.6112695
smerodch <- sd(xd) # 0.7818373
```

Pozorované testovacie štatistiky:
Waldova testovacia štatistika $t_W^{(d)} = \frac{\overline{x}_d - \mu_0}{s_d}\sqrt{n} \doteq \frac{-0.029-0}{0.782}\sqrt{40} \doteq -0.23$.
Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}}^{(d)} = n \ln\left(1 + \frac{(t_W^{(d)})^2}{n-1}\right) \doteq 0.056$.

```
mu0 <- 0
tdW.obs <- (priemer-mu0)/smerodch*sqrt(n) # -0.2345912
sigmad.sq.hat <- (n-1)/n*var(xd) # 0.5959878
sigmad0.sq.hat <- sum((xd-mu0)^2)/n # 0.5968288
udLR.obs <- n*log(sigmad0.sq.hat/sigmad.sq.hat) # 0.05640433
```

3. **Zamietacia oblasť** –
Waldov test:
kritické hodnoty $t_{n-1}(1-\alpha/2) = t_{39}(1-0.025) \doteq -2.023$ a $t_{n-1}(\alpha/2) = t_{39}(0.025) \doteq 2.023$;
kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-1}(1-\alpha/2)) \cup (t_{n-1}(\alpha/2), t_{\max}) = (-\infty, -2.023) \cup (2.023, \infty)$.
Test pomerom vierohodnosti:
kritická hodnota $\chi_1^2(\alpha) \doteq 3.841$;
kritický obor $\mathcal{W} = (\chi_1^2(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
t.krit.d <- qt(0.025,df=n-1) # -2.022691
t.krit.h <- qt(0.975,df=n-1) # 2.022691
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
Waldov $100 \times (1-\alpha)\%$ empirický DIS = 95% empirický DIS pre μ_d:
$(d, h) = (\overline{x}_d - t_{n-1}(1-\alpha/2)\frac{s_d}{\sqrt{n}}, \overline{x}_d + t_{n-1}(\alpha/2)\frac{s_d}{\sqrt{n}}) \doteq (-0.279, 0.221)$.
Vierohodnostný 95% empirický DIS pre μ_d:
$\mathcal{CS}_{0.95} = \left\{\mu_0 : u_{\mathrm{LR}}^{(d)}(\mu_0) < \chi_1^2(\alpha)\right\} \doteq (-0.274, 0.216)$.

```
IS.dW <- priemer+c(-1,1)*t.krit.h*smerodch/sqrt(n) # -0.2790437  0.2210437
min.mu <- priemer-1*sd(xd)
max.mu <- priemer+1*sd(xd)
mu.0i <- seq(min.mu,max.mu,by=0.0001)
tdW.obs.i <- (priemer-mu.0i)/sd(xd)*sqrt(n)
udW.obs.i <- tdW.obs.i^2
udLR.i <- n*log(1+udW.obs.i/(n-1))
IS.dLR <- range(mu.0i[which(udLR.i<lr.krit.hodn)]) # -0.2740373  0.2160627
```

5. **Štatistický záver** –
Waldov test: p-hodnota $\doteq 2\Pr(T_W^{(d)} \geq |-0.235| | H_0) \doteq 0.816$.
Test pomerom vierohodnosti: p-hodnota $\doteq \Pr(U_{\mathrm{W}}^{(d)} \geq 0.056 | H_0) \doteq 0.812$.

```
p.hodn.dW <- 2*(1-pt(abs(tdW.obs),df=n-1)) # 0.8157534
p.hodn.dLR <- 1-pchisq(udLR.obs,df=1) # 0.8122721
```

H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\mu_0 = 0$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že stredná hodnota rozdielu vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na ľavej a pravej strane je rovná nule.

7. **Antropologický slovný záver** – Napriek tomu, že na rozmeroch kľúčnej kosti sú popisované výrazné systematické stranové rozdiely (kľúčna kosť ľavej strany je v priemere dlhšia a gracilnejšia), v sledovanom rozmere sme žiadnu smerovú asymetriu nedokázali. Pravdepodobne je v tomto rozmere, na rozdiel od iných, vzorec biomechnického zaťažovania na oboch stranách tela podobný.

Riešenie pomocou funkcie `t.test()`:

```
Wd.test <- t.test(simd.xbar.R,simd.xbar.L,mu=0,paired=TRUE)
Wd.test$estimate # -0.029
Wd.test$conf.int # -0.2790437  0.2210437
Wd.test$stat # -0.2345912
Wd.test$p.val # 0.8157534
```

Príklad 190 (programovanie TEM, TEM$_{\mathbf{rel}}$ a CR) *Naprogramujte v vzorce na výpočet*
(a) technickej chyby merania TEM
(b) relatívnej technickej chyby merania TEM_{rel} a
(c) koeficientu reliability merania CR.

Riešenie

$\mathrm{TEM} = \sqrt{\frac{\sum_{i=1}^{n} x_{d,i}^2}{2n}}$, kde $x_{d,i} = x_{i1} - x_{i2}$ je rozdiel medzi meraniami,
$\mathrm{TEM}_{\mathrm{rel}} = \frac{\mathrm{TEM}}{\overline{x}} \times 100$, kde $\overline{x}$ je celková priemerná hodnota,
$\mathrm{CR} = (1 - \frac{\mathrm{TEM}^2}{s^2}) \times 100$, kde s je celková smerodajná odchýlka.

```
"tem" <- function(x1,x2) sqrt(sum((x1-x2)^2)/(2*length(x1)))
"tem.rel" <- function(x1,x2) tem(x1,x2)/mean(c(x1,x2))
"cr" <- function(x1,x2) (1-(tem(x1,x2))^2/var(c(x1,x2)))
```

Príklad 191 (výpočet TEM, TEM$_{\mathbf{rel}}$ a CR) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v podobe jej prvého (*simd.1*) a druhého (*simd.2*) merania v mm na pravej strane (*side *je rovná* R*). Zaujíma nás rozdiel prvého a druhého merania na pravej strane. Vypočítajte (a) intraindividuálnu technickú chybu merania TEM, (b) intraindividuálnu relatívnu technickú chybu merania TEM_{rel} a (c) intraindividuálny koeficient reliability merania CR.*

Riešenie v

```
DATA <- read.table("paired-means-clavicle.txt",header=TRUE)
attach(DATA)
tem(simd.1[side=="R"],simd.2[side=="R"]) # 0.1401205
tem.rel(simd.1[side=="R"],simd.2[side=="R"])*100 # 1.469326
cr(simd.1[side=="R"],simd.2[side=="R"])*100 # 99.14755
```

Technická chyba merania predstavuje 0.14 mm na jedno meranie, čo zodpovedá 1.47 % aritmetického priemeru vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane. Koeficient reliability je rovný 99.15 %, čo zodpovedá vysokej presnosti a opakovateľnosti merania tohoto výskumníka.

Príklad 192 (výpočet TEM, TEM$_{\mathbf{rel}}$ a CR) *Majme dáta* paired-means-clavicle.txt *a premennú vertikálny priemer v strede dĺžky tela kľúčnej kosti v mm (priemer prvého a druhého merania* simd.1 *a* simd.2*) na ľavej (*side *sa rovná* L*) a pravej (*side *sa rovná* R*). Tieto merania vykonal prvý výskumník. Druhý výskumník zmeral vertikálny priemer v strede dĺžky tela kľúčnej kosti jedenkrát (*simd*). Zaujíma nás rozdiel meraní oboch výskumníkov na pravej strane. Vypočítajte (a) interindividuálnu technickú chybu merania TEM, (b) interindividuálnu relatívnu technickú chybu merania TEM_{rel} a (c) interindividuálny koeficient reliability merania CR.*

Riešenie v

```
DATA <- read.table("paired-means-clavicle.txt", header=TRUE)
attach(DATA)
id <- !is.na(simd[side=="R"])
simd2 <- apply(cbind(simd.1[side=="R"],simd.2[side=="R"]),1,mean)
simd2 <- simd2[id]
simd1 <- simd[side=="R"]
simd1 <- simd1[id]

tem(simd1,simd2) # 0.6705055
tem.rel(simd1,simd2)*100 # 6.834918
cr(simd1,simd2)*100 # 83.29
```

Technická chyba merania predstavuje 0.67 mm na jedno meranie, čo zodpovedá 6.83 % aritmetického priemeru vertikálneho priemeru v strede dĺžky tela kľúčnej kosti na pravej strane. Relatívna technická chyba merania je teda vyššia ako odporúčaných 5 %. Koeficient reliability je rovný 83.29 %, čo zodpovedá nízkej presnosti a veľkým rozdielom medzi oboma výskumníkmi. Koeficient reliability je teda oveľa nižší ako odporúčaných 95 %. Súbor nameraný druhým výskumníkom nie je možné použiť ako zdrojové dáta v antropologickom výskume.

6.2 Asymptotické testy o rozptyle

V biologickej antropológii je niekedy potrebné zistiť, aký je rozptyl sledovanej veličiny v danej vzorke, a či je zhodný s rozptylom v referenčnom súbore alebo publikovaným údajom v porovnávacej populácii. Rozdiely medzi populáciami v rozptyle môžu mať celý rad vysvetlení. Rozptyl biologického znaku, napr. veku menarché, dĺžky hlavy alebo stranových rozdielov v prípade bilaterálne symetrických znakov, je odrazom variability premenných (genetických i environmentálnych), ktorých odlišnosti spôsobia, že sa populácie v rozptyle tohto znaku líšia. Populácie s väčším rozptylom daného znaku môžu byť, napr. vo veľkej, reprodukčne málo obmedzenej populácii, ovplyvnené väčšou mierou genetického polymorfizmu a väčšou genetickou variabilitou. Väčší rozptyl stranových rozdielov párových znakov (vyššia miera fluktuačnej asymetrie) zas môže na úrovni populácie indikovať vyššiu vývinovú nestabilitu jedinca a horšie životné podmienky celej populácie. Znaky s malým rozptylom boli v evolúcii – alebo ešte aktuálne sú – pod silným vplyvom stabilizujúcej selekcie (u človeka napr. znaky súvisiace s bipédiou). Ak napríklad populácia žije v prostredí so špecifickými, úzko definovanými zdrojmi, ktoré obmedzujú výber jedincov na tých s vlastnosťami najvýhodnejšie prispôsobenými danému prostrediu, či už mechanizmom selekcie sociálnej (pohrebisko vojakov vyberaných do jednotky na základe výšky postavy), selekcie prírodnej alebo migrácie (migrujú len jedinci s určitými vlastnosťami), stredná hodnota sledovaného znaku sa posunúť nemusí, hodnota rozptylu sa však zníži. Podobným spôsobom ako uvedené systematické vplyvy môže rozptyl znížiť aj náhodné zmenšenie veľkosti populácie (evolučné javy známe ako efekt zakladateľa (*founder effect*), efekt hrdla fľaše (*bottleneck effect*)).

Okrem biologických premenných však rozptyl veľmi ovplyvňuje tiež spôsob vzniku vzorky (náhodný výber) a rôzne nenáhodné vplyvy vzorkovania. Ak je napríklad kostrové pohrebisko usporiadané na základe príbuzenských princípov, je možné, že vzorka záchranného výskumu z jednej obmedzenej časti pohrebiska bude obsahovať veľký podiel príbuzných jedincov s menším rozptylom silno dedičných znakov. Vzhľadom na to, že informácie o všetkých faktoroch (biologických a metodických) potenciálne ovplyvňujúcich výsledný rozptyl bývajú pri popise vzorky publikované zriedkavo, interpretácia testom zistených rozdielov v rozptyle býva neľahká. Môže totiž ísť buď o jasné biologické rozdiely medzi populáciami, alebo metodické nezrovnalosti, alebo, a to je v reálnej situácii najpravdepodobnejšie, o nejakú kombináciu oboch skupín príčin.

Nech $X \sim N(\mu, \sigma^2)$. Majme dvojice hypotéz $H_{01} : \sigma^2 = \sigma_0^2$ oproti $H_{11} : \sigma^2 \neq \sigma_0^2$, $H_{02} : \sigma^2 \leq \sigma_0^2$ oproti $H_{12} : \sigma^2 > \sigma_0^2$, $H_{03} : \mu \geq \sigma_0^2$ oproti $H_{13} : \sigma^2 < \sigma_0^2$, ktoré chceme testovať.

Ak H_0 platí, potom

$$F_W = \frac{(n-1)\,S^2}{\sigma_0^2} \overset{\mathcal{D}}{\sim} \chi^2_{n-1},$$

kde $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ a rozdelenie χ^2_{df}, kde $df = n-1$, sa nazýva **centrálne χ^2-rozdelenie** s $df = n-1$ stupňami voľnosti. F_W sa nazýva **Waldova testovacia štatistika** (alebo **χ^2-štatistika**) a test **jednovýberový χ^2-test o rozptyle σ^2**.

Chceme nájsť kritickú hodnotu $d_S^\star$ za platnosti H_{02}, tak aby platilo $\Pr\left(S^2 < d_S^\star | H_{02}\right) = \alpha$. Teda

$$\alpha = \Pr\left(\frac{(n-1)\,S^2}{\sigma_0^2} < d_S^\star \frac{n-1}{\sigma_0^2} | H_{02}\right) = \Pr\left(F_W < d_S^\star \frac{n-1}{\sigma_0^2} | H_{02}\right)$$

a z toho vyplýva, že $d_S^\star = \chi^2_{n-1}(1-\alpha)\,\frac{\sigma_0^2}{n-1}$. Ak H_{02} neplatí, t.j. skutočný rozptyl je σ_1^2 (platí H_{12}), potom chceme nájsť pravdepodobnosť zamietnutia H_{02} (silu testu). To je pravdepodobnosť, že S^2 neprekročí $d_S^\star$, teda

$$\Pr\left(S^2 < d_S^\star | H_{12}\right) = \Pr\left(\frac{(n-1)\,S^2}{\sigma_1^2} < d_S^\star \frac{n-1}{\sigma_1^2} | H_{12}\right) = \Pr\left(F_W^{(\mathrm{alt})} < \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{n-1}(1-\alpha)\,| H_{12}\right),$$

kde $F_W^{(\mathrm{alt})} = \frac{(n-1)S^2}{\sigma_1^2}$.

Chceme nájsť kritickú hodnotu $h_S^\star$ za platnosti H_{03}, tak aby platilo $\Pr\left(S^2 < h_S^\star | H_{03}\right) = 1-\alpha$. Teda

$$1-\alpha = \Pr\left(\frac{(n-1)\,S^2}{\sigma_0^2} < h_S^\star \frac{n-1}{\sigma_0^2} | H_{03}\right) = \Pr\left(F_W < h_S^\star \frac{n-1}{\sigma_0^2} | H_{03}\right)$$

a z toho vyplýva, že $h_S^\star = \chi^2_{n-1}(\alpha)\,\frac{\sigma_0^2}{n-1}$. Ak H_{03} neplatí, t.j. skutočný rozptyl je σ_1^2 (platí H_{13}), potom chceme nájsť pravdepodobnosť zamietnutia H_{03} (silu testu). To je pravdepodobnosť, že S^2 prekročí $h_S^\star$, teda

$$\Pr\left(S^2 > h_S^\star | H_{13}\right) = \Pr\left(\frac{(n-1)\,S^2}{\sigma_1^2} > h_S^\star \frac{n-1}{\sigma_1^2} | H_{13}\right) = \Pr\left(F_W^{(\mathrm{alt})} > \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{n-1}(\alpha)\,| H_{13}\right).$$

Ak by sme chceli nájsť kritické hodnoty d_S a h_S pre test H_{01} oproti H_{11}, použili by sme podobnú úvahu ako vyššie.

Príklad 193 *Odvoďte vzorec na výpočet kritických hodnôt d_S a h_S pre H_{01}.*

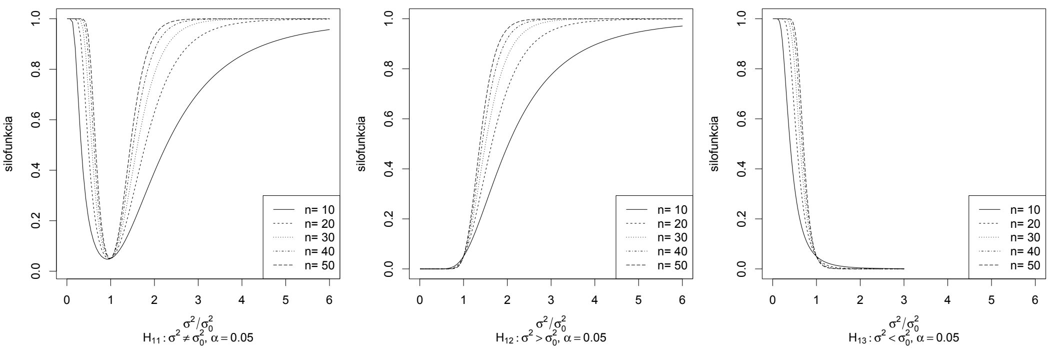
Definícia 55 (Kritický obor a silofunkcia F_W testu o σ^2) *Kritický obor a silofunkcia sú definované (za platnosti $d_{\alpha/2} = \chi^2_{n-1}(1-\alpha/2), h_{\alpha/2} = \chi^2_{n-1}(\alpha/2)$, $d_\alpha = \chi^2_{n-1}(1-\alpha)$ a $h_\alpha = \chi^2_{n-1}(\alpha)$) nasledovne:*

H_0	H_1	$\mathcal{W}$	$1-\beta\left(\sigma^2\right)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\mathcal{W}_1 = \left\{F_W; F_W \notin \left(d_{\alpha/2}, h_{\alpha/2}\right)\right\}$	$1 - \Pr\left(\frac{\sigma_0^2}{\sigma^2} d_{\alpha/2} < F_W^{(alt)} < \frac{\sigma_0^2}{\sigma^2} h_{\alpha/2}\right)$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\mathcal{W}_2 = \{F_W; F_W \geq h_\alpha\}$	$\Pr\left(F_W^{(alt)} \geq \frac{\sigma_0^2}{\sigma^2} h_\alpha\right)$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\mathcal{W}_3 = \{F_W; F_W \leq d_\alpha\}$	$\Pr\left(F_W^{(alt)} \leq \frac{\sigma_0^2}{\sigma^2} d_\alpha\right)$

Príklad 194 (silofunkcia) *Nakreslite v R do jedného obrázka silofunkcie $1-\beta(\sigma^2)$ pri $\alpha = 0.05$ a $n = 10, 20, 30, 40$ a 50. Rozumne zvoľte podiel σ^2/σ_0^2, ktorý bude na osi x. Obrázky vytvorte pre (a) H_{11}, (b) H_{12} a (c) H_{13}. Riešenie pozri na obrázku 6.6.*

Definícia 56 (p-hodnota F_W testu o σ^2) *Nech F_W je* ***nejaká testovacia štatistika*** *a $F_{obs} = \frac{(n-1)s^2}{\sigma_0^2}$ je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-hodnota} = \begin{cases} 2\min(\Pr(F_W \leq F_{obs}|H_{01}), \Pr(F_W \geq F_{obs}|H_{01})),\ ak & H_{11}: \sigma^2 \neq \sigma_0^2 \\ Pr(F_W \geq F_{obs}|H_{02}),\ ak & H_{12}: \sigma^2 > \sigma_0^2 \\ Pr(F_W \leq F_{obs}|H_{03}),\ ak & H_{13}: \sigma^2 < \sigma_0^2 \end{cases}$$



Obr. 6.6: Silofunkcie asymptotického testu o σ^2 pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 57 (Waldove $100 \times (1-\alpha)\%$ empirické IS pre σ^2) *Waldove $100 \times (1-\alpha)\%$ empirické IS pre σ^2 pre všetky tri typy hypotéz majú nasledovný tvar:*

$$\begin{array}{lll} H_0 & H_1 & \textit{hranice } (d,h) \textit{ pre } 100(1-\alpha)\% \textit{ empirický IS} \\ \sigma^2 = \sigma_0^2 & \sigma^2 \neq \sigma_0^2 & \mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : \sigma_0^2 \in \left(\frac{(n-1)s^2}{H_{\alpha/2}}, \frac{(n-1)s^2}{D_{\alpha/2}}\right)\right\} \\ \sigma^2 \leq \sigma_0^2 & \sigma^2 > \sigma_0^2 & \mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : \sigma_0^2 \in \left(\frac{(n-1)s^2}{H_{\alpha}}, \infty\right)\right\} \\ \sigma^2 \geq \sigma_0^2 & \sigma^2 < \sigma_0^2 & \mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : \sigma_0^2 \in \left(0, \frac{(n-1)s^2}{D_{\alpha}}\right)\right\} \end{array}$$

Príklad 195 (horná hranica IS pre ^{2}) *Nech $X \sim N(\mu, \sigma^2)$, $\mu = 0$ a $\sigma^2 = 4$. Vypočítajte v R hornú hranicu pravostranného (horného) 95% JIS pre σ^2, t.j. $\frac{(n-1)s^2}{\chi^2_{n-1}(1-\alpha)}$, ak $n = 20$ a $s^2 = 4$.*

Riešenie (aj v R)
$\Pr\left(\frac{(n-1)S^2}{\chi^2_{n-1}(1-\alpha)} < \frac{(n-1)s^2}{\chi^2_{n-1}(1-\alpha)}\right) = 1-\alpha$, kde $\chi^2_{n-1}(1-\alpha) = 10.117$, $n = 20$ a $\alpha = 0.05$. Potom $\frac{(n-1)s^2}{\chi^2_{n-1}(1-\alpha)} = 7.512$.

```
n <- 20
alfa <- .05
HH <- (n-1)*4/qchisq(alfa, df=n-1) # 7.512099
```

Príklad 196 (MC odhad koeficientu spoľahlivosti $1-\alpha$) *Vypočítajte v R MC odhad koeficientu spoľahlivosti (pravdepodobnosti pokrytia) pre pravostranný (horný) 95% JIS pre σ^2 pri $M = 1000$ a $n = 20$. Tento JIS je ekvivalentný s testom H_{02} oproti H_{12}. (a) Nech $X \sim N(0,4)$, (b) $X \sim \chi^2(2)$ a (c) $X \sim [pN(0,4) + (1-p)N(0,9)]$, kde $p = 0.9$.*

Riešenie v R
(a)

```
M <- 1000
n <- 20
alfa <- .05
HH <- replicate(M, expr={
          x <- rnorm(n, mean=0, sd=2)
          (n-1) * var(x)/qchisq(alfa, df=n-1)
          })
sum(HH>4) # 953
alfa.hat <- mean(HH>4) # 0.953
SE.alfa.hat <- sqrt(alfa.hat*(1-alfa.hat)/M) # 0.006692608
alfa.hat+c(-1,1)*qnorm(0.975)*SE.alfa.hat # 0.9398827 0.9661173
```

(a) Empirický koeficient spoľahlivosti $1-\widehat{\alpha} = 0.953$ a $\widehat{SE[1-\widehat{\alpha}]} = \sqrt{\widehat{\alpha}(1-\widehat{\alpha})/M} = 0.0067$. 95% DIS pre $1-\alpha$ je rovný $(d,h) = (0.940, 0.966)$. Tento DIS obsahuje skutočnú hodnotu $1-\alpha = 0.95$.
(b) Empirický koeficient spoľahlivosti $1-\widehat{\alpha} = 0.788$ a $\widehat{SE[1-\widehat{\alpha}]} = \sqrt{\widehat{\alpha}(1-\widehat{\alpha})/M} = 0.013$. 95% DIS

pre $1-\alpha$ je rovný $(d,h) = (0.763, 0.813)$. Tento DIS neobsahuje skutočnú hodnotu $1-\alpha = 0.95$.
(c) Empirický koeficient spoľahlivosti $1-\widehat{\alpha} = 0.966$ a $\widehat{SE[1-\widehat{\alpha}]} = \sqrt{\widehat{\alpha}(1-\widehat{\alpha})/M} = 0.0057$. 95% DIS pre $1-\alpha$ je rovný $(d,h) = (0.955, 0.977)$. Tento DIS neobsahuje skutočnú hodnotu $1-\alpha = 0.95$.

Minimálny rozsah n pre nejaký podiel $\frac{\sigma_0^2}{\sigma^2}$ (pri nejakom α a β) vypočítame iteračne pomocou nasledovných rovností:

$$\begin{aligned} \frac{\chi^2_{n-1}(1-\beta)}{\chi^2_{n-1}(\alpha/2)} &\approx \frac{\sigma_0^2}{\sigma^2} \text{ za platnosti } H_{11}; \\ \frac{\chi^2_{n-1}(1-\beta)}{\chi^2_{n-1}(\alpha)} &= \frac{\sigma_0^2}{\sigma^2} \text{ za platnosti } H_{12}; \\ \frac{\chi^2_{n-1}(\beta)}{\chi^2_{n-1}(1-\alpha)} &= \frac{\sigma_0^2}{\sigma^2} \text{ za platnosti } H_{13}. \end{aligned}$$

Pri obojstrannej alternatíve ide o približný výpočet, kedy sa zanedbáva časť sily, ktorá predstavuje obsah pod dolnou kritickou hodnotou.

Príklad 197 (minimálny rozsah súboru) *Vypočítajte v R minimálny rozsah náhodného výberu pre test $H_{03}: \sigma^2 \geq \sigma_0^2$ oproti $H_{13}: \sigma^2 < \sigma_0^2$ pri $\alpha = 0.05$ a $1-\beta = 0.8$, ak podiel $\frac{\sigma_0^2}{\sigma^2}$ je rovný (a) 1.1, (b) 1.5 a (c) 5.*

Test pomerom vierohodnosti pre σ^2.

Majme H_{01} oproti H_{11}. Nech $X \sim N(\mu, \sigma^2)$, kde $\boldsymbol{\theta} = (\mu, \sigma^2)^T$. Funkcia vierohodnosti je rovná

$$L(\boldsymbol{\theta}|\mathbf{x}) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2\right).$$

MLE $\boldsymbol{\theta}$ je rovný

$$\widehat{\boldsymbol{\theta}} = (\overline{x}, \widehat{\sigma}^2)^T = \left(\frac{1}{n}\sum_{i-1}^n x_i, \frac{1}{n}\sum_{i=1}^n (x_i-\overline{x})^2\right),$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \sigma^2 \neq \sigma_0^2\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\overline{x}, \sigma_0^2)^T$, t.j. $\Theta_0 = \{\boldsymbol{\theta} : \sigma^2 = \sigma_0^2\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{x})) &= l\left(\widehat{\boldsymbol{\theta}}|\mathbf{x}\right) - l\left(\boldsymbol{\theta}_0|\mathbf{x}\right) \\ &= \left(-\frac{n}{2}(\ln(2\pi) + \ln\widehat{\sigma}^2 + \frac{\widehat{\sigma}^2}{\widehat{\sigma}^2})\right) - \left(-\frac{n}{2}(\ln(2\pi) + \ln\sigma_0^2 + \frac{\widehat{\sigma}^2}{\sigma_0^2})\right) \\ &= \frac{n}{2}\left(\frac{\widehat{\sigma}^2}{\sigma_0^2} - 1 - \ln(\frac{\widehat{\sigma}^2}{\sigma_0^2})\right). \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln(\lambda(\mathbf{X})) \overset{\mathcal{D}}{\sim} \chi_1^2$, kde H_0 bude zamietnutá pre malé alebo veľké hodnoty podielu $\frac{\widehat{\sigma}^2}{\sigma_0^2}$.

Dá sa ukázať, že podiel $\frac{\widehat{\sigma}^2}{\sigma_0^2}$ je rastúcou funkciou $F_{\mathrm{obs}} = \frac{n\widehat{\sigma}^2}{\sigma_0^2}$ a potom

$$u_{\mathrm{LR}} = -2\ln(\lambda(\mathbf{x})) \approx F_{\mathrm{obs}} - n\left(1 + \ln\left(\frac{F_{\mathrm{obs}}}{n}\right)\right).$$

Vierohodnostný $100 \times (1-\alpha)\%$ **empirický DIS pre** σ^2 bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : U_{\mathrm{LR}} < \chi_1^2(\alpha)\right\}.$$

Majme H_{02} oproti H_{12}. Potom $u_{\mathrm{LR}} = 0$ ak $\frac{\widehat{\sigma}^2}{\sigma_0^2} \leq 1$ a $u_{\mathrm{LR}} \approx F_{\mathrm{obs}} - n(1 + \ln\left(\frac{F_{\mathrm{obs}}}{n}\right)$ pre $\frac{\widehat{\sigma}^2}{\sigma_0^2} > 1$.
Majme H_{03} oproti H_{13}. Potom $u_{\mathrm{LR}} = 0$ ak $\frac{\widehat{\sigma}^2}{\sigma_0^2} \geq 1$ a $u_{\mathrm{LR}} \approx F_{\mathrm{obs}} - n(1 + \ln\left(\frac{F_{\mathrm{obs}}}{n}\right)$ pre $\frac{\widehat{\sigma}^2}{\sigma_0^2} < 1$.

Príklad 198 (test rozptylu σ^2) *Majme dáta* one-sample-variance-skull-mf.txt *a premennú dĺžku lebky* skull.L *v mm starovekej egyptskej mužskej populácie, o ktorej predpokladáme, že má normálne rozdelenie $N(\mu, \sigma^2)$. (a) Otestujte hypotézu o zhode rozptylu dĺžky lebky tejto populácie s rozptylom dĺžky lebky novovekej egyptskej mužskej populácie 7.526^2 na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre rozptyl dĺžky lebky, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku F_W, (2) testovaciu štatistiku pomeru vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: rozptyl dĺžky lebky starovekej egyptskej mužskej populácie je zhodný s rozptylom dĺžky lebky novovekej egyptskej mužskej populácie oproti H_1: rozptyl dĺžky lebky starovekej egyptskej mužskej populácie nie je zhodný s rozptylom dĺžky lebky novovekej egyptskej mužskej populácie.
   - **matematická formulácia** – $H_0 : \sigma^2 = \sigma_0^2$ oproti $H_1 : \sigma^2 \neq \sigma_0^2$, kde $\sigma_0 = 7.526^2$.

2. **Testovacia štatistika** – Rozptyl σ^2 nepoznáme, a preto ho musíme odhadnúť. Jeho odhadom je $s^2 = 40.582$ (smerodajná odchýlka $s = 6.370$).

```
DATA <- read.table("one-sample-variance-skull-mf.txt", header=TRUE)
attach(DATA)
x <- na.omit(skull.L[sex=="m"])
n <- length(x) # 217
rozptyl <- var(x) # 40.58197
smerodch <- sd(x) # 6.370398
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $F_{\mathrm{obs}} = \frac{(n-1)s^2}{\sigma_0^2} \doteq \frac{217 \times 6.370^2}{7.526^2} \doteq 154.76$.
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} = F_{\mathrm{obs}} - n\left(1 + \ln(\frac{F_{\mathrm{obs}}}{n})\right)$, kde $F_{\mathrm{obs}} = \frac{ns^2}{\sigma_0^2}$, t.j. keďže $F_{\mathrm{obs}} \doteq \frac{217 \times 6.370^2}{7.526^2} \doteq 155.476$, potom $u_{\mathrm{LR}} \doteq 155.476 - 217(1 + \ln(\frac{155.476}{n})) \doteq 10.825$.

```
sigma0.sq <- 7.526^2 # 56.64068
F.obs <- (n-1)*rozptyl/sigma0.sq # 154.7599
Fstar.obs <- n*rozptyl/sigma0.sq # 155.4764
uLR.obs <- Fstar.obs-n*(1+log(Fstar.obs/n)) # 10.82495
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $\chi^2(1-\alpha) \doteq 177.189$ a $\chi^2(\alpha) \doteq 258.597$;
   kritický obor $\mathcal{W}_1 = (-\infty, -1.971) \cup (1.971, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi_1^2(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (3.841, \infty)$.

```
chisq.krit.d <- qchisq(0.025, df=n-1) # 177.189
chisq.krit.h <- qchisq(0.975, df=n-1) # 258.5971
lr.krit.hodn <- qchisq(0.95, df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100 \times (1-\alpha)\%$ empirický DIS = 95% empirický DIS pre σ^2:
   $(d, h) = \left(\frac{(n-1)s^2}{\chi_{n-1}^2(\alpha/2)}, \frac{(n-1)s^2}{\chi_{n-1}^2(1-\alpha/2)}\right) \doteq (33.897, 49.471)$.
   Vierohodnostný 95% empirický DIS pre σ^2:
   $\mathcal{CS}_{0.95} = \left\{\sigma_0^2 : u_{\mathrm{LR}}(\sigma_0^2) < \chi_1^2(\alpha)\right\} \doteq (33.815, 49.280)$.

```
DH.W <- (n-1)*rozptyl/chisq.krit.h # 33.89715
HH.W <- (n-1)*rozptyl/chisq.krit.d # 49.47094
min.sigma.sq <- 0.01
max.sigma.sq <- 2
podiel.i <- seq(min.sigma.sq,max.sigma.sq,by=0.0001)
sigma0.sq.i <- rozptyl/podiel.i
Fstar.obs.i <- n*podiel.i
uLR.i <- Fstar.obs.i-n*(1+log(Fstar.obs.i/n))
IS.LR <- range(sigma0.sq.i[which(uLR.i<lr.krit.hodn)]) # 33.81549 49.27986
```

5. **Štatistický záver** –
Waldov test: p-hodnota $\doteq 2\Pr(F \geq F_{\text{obs}}|H_0) = 2\Pr(F \geq 154.760|H_0) \doteq 0.001$.
Test pomerom vierohodnosti: p-hodnota $\doteq \Pr(U_{\text{LR}} \geq 10.825|H_0) \doteq 0.001$.

```
p.hodn.W <- 2*(pchisq(F.obs, df=n-1)) # 0.00115558
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # 0.001001415
```

H_0 na hladine významnosti zamietame, pretože (1) testovacia štatistika patrí kritickému oboru, (2) $\sigma_0^2 = 56.641$ nepatrí DIS a nakoniec (3) p-hodnota je menšia ako 0.05.

6. **Slovný záver** – Zamietame nulovú hypotézu o tom, že rozptyl dĺžky lebky starovekej egyptskej mužskej populácie je zhodný s rozptylom dĺžky lebky novovekej egyptskej mužskej populácie.

7. **Antropologický záver** – Hodnota rozptylu dĺžky lebky novovekej egyptskej populácie je štatisticky významne vyššia než v starovej egyptskej populácii. Rozdiely v rozptyloch môžu byť dôsledkom skutočných biologických rozdielov vo variabilite v dĺžke lebky v dôsledku odlišného zloženia populácie v priebehu času – ak staroveká populácia bola celkovo menšia a v menších skupinách sú si jedinci podobnejší, tak variabilita je nižšia. Nulová hypotéza však môže byť zamietnutá aj z dôvodu odlišného vzorkovania a v prípade kostrových sérií aj z dôvodu obmedzenej reprezentatívnosti a celistvosti dát a ďalších okolností, ktoré sa spájajú s výskumom minulých populácií z kostrových pozostatkov.

Príklad 199 (test rozptylu σ^2) *Majme dáta* one-sample-variance-skull-mf.txt *a premennú šírku lebky* skull.B *v mm starovekej egyptskej mužskej populácie, o ktorej predpokladáme, že má normálne rozdelenie $N(\mu, \sigma^2)$. (a) Otestujte hypotézu o zhode rozptylu šírky lebky tejto populácie s rozptylom šírky lebky novovekej egyptskej mužskej populácie 6.411^2 na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre rozptyl šírky lebky, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku F_W, (2) testovaciu štatistiku pomeru vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2).*

6.3 Asymptotické testy o korelačnom koeficiente

Korelačná analýza slúži na hodnotenie sily závislosti medzi dvoma spojitými premennými. Korelačný koeficient vyjadruje mieru asociácie. Vyjadruje však iba mieru lineárneho vzťahu, a pokiaľ je závislosť medzi premennými nelineárna, nemusí vzťah popisovať správne alebo ho neodhalí vôbec. V biologickej antropológii existuje celý rad situácií, v ktorých je potrebné zistiť existenciu, mieru (veľkosť) a smer (kladná, záporná) lineárnej závislosti medzi dvoma premennými. Klasickou situáciou je sledovanie vzťahu celku a jeho častí (napr. zistenie vzťahu výšky postavy a dĺžky končatiny), hľadanie väzby medzi faktorom považovaným za efektor metabolického alebo ontogenetického procesu a jeho produktom/dopadom (hladina prenatálneho testosterónu a miera agresivity v dospelosti) alebo hľadanie súvislostí medzi nejakým faktorom prostredia a veľkosťou, tvarom a zložením ľudského tela (napr. medzi energetickým obsahom stravy a indexom telesnej hmotnosti (BMI)). Korelácia hodnoteného znaku medzi dvojčatami je tiež základom výpočtu heritability (dedivosti). Korelačné koeficienty týchto súvislostí sú často publikovanými ukazovateľmi a možno ich využiť na štatistické porovnanie publikovaných výsledkov s koreláciami nášho výskumu. Pri použití korelačného koeficientu je potrebné pamätať na to, že vyjadruje mieru štatistickej závislosti a nie mieru biologickej príčinnej (kauzálnej) závislosti. Bez ďalších údajov nie je možné z korelačného koeficientu vypočítaného z dát získaných pozorovaním (neexperimentálne) zistiť mechanizmus, či spolu korelované znaky priamo súvisia, alebo ich korelácia vzniká pod vplyvom neznámej vlastnosti tretej (mätúcej premennej), ktorou často býva nejaká generalizovaná, všetko ovplyvňujúca vlastnosť, napr. veľkosť tela. Pokiaľ spolu vlastnosti priamo súvisia, nie je možné len na základe korelačného koeficientu zistiť, v akom zmysle, t.j. ktorý znak je príčina a ktorý dôsledok. Tiež treba mať na pamäti, že celý rad príčinných biologických väzieb a procesov v ľudskom tele prebieha nelineárne. Napríklad procesy vývoja a rastu ľudského tela sú zložitým reťazcom mnohých odlišne prebiehajúcich fáz a sledovanie súvislostí dvoch znakov v priebehu rastu pomocou lineárneho koeficientu korelácie môže poskytnúť falošnú predstavu o miere závislosti.

Korelačný koeficient $\rho \in \langle -1, 1\rangle$ meria silu nekauzálneho lineárneho vzťahu (asociácie, závislosti) dvoch spojitých premenných X a Y pochádzajúcich z dvojrozmerného normálneho rozdelenia, t.j. $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde tieto predstavujú (1) dve rôzne premenné, (2) párovú pemennú alebo (3) jednu premennú z dvoch rôznych časových bodov. Pri interpretácii tohoto vzťahu je dôležité zamerať sa nielen na veľkosť ale aj znamienko ρ, kde

- 0 predstavuje nulovú (žiadnu) asociáciu medzi dvoma premennými,
- ± 1 predstavuje perfektnú asociáciu medzi dvoma premennými,
- kladné číslo predstavuje pozitívnu asociáciu (keď rastú hodnoty jednej premennej, rastú aj hodnoty druhej premennej),
- záporné číslo predstavuje negatívnu asociáciu (keď rastú hodnoty jednej premennej, hodnoty druhej premennej klesajú).

V prírodných vedách sa zaužívala nasledovná interpretácia $|\rho|$:

- $|\rho| \in \langle 0, 0.4)$ – žiadna alebo takmer žiadna asociácia,
- $|\rho| \in \langle 0.4, 0.6)$ – slabá asociácia,
- $|\rho| \in \langle 0.6, 0.8)$ – mierna asociácia,
- $|\rho| \in \langle 0.8, 1\rangle$ – silná asociácia.

V sociálnych vedách sa zaužívala nasledovná interpretácia $|\rho|$:

- $|\rho| \in \langle 0, 0.15)$ – žiadna alebo takmer žiadna asociácia,
- $|\rho| \in \langle 0.15, 0.3)$ – slabá asociácia,
- $|\rho| \in \langle 0.3, 0.5)$ – mierna asociácia,
- $|\rho| \in \langle 0.5, 0.6)$ – dosť (celkom) silná asociácia,
- $|\rho| \in \langle 0.6, 0.8)$ – silná asociácia,
- $|\rho| \in \langle 0.8, 1\rangle$ – veľmi silná asociácia.

Označme odhad ρ ako $\widehat{\rho} = r$. Potom rozlišujeme tri typy korelačných koeficientov $r = r(x, y)$ podľa typu premenných (Mardia, 1976; Mardia a Jupp, 2000; Jammalamadaka a Lund, 2006)

1. **klasický/lineárny Pearsonov korelačný koeficient** (**lineárna–lineárna premenná**), kde

$$r = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{(n-1)s_x s_y} = \frac{1}{n-1}\sum_{i=1}^n \left(\frac{x_i - \overline{x}}{s_x}\right)\left(\frac{y_i - \overline{y}}{s_y}\right),$$

kde $s_x^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$ a $s_y^2 = \frac{1}{n-1}\sum_{i=1}^n (y_i - \overline{y})^2$.

2. **lineárno-uhlový Pearsonov korelačný koeficient** (**lineárna–uhlová premenná**), kde

$$r = \sqrt{\frac{r_{xc}^2 + r_{xs}^2 - 2r_{xc}r_{xs}r_{cs}}{1 - r_{cs}^2}},$$

x je lineárna a y uhlová premenná meraná v radiánoch (1 rad$\doteq \frac{180}{\pi} = 57.3$ stupňa; uhol v radiánoch = uhol v stupňoch $\times\ \pi/180$), $r_{xc} = r(x, \cos y)$, $r_{xs} = r(x, \sin y)$ a $r_{sc} = r(\cos y, \sin y)$;

3. **uhlový Pearsonov korelačný koeficient** (**uhlová−uhlová premenná**), kde

$$r = \frac{\sum_{i=1}^{n} \sin(x_i - \overline{x}) \sin(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} \sin^2(x_i - \overline{x}) \sin^2(y_i - \overline{y})}},$$

x a y sú uhlové premenné merané v radiánoch[1], kvadrant špecifický priemerný uhol

$$\overline{x} = \begin{cases} \arctan\left(\frac{\sum_{i=1}^{n} \sin x_i}{\sum_{i=1}^{n} \cos x_i}\right), & \text{ak } \sum_{i=1}^{n} \cos x_i \geq 0, \\ \arctan\left(\frac{\sum_{i=1}^{n} \sin x_i}{\sum_{i=1}^{n} \cos x_i}\right) + \pi, & \text{ak } \sum_{i=1}^{n} \cos x_i < 0, \\ \text{nedefinované}, & \text{ak } \sqrt{(\sum_{i=1}^{n} \sin x_i)^2 + (\sum_{i=1}^{n} \cos x_i)^2} = 0. \end{cases}$$

Podobne definujeme aj $\overline{y}$.

Korelačný koeficient má asymptoticky normálne rozdelenie, t.j. $R \sim N\left(\rho, \left(1-\rho^2\right)^2 / (n-1)\right)$. Konvergencia k normalite je však pomalá, preto sa používa **Fisherova Z-premenná** (rýchlejšie konvergujúca k normalite)

$$Z_R = \frac{1}{2} \ln \frac{1+R}{1-R}, Z_R \overset{\mathcal{D}}{\sim} N\left(\frac{1}{2} \ln \frac{1+\rho}{1-\rho}, \frac{1}{n-3}\right),$$

kde $n > 3$. Odporúča sa uvažovať o Z_R ako o normálnej premennej už pre $n \geq 10$, ak ρ nie je blízko ± 1.

Majme dvojice hypotéz $H_{01} : \rho = \rho_0$ oproti $H_{11} : \rho \neq \rho_0$, $H_{02} : \rho \leq \rho_0$ oproti $H_{12} : \rho > \rho_0$, $H_{03} : \rho \geq \rho_0$ oproti $H_{13} : \rho < \rho_0$, ktoré chceme testovať.

Ak H_0 platí, potom

$$Z_W = \sqrt{n-3}\,(Z_R - \xi_0) \overset{\mathcal{D}}{\sim} N(0,1),$$

kde $\xi_0 = \frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0}$. Z_W sa nazýva **Waldova testovacia štatistika** a test **jednovýberový Z-test o korelačnom koeficiente** ρ.

Definícia 58 (Kritický obor a silofunkcia Z_W testu o ρ) *Kritický obor a silofunkcia (za platnosti $\xi = \frac{1}{2} \ln \frac{1+\rho}{1-\rho}$) sú definované nasledovne (pozri obrázok 6.7):*

H_0	H_1	$\mathcal{W}$	$1-\beta(\xi)$
$\rho = \rho_0$	$\rho \neq \rho_0$	$\mathcal{W}_1 = \left\{Z_W; \lvert Z_W\rvert \geq u_{\alpha/2}\right\}$	$1 - \Phi\left(u_{\alpha/2} - \sqrt{n-3}\,\lvert\xi_0 - \xi\rvert\right)$
$\rho \leq \rho_0$	$\rho > \rho_0$	$\mathcal{W}_2 = \{Z_W; Z_W \geq u_\alpha\}$	$1 - \Phi\left(u_\alpha + \sqrt{n-3}\,(\xi_0 - \xi)\right)$
$\rho \geq \rho_0$	$\rho < \rho_0$	$\mathcal{W}_3 = \{Z_W; Z_W \leq -u_\alpha\}$	$1 - \Phi\left(u_\alpha - \sqrt{n-3}\,(\xi_0 - \xi)\right)$

Definícia 59 (p-hodnota Z_W testu o ρ) *Nech Z_W je* ***nejaká testovacia štatistika*** *a z_W je jej realizácia (****pozorovaná, vypočítaná testovacia štatistika****), potom p-hodnotu počítame nasledovne:*

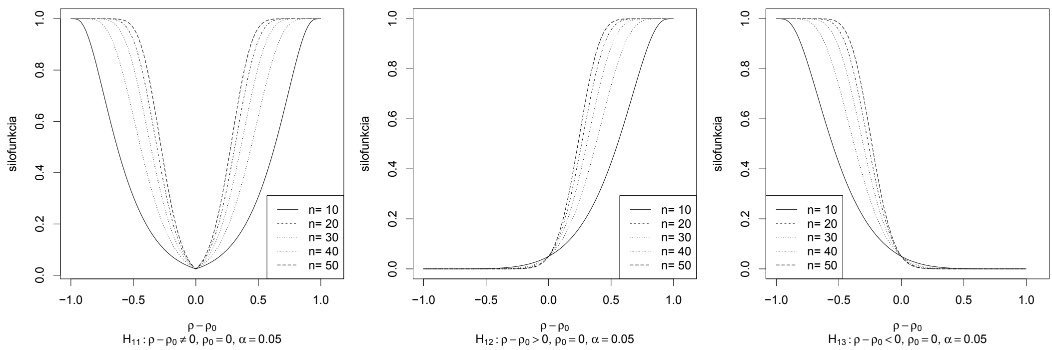
$$p\text{-hodnota} = \begin{cases} 2Pr(Z_W \geq |z_W| | H_{01}), \text{ ak } & H_{11} : \rho \neq \rho_0 \\ Pr(Z_W \geq z_W | H_{02}), \text{ ak } & H_{12} : \rho > \rho_0 \\ Pr(Z_W \leq z_W | H_{03}), \text{ ak } & H_{13} : \rho < \rho_0 \end{cases}$$

Definícia 60 (Waldov $100 \times (1-\alpha)\%$ empirický DIS pre ξ a IS pre ρ) *Keďže* $\Pr\left(\sqrt{n-3}\,|Z_R - \xi| \leq u(\alpha/2)\right) = 1 - \alpha$, *potom Waldov $100 \times (1-\alpha)\%$ empirický DIS pre ξ má tvar*

$$\mathcal{CS}_{1-\alpha} = \left\{\xi_0; \xi_0 \in \left(z_R - \frac{u_{\alpha/2}}{\sqrt{n-3}}, z_R + \frac{u_{\alpha/2}}{\sqrt{n-3}}\right)\right\}.$$

Vieme, že $\rho = \tanh(\xi)$ je monotónna funkcia ξ. Potom môžeme písať spätne transformované Waldove $100 \times (1-\alpha)\%$ empirické IS pre ρ ako:

[1] Rozptyl uhlovej premennej definujeme ako $1 - \frac{1}{n}\sqrt{(\sum_{i=1}^{n} \sin x_i)^2 + (\sum_{i=1}^{n} \cos x_i)^2} \in \langle 0, 1 \rangle$.



Obr. 6.7: Silofunkcie asymptotického testu o ρ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

H_0	H_1	*hranice* (d,h) *pre* $100(1-\alpha)\%$ *empirický IS*
$\rho = \rho_0$	$\rho \neq \rho_0$	$\mathcal{CS}_{1-\alpha} = \left\{\rho_0 : \rho_0 \in \left(\tanh\left[z_R - \frac{u_{\alpha/2}}{\sqrt{n-3}}\right], \tanh\left[z_R + \frac{u_{\alpha/2}}{\sqrt{n-3}}\right]\right)\right.$
$\rho \leq \rho_0$	$\rho > \rho_0$	$\mathcal{CS}_{1-\alpha} = \left\{\rho_0 : \rho_0 \in \left(\tanh\left[z_R - \frac{u_{\alpha}}{\sqrt{n-3}}\right], 1\right)\right\}$
$\rho \geq \rho_0$	$\rho < \rho_0$	$\mathcal{CS}_{1-\alpha} = \left\{\rho_0 : \rho_0 \in \left(-1, \tanh\left[z_R + \frac{u_{\alpha/2}}{\sqrt{n-3}}\right]\right)\right\}$

Príklad 200 (DIS pre ρ použitím Fisherovej Z-premennej) *Naprogramujte DIS pre ρ.*

Riešenie v R

```
"ISkor" <- function(x,y,conf.level=0.95){
z <- atanh(cor(x,y))
n <- length(x)
a <- qnorm(1-(1-conf.level)/2)/(n-3)^0.5
IS.xi <- c(z-a,z+a)
IS <- tanh(IS.xi)
return(IS)
}
```

Majme H_{11}. Použitím Fisherovej Z-premennej pre nejaký rozdiel $\xi - \xi_0$, α a β bude **minimálny rozsah** n definovaný ako

$$n \geq 3 + \left(\frac{u_{\alpha/2} + u_\beta}{\xi - \xi_0}\right)^2.$$

Pre H_{12} a H_{13} zameníme $u_{\alpha/2}$ za u_α.

Príklad 201 (minimálny rozsah n) *Vypočítajte minimálny rozsah n pre $\rho = 0.1, 0.2, \ldots, 0.9$, $\rho_0 = 0$ pri $\alpha = 0.05$, $1-\beta = 0.8$ a obojstrannej alternatíve H_{11}.*

Riešenie (pozri tabuľku 6.1)

Tabuľka 6.1: Minimálne rozsahy n pri vybraných rozdieloch ρ a ρ_0, kde $\rho_0 = 0$

ρ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
n	783	194	85	47	30	20	14	10	7

Príklad 202 (minimálny rozsah n) *Vypočítajte minimálny rozsah n pre $\rho = 0.1, 0.2, \ldots, 0.9$, ρ_0 vždy o 0.1 menšie ako ρ, pri $\alpha = 0.05$, $\beta = 0.8$ a obojstrannej alternatíve H_{11}.*

Tabuľka 6.2: Minimálne rozsahy n pri vybraných rozdieloch ρ a ρ_0 spolu s rozdielom $z_R - \xi_0$, ktorý je funkciou ρ a ρ_0

ρ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
ρ_0	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$z_R - \xi_0$	0.1003	0.1024	0.1068	0.1141	0.1257	0.1438	0.1742	0.2313	0.3736
n	783	752	692	606	501	383	262	150	60

Riešenie (pozri tabuľku 6.2)

Jednovýberový t-test o korelačnom koeficiente ρ v R (funkcia **cor.test()**).
Argumenty (vstupy) funkcie:

1. dáta x a y;
2. spoľahlivosť conf.level=p, prednastavené je conf.level=.95;
3. metóda method="pearson" (prednastavené);
4. formulácia alternatívy alternative="two.sided" je prednastavené, ďalšie voľby sú "greater", "less".

Výstupy funkcie:

1. názov použitého testu method;
2. testovacia štatistika statistic;
3. stupne voľnosti df parameter;
4. p-hodnota p.value;
5. alternatívna hypotéza alternative hypothesis;
6. interval spoľahlivosti conf.int (vypočítaný na základe Fisherovej Z-premennej);
7. bodový odhad (korelačný koeficient) estimate.

Test pomerom vierohodnosti pre ρ.

Majme H_{01} oproti H_{11}. Nech $(X,Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)^T$. Logaritmus funkcie vierohodnosti bude mať tvar

$$\begin{aligned} l(\boldsymbol{\theta}|\mathbf{x},\mathbf{y}) &= -n\ln\left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right) \\ &- \frac{1}{2(1-\rho^2)}\left(\frac{\sum_{i=1}^n (x_i-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{\sum_{i=1}^n (x_i-\mu_1)(y_i-\mu_2)}{\sigma_1\sigma_2} + \frac{\sum_{i=1}^n (y_i-\mu_2)^2}{\sigma_2^2}\right). \end{aligned}$$

MLE $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2, \widehat{\rho})^T$, kde

$$\widehat{\mu}_1 = \overline{x}, \widehat{\mu}_2 = \overline{y}, \widehat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^n (x_i-\overline{x})^2, \widehat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^n (y_i-\overline{y})^2, \widehat{\rho} = \frac{\sum_{i=1}^n (x_i-\overline{x})(y_i-\overline{y})}{n\widehat{\sigma}_1\widehat{\sigma}_2}.$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \rho \neq 0\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2, 0)^T$, t.j. $\Theta_0 = \{\boldsymbol{\theta} : \rho = 0\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{x},\mathbf{y})) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x},\mathbf{y}) - l(\boldsymbol{\theta}_0|\mathbf{x},\mathbf{y}) \\ &= \left(-n\ln(2\pi\widehat{\sigma}_1\widehat{\sigma}_2) - \frac{n}{2}\ln(1-\widehat{\rho}^2) - n\right) - \left(-n\ln(2\pi\widehat{\sigma}_1\widehat{\sigma}_2) - n\right) \\ &= -\frac{n}{2}\ln(1-\widehat{\rho}^2). \end{aligned}$$

Z vyššie uvedeného je vidieť, že $-\ln(\lambda(\mathbf{x},\mathbf{y}))$ rastúcou funkciou $\widehat{\rho}^2$. Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{X},\mathbf{Y}\right)\right) \overset{\mathcal{D}}{\sim} \chi^2_1$, kde H_0 bude zamietnutá pre veľké hodnoty $|\widehat{\rho}|$.

Keďže platí

$$T_W = \frac{\sqrt{n-2}R}{\sqrt{1-R^2}} \overset{\mathcal{D}}{\sim} t_{n-2},$$

môžeme zapísať u_{LR} pomocou t_W. $|t_W|$ je rastúcou funkciou $|\widehat{\rho}|$, preto

$$u_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{x},\mathbf{y}\right)\right) = -n\ln(1-\widehat{\rho}^2) = -n\ln\left(1-\frac{t_W^2}{t_W^2+(n-2)}\right).$$

T_W sa nazýva **Waldova testovacia štatistika** a test **jednovýberový t-test o korelačnom koeficiente** ρ. Ak H_0 neplatí, rozdelenie T_W ako aj $\widehat{\rho}$, neexistuje v jednoduchej forme, a preto sa používa normálna aproximácia pomocou Fisherovej Z-premennej spomenutá vyššie. Tiež je dôležité upozorniť na fakt, že ak sa ρ výrazne líši od 0, konverguje jeho rozdelenie k normálnemu rozdeleniu veľmi pomaly. Ďalšou nevýhodou je, že rozptyl ρ zavisí na hodnote ρ (rozptyl ρ je funkciou ρ).

Ak $\rho_0 \neq 0$, T_W nie je možné použiť. Rovnako nie je možné použiť U_{LR} na vytvorenie vierohodnostného IS. Preto musíme vypočítať U_{LR} pre nejaké ρ_0. Nech $\boldsymbol{\theta}_0 = (\widehat{\mu}_1, \widehat{\mu}_2, \widetilde{\sigma}^2_1, \widetilde{\sigma}^2_2, \rho_0)^T$, kde $\widetilde{\sigma}^2_j = \frac{\widehat{\sigma}^2_j(1-\rho_0\widehat{\rho})}{1-\rho_0^2}$, $j=1,2$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$\begin{aligned}
-\ln(\lambda(\mathbf{x},\mathbf{y})) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x},\mathbf{y}) - l\left(\boldsymbol{\theta}_0|\mathbf{x},\mathbf{y}\right) = -\frac{n}{2}\ln(1-\widehat{\rho}^2) - \frac{n}{2}\ln(\widehat{\sigma}^2_1\widehat{\sigma}^2_2) + \frac{n}{2}\ln(1-\rho_0^2) + \frac{n}{2}\ln(\widetilde{\sigma}^2_1\widetilde{\sigma}^2_2) \\
&= -\frac{n}{2}\left(\ln(1-\widehat{\rho}^2) + \ln(\widehat{\sigma}^2_1\widehat{\sigma}^2_2) - \ln\left(\widehat{\sigma}^2_1\widehat{\sigma}^2_2\left(\frac{1-\rho_0\widehat{\rho}}{1-\rho_0^2}\right)^2\right) - \ln(1-\widehat{\rho}^2_0)\right)
\end{aligned}$$

a

$$u_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{x},\mathbf{y}\right)\right) = n\left(\ln\frac{(1-\rho_0\widehat{\rho})^2}{(1-\rho_0^2)(1-\widehat{\rho}^2)}\right).$$

Vierohodnostný $100\times(1-\alpha)\%$ **empirický DIS pre** ρ bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{\rho_0 : U_{\mathrm{LR}} < \chi^2_1\left(\alpha\right)\right\}.$$

Príklad 203 (konvergencia ρ a ξ k normálnemu rozdeleniu) *Urobte v R simuláciu pseudonáhodných čísel z $N_2\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$, kde $\mu_1=0, \mu_2=0$, $\sigma_1=1, \sigma_2=1$ (pozri príklad 87), kde $n=5,10,20,50$ a 100, $M=10000$. Použite (a) $\rho=0$, (b) $\rho=0.50$ a (c) $\rho=0.9$. Pre každé $m=1,2,\ldots,M$, vypočítajte Pearsonov korelačný koeficient r_m a Fisherovu Z-premennú $z_{R,m}$. Zobrazte histogramy simulovaných r_m a $z_{R,m}$ a superponujte ich teoretickými hustotami prislúchajúcich normálnych rozdelení.*

Príklad 204 (test o korelačnom koeficiente ρ) *Majme dáta* one-sample-correlation-skull-mf.txt *a premenné najväčšia výška mozgovne* skull.pH *a morfologická výška tváre* face.H *(obe v mm) starovekej egyptskej mužskej populácie, o ktorých predpokladáme, že majú dvojrozmerné normálne rozdelenie $N_2(\boldsymbol{\mu},\boldsymbol{\Sigma})$. (a) Otestujte hypotézu o tom, že korelačný koeficient najväčšej výšky mozgovne a morfologickej výšky tváre je rovný 0.251. (b) Vypočítajte $100\times(1-\alpha)\%$ empirický DIS pre korelačný koeficient, kde koeficient spoľahlivosti $1-\alpha=0.95$. Použite (1) Waldovu testovaciu štatistiku Z_W, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2). Výsledky Waldovho testu skontrolujte pomocou funkcie* cor.test()*. Porovnajte empirický DIS vypočítaný funkciou* cor.test() *s výsledkom vypočítaným funkciou* lScor()*.*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: korelačný koeficient najväčšej výšky mozgovne a morfologickej výšky tváre v starovekej egyptskej mužskej populácii je rovný korelačnému koeficientu 0.251 týchto dvoch premenných v novovekej egyptskej mužskej populácii oproti H_1: korelačný koeficient najväčšej výšky mozgovne a morfologickej výšky tváre v starovekej egyptskej mužskej populácii nie je rovný korelačnému koeficientu 0.251 týchto dvoch premenných v novovekej egyptskej mužskej populácii.
   - **matematická formulácia** – $H_0 : \rho = \rho_0$ oproti $H_1 : \rho \neq \rho_0$, kde $\rho_0 = 0.251$.

2. **Testovacia štatistika** – Korelačný koeficient ρ je neznámy, preto ho musíme odhadnúť. Jeho odhadom je $\widehat{\rho} = 0.331$.

```
DATA <- read.table("one-sample-correlation-skull.txt",header=TRUE)
DATA <- na.omit(DATA)
attach(DATA)
n <- dim(DATA)[1] # 164
rho.hat <- cor(skull.pH,face.H) # 0.3306431
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $z_W = \sqrt{n-3}\,(z_R - \xi_0) \doteq 1.105$.
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} \doteq 4.459$ (ak $H_0 : \rho = 0$).
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} \doteq 1.242$ (ak $H_0 : \rho = \rho_0$).

```
zR <- 1/2*log((1+rho.hat)/(1-rho.hat)) # 0.3435501
rho0 <- 0.251
xi0 <- 1/2*log((1+rho0)/(1-rho0)) # 0.2564798
zW.obs <- sqrt(n-3)*(zR-xi0) # 1.104799
tW.obs <- sqrt(n-2)*rho.hat/(sqrt(1-rho.hat^2)) # 4.459204
uLR.obs <- -n*log(1-tW.obs^2/(tW.obs^2+(n-2))) # 18.98718
 # tW.obs a ani uLR.obs nie je mozne pouzit, lebo rho0 sa nerovna 0
uLR.obs <- -n*(log((1-rho0^2)*(1-rho.hat^2)/(1-rho0*rho.hat)^2)) # 1.241757
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $u(1-\alpha/2) \doteq -1.960$ a $u(\alpha/2) \doteq 1.960$;
   kritický obor $\mathcal{W}_1 = (z_{\min}, u(1-\alpha/2)) \cup (u(\alpha/2), z_{\max}) \doteq (-\infty, -1.960) \cup (1.960, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi^2_1(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi^2_1(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
z.krit.d <- qnorm(0.025) # -1.959964
z.krit.h <- qnorm(0.975) # 1.959964
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100(1-\alpha)\%$ empirický DIS = 95% empirický DIS pre ρ:
   $(d, h) \doteq (0.187, 0.461)$.
   Vierohodnostný 95% empirický DIS pre ρ:
   $\mathcal{CS}_{0.95} = \left\{\rho_0 : u_{\mathrm{LR}}(\rho_0) < \chi^2_1(\alpha)\right\} \doteq (0.188, 0.460)$.

```
ISz <- IScor(skull.pH,face.H) # 0.1868617 0.4605561
rho.0i <- seq(-0.999,0.999,by=0.0001)
uLR.i <- -n*(log((1-rho0i^2)*(1-rho.hat^2)/(1-rho0i*rho.hat)^2))
IS.LR <- range(rho.0i[which(uLR.i<lr.krit.hodn)]) # 0.1880 0.4596
```

5. **Štatistický záver** –
   Waldov test: p-hodnota $\doteq 2\Pr(Z_W \geq |10.334||H_0) \doteq 0.269$.
   Test pomerom vierohodnosti: p-hodnota $= \Pr(U_{\mathrm{LR}} \geq 1.242|H_0) \doteq 0.265$.

```
p.hodn.zW <- 2*(1-pnorm(abs(zW.obs))) # 0.2692467
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # 0.2651328
```

   H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\rho_0 = 0.251$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že korelačný koeficient najväčšej výšky mozgovne a morfologickej výšky tváre v starovekej egyptskej mužskej populácii je rovný korelačnému koeficientu 0.251 týchto dvoch premenných v novovekej egyptskej mužskej populácii.

7. **Antropologický záver** – Vo vzťahu týchto dvoch premenných sa staroveká a novoveká egyptská populácia nelíšia. Ani v jednej z nich sledované rozmery mozgovej a tvárovej časti lebky spolu nesúvisia, čo naznačuje, že ide o samostatné vývinové moduly riadené počas ontogenézy inými faktormi.

Riešenie pomocou funkcie cor.test() – touto funkciou nie je možné testovať inú H_0 ako $\rho = 0$ (pri všetkých troch alternatívach). Môžeme ale pomocou nej vypočítať empirický DIS pre ρ, ktorý je identický s nami vypočítaným DIS (Waldov princíp).

```
Ttest <- cor.test(skull.pH, face.H)
Ttest$conf.int # 0.1868617 0.4605561
```

Príklad 205 (test o lineárno-uhlovom ρ) *Naprogramujte v R funkcie na výpočet (a) odhadu lineárno-uhlového korelačného koeficientu a (b) $100\times(1-\alpha)\%$ intervalu spoľahlivosti lineárno-uhlového ρ.*

Riešenie v R

```
"lin.uhl.r" <- function(x,y){
# x ... linearna premenna
# y ... uhlova premenna v stupnoch
y.rad <- (y*pi)/180
y.cos <- cos(y.rad)
y.sin <- sin(y.rad)
r.xc <- cor(x,y.cos)
r.xs <- cor(x,y.sin)
r.cs <- cor(y.cos,y.sin)
citatel <- r.xc^2+r.xs^2-2*r.xc*r.xs*r.cs
menovatel <- 1-r.cs^2
r.xy <- sqrt(citatel/menovatel)
return(r.xy)
}
"ISkor.uhl" <- function(r.xy,n, conf.level = 0.95){
z <- atanh(r.xy)
a <- qnorm(1 - (1 - conf.level)/2)/(n - 3)^0.5
IS.xi <- c(z - a, z + a)
IS <- tanh(IS.xi)
return(IS)
}
```

Príklad 206 (test o lineárno-uhlovom ρ) *Majme dáta lin-uhl-fm.txt a premenné basion-bregmatická výška lebky skull.H (v mm) a uhol, ktorý zvierajú línie prechádzajúce bodom basion a pravostranným a ľavostranným bodom porion base.A (v stupňoch) u mužov. (a) Otestujte hypotézu o tom, že korelačný koeficient basion-bregmatickej výšky lebky a vyššie spomenutého uhla je rovný nule. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre korelačný koeficient, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku Z_W, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: korelačný koeficient basion-bregmatickej výšky lebky a uhla, ktorý zvierajú línie prechádzajúce bodom *basion* a pravostranným a ľavostranným bodom *porion* u mužov je rovný nule oproti H_1: korelačný koeficient basion-bregmatickej výšky lebky a uhla, ktorý zvierajú línie prechádzajúce bodom *basion* a pravostranným a ľavostranným bodom *porion* u mužov nie je rovný nule.
   - **matematická formulácia** – $H_0 : \rho = \rho_0$ oproti $H_1 : \rho \neq \rho_0$, kde $\rho_0 = 0$.

2. **Testovacia štatistika** – Korelačný koeficient ρ je neznámy, preto ho musíme odhadnúť. Jeho odhadom je $\widehat{\rho} \doteq 0.662$.

```
DATA <- read.table("lin-uhl-fm.txt", header=TRUE)
attach(DATA)
n.m <- length(skull.H[sex=="m"]) # 40
rho.hat <- lin.uhl.r(skull.H[sex=="m"], base.A[sex=="m"]) # 0.6618232
```

Pozorované testovacie štatistiky:
Waldova testovacia štatistika $z_W \doteq 4.842$.
Testovacia štatistika t-testu $t_W \doteq 5.442$.
Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} \doteq 23.050$.

```
zR <- 1/2*log((1+rho.hat)/(1-rho.hat)) # 0.7960509
rho0 <- 0
xi0 <- 1/2*log((1+rho0)/(1-rho0)) # 0
zW.obs <- sqrt(n.m-3)*(zR-xi0) # 4.842188
tW.obs <- sqrt(n.m-2)*rho.hat/(sqrt(1-rho.hat^2)) # 5.442137
uLR.obs <- -n.m*log(1-tW.obs^2/(tW.obs^2+(n-2)))  # 23.05084
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $u(1-\alpha/2) \doteq -1.960$ a $u(\alpha/2) \doteq 1.960$;
   kritický obor $\mathcal{W}_1 = (z_{\min}, u(1-\alpha/2)) \cup (u(\alpha/2), z_{\max}) \doteq (-\infty, -1.960) \cup (1.960, \infty)$.
   t-test:
   kritické hodnoty $t_{n-2}(1-\alpha/2) = t_{38}(1-0.025) \doteq -2.024$ a $t_{n-2}(\alpha/2) = t_{38}(0.025) \doteq 2.024$;
   kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-2}(1-\alpha/2)) \cup (t_{n-2}(\alpha/2), t_{\max}) = (-\infty, -2.024) \cup (2.024, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi^2_1(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi^2_1(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
z.krit.d <- qnorm(0.025) # -1.959964
z.krit.h <- qnorm(0.975) # 1.959964
t.krit.d <- qt(0.025, df=n.m-2) # -2.024394
t.krit.h <- qt(0.975, df=n.m-2) # 2.024394
lr.krit.hodn <- qchisq(0.95, df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100(1-\alpha)\%$ empirický DIS = 95% empirický DIS pre ρ:
   $(d, h) \doteq (0.441, 0.807)$.
   Vierohodnostný 95% empirický DIS pre ρ:
   $\mathcal{CS}_{0.95} = \left\{\rho_0 : u_{\mathrm{LR}}(\rho_0) < \chi^2_1(\alpha)\right\} \doteq (0.449, 0.804)$.

```
ISz.xy <- IScor.uhl(rho.hat, n.m) # 0.4412925 0.8069653
rho.0i <- seq(-0.999, 0.999, by=0.0001)
uLR.i <- -n.m*(log((1-rho.0i^2)*(1-rho.hat^2)/(1-rho.0i*rho.hat)^2))
IS.LR <- range(rho.0i[which(uLR.i<lr.krit.hodn)]) # 0.4492 0.8035
```

5. **Štatistický záver** –
   Waldov test: p-hodnota=$2\Pr(Z_W \geq |6.010||H_0) < 0.0001$
   t-test: p-hodnota=$2\Pr(T_W \geq |6.723||H_0) < 0.0001$.
   Test pomerom vierohodnosti: p-hodnota=$\Pr(U_{\mathrm{LR}} \geq 23.050|H_0) < 0.0001$.

```
p.hodn.zW <- 2*(1-pnorm(abs(zW.obs))) # 1.854635e-09
p.hodn.tW <- 2*(1-pnorm(abs(tW.obs))) # 1.774803e-11
p.hodn.LR <- 1-pchisq(uLR.obs, df=1) # 4.098731e-09
```

   H_0 na hladine významnosti zamietame, pretože (1) testovacia štatistika patrí kritickému oboru, (2) $\rho_0 = 0$ nepatrí DIS a nakoniec (3) p-hodnota je menšia ako 0.05.

6. **Slovný záver** – Zamietame nulovú hypotézu o tom, že korelačný koeficient basion-bregmatickej výšky lebky a uhla, ktorý zvierajú línie prechádzajúce bodom *basion* a pravostranným a ľavostranným bodom *porion* u mužov, je rovný nule.

7. **Antropologický záver** – Veľkosť uhla s basion-bregmatickou výškou lebky pozitívne koreluje (r=0.66), t. j. so zväčšujúcou sa basion-bregmatickou výškou lebky sa zväčšuje aj veľkosť sledovaného uhla. Je otázkou, či je to v dôsledku veľkostných zmien a vyššie lebky sú súčasne aj širšie (oba body *porion* sú od seba vzdialenejšie) alebo sa so zvyšujúcou sa basion-bregmatickou výškou lebky znižuje výška lebečnej bázy (minimálna vzdialenosť bodu *basion* k spojnici pravostranného a ľavostranného bodu *porion*), vylúčená nie je ani kombinácia týchto faktorov.

Príklad 207 (test o uhlovom ρ) *Naprogramujte v R funkcie na výpočet odhadu uhlového korelačného koeficientu.*

Riešenie v R

```
"priem.uhol" <- function(x.cos,x.sin){
# x.cos ... kosinus uhlovej premennej
# x.sin ... sinus uhlovej premennej
sum.x.cos <- sum(x.cos)
sum.x.sin <- sum(x.sin)
if (sum.x.cos>0) x.bar <- atan(sum.x.sin/sum.x.cos)
if (sum.x.cos<0) x.bar <- atan(sum.x.sin/sum.x.cos)+pi
if (sqrt(sum.x.sin^2+sum.x.cos^2)==0) x.bar <- NaN
return(x.bar)
}
"sin.cos.uhla" <- function(x){
x.rad <- (x*pi)/180
x.cos <- cos(x.rad)
x.sin <- sin(x.rad)
VYSL <- list(x.cos=x.cos,x.sin=x.sin)
return(VYSL)
}
"uhl.uhl.r" <- function(x,y){
# x ... uhlova premenna v stupnoch
# y ... uhlova premenna v stupnoch
sin.cos.x <- sin.cos.uhla(x)
x.cos <- sin.cos.x$x.cos
x.sin <- sin.cos.x$x.sin
x.bar <- priem.uhol(x.cos,x.sin)
x.sin.centr <- sin(x.rad-x.bar)
sin.cos.y <- sin.cos.uhla(y)
y.cos <- sin.cos.y$x.cos
y.sin <- sin.cos.y$x.sin
y.bar <- priem.uhol(y.cos,y.sin)
y.sin.centr <- sin(y.rad-y.bar)
citatel <- sum(x.sin.centr*y.sin.centr)
menovatel <- sqrt(sum((x.sin.centr^2)))*sqrt(sum((y.sin.centr^2)))
r.xy <- citatel/menovatel
return(r.xy)
}
```

Príklad 208 (test o uhlovom ρ) *Majme dáta* uhl-uhl-fm.txt *a premenné uhol v bode nasion* front.A *(v stupňoch) a uhol tvárového trojuholníka v bode prosthion* prog.A *(v stupňoch) u žien. (a) Otestujte hypotézu o tom, že korelačný koeficient týchto dvoch uhlov je rovný nule. (b) Vypočítajte $100\times(1-\alpha)\%$ empirický DIS pre korelačný koeficient, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku Z_W, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} a DIS prislúchajúce (1) a (2). Výsledky Waldovho testu skontrolujte pomocou funkcie* cor.test().

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: korelačný koeficient uhla v bode *nasion* a uhla tvárového trojuholnḱa v bode *prosthion* u žien je rovný nule oproti H_1: korelačný koeficient uhla v bode *nasion* a uhla tvárového trojuholnḱa v bode *prosthion* u žien nie je rovný nule.
   - **matematická formulácia** – $H_0 : \rho = \rho_0$ oproti $H_1 : \rho \neq \rho_0$, kde $\rho_0 = 0$.
2. **Testovacia štatistika** – Korelačný koeficient ρ je neznámy, preto ho musíme odhadnúť. Jeho odhadom je $\widehat{\rho} \doteq 0.335$.

```
DATA <- read.table("uhl-uhl-fm.txt",header=TRUE)
attach(DATA)
n.f <- length(front.A[sex=="f"]) # 20
rho.hat <- uhl.uhl.r(front.A[sex=="f"],prog.A[sex=="f"]) # 0.3347396
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $z_W \doteq 1.435$.
   Testovacia štatistika t-testu $t_W \doteq 1.507$.
   Testovacia štatistika pomerom vierohodnosti $U_{\mathrm{LR}} \doteq 2.377$.

```
zR <- 1/2*log((1+rho.hat)/(1-rho.hat)) # 0.3481565
rho0 <- 0
xi0 <- 1/2*log((1+rho0)/(1-rho0)) # 0
zW.obs <- sqrt(n.f-3)*(zR-xi0) # 1.435486
tW.obs <- sqrt(n.f-2)*rho.hat/(sqrt(1-rho.hat^2)) # 1.507125
uLR.obs <- -n.f*(log((1-rho0^2)*(1-rho.hat^2)/(1-rho0*rho.hat)^2))
#  2.376811
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $u(1-\alpha/2) \doteq -1.960$ a $u(\alpha/2) \doteq 1.960$;
   kritický obor $\mathcal{W}_1 = (z_{\min}, u(1-\alpha/2)) \cup (u(\alpha/2), z_{\max}) \doteq (-\infty, -1.960) \cup (1.960, \infty)$.
   t-test:
   kritické hodnoty $t_{n-2}(1-\alpha/2) = t_{18}(1-0.025) \doteq -2.101$ a $t_{n-2}(\alpha/2) = t_{18}(0.025) \doteq 2.101$;
   kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-2}(1-\alpha/2)) \cup (t_{n-2}(\alpha/2), t_{\max}) \doteq (-\infty, -2.101) \cup (2.101, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi^2_1(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi^2_1(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
z.krit.d <- qnorm(0.025) # -1.959964
z.krit.h <- qnorm(0.975) # 1.959964
t.krit.d <- qt(0.025,df=n.f-2) # -2.100922
t.krit.h <- qt(0.975,df=n.f-2) # 2.100922
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100(1-\alpha)\%$ empirický DIS $= 95\%$ empirický DIS pre ρ:
   $(d, h) \doteq (-0.127, 0.677)$.
   Vierohodnostný 95% empirický DIS pre ρ:
   $\mathcal{CS}_{0.95} = \left\{\rho_0 : u_{\mathrm{LR}}(\rho_0) < \chi^2_1(\alpha)\right\} \doteq (-0.097, 0.660)$.

```
ISz.xy <- IScor.uhl(rho.hat,n.f) # -0.1265229  0.6769799
rho.0i <- seq(-0.999,0.999,by=0.0001)
uLR.i <- -n.f*(log((1-rho.0i^2)*(1-rho.hat^2)/(1-rho.0i*rho.hat)^2))
IS.LR <- range(rho.0i[which(uLR.i<lr.krit.hodn)]) # -0.0968  0.6603
```

5. **Štatistický záver** –
   Waldov test: p-hodnota $\doteq 2\Pr(Z \geq |1.435||H_0) \doteq 0.151$
   t-test: p-hodnota $\doteq 2\Pr(T \geq |1.507||H_0) \doteq 0.132$.
   Test pomerom vierohodnosti: p-hodnota$=\Pr(U \geq 2.377|H_0) \doteq 0.123$.

```
p.hodn.zW <- 2*(1-pnorm(abs(zW.obs))) #  0.1511486
p.hodn.tW <- 2*(1-pnorm(abs(tW.obs))) # 0.1317786
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # 0.1231487
```

   H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\rho_0 = 0$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že korelačný koeficient medzi veľkosťou uhla v bode *nasion* a veľkosťou uhla tvárového trojuholníka v bode *prosthion* je u žien rovný nule.

7. **Antropologický záver** – Nedokázali sme súvislosť medzi tvarom lebky v oblasti bodu *nasion* a bodu *prosthion*. Sklon čela (zvislosť šupiny čelovej kosti) v mediosagitálnej rovine teda nesúvisí s analogicky definovaným uhlom v hornej časti tváre v bode *prosthion*, i keď korelácia je pozitívna a pri väčšej vzorke by sme závislosť nájsť mohli. Obe časti lebky (predná časť neurokrania a horná časť splanchnokrania) sú teda pomerne samostatnými jednotkami a ich tvar je ovplyvnený inými faktormi, napriek tomu, že spolu anatomicky úzko súvisia.

6.4 Asymptotické testy o pravdepodobnosti

Testy pravdepodobností (proporcií) sa v antropológii používajú pri hodnotení kvalitatívnych znakov u živých ľudí i na kostre človeka. Kvalitatívne znaky sú znaky, ktoré nehodnotíme meraním, ale posudzujeme ich prítomnosť či neprítomnosť. Populácie sa potom medzi sebou líšia proporciou prípadov s výskytom znaku a bez výskytu znaku. Medzi kvalitatívne znaky môžeme zaradiť napr.

krvné skupiny s jedným antigénom, v prípade ktorých ide o frekvenciu jedincov s prítomnosťou alebo neprítomnosťou danej protilátky. Kvalitatívne sa tradične popisuje počet jedincov každého pohlavia v populácii (pomer pohlaví), stranová asymetria (väčšia hodnota vpravo alebo väčšia hodnota vľavo) či stranové preferencie (praváci alebo ľaváci). Iným príkladom kvalitatívnych znakov sú epigenetické znaky na lebke (napr. metopizmus – perzistencia *sutura interfrontalis* v dospelosti), kde väčšinou jedinca zaraďujeme do jednej z dvoch kategórií (znak prítomný alebo neprítomný) alebo dermatoglyfické znaky, u ktorých jedincov radíme do dvoch kategórií (napr. prítomnosť či neprítomnosť digitálneho trirádia c). Ďalším príkladom kvalitatívneho znaku môže byť chuťová vnímavosť fenylthiokarbamidu, kde jedincov rozdeľujeme do dvoch skupín podľa toho, či horkú chuť látok s touto zložkou vnímajú alebo nie. Kvalitatívnych znakov sa v antropológii rozlišuje veľké množstvo a výsledky ich hodnotenia sa ľahko publikujú v texte či tabuľkách, čo umožňuje rozsiahle porovnania s publikovanými výsledkami rôznych výskumov (na rozdiel od spojitých znakov, kde sa v súčasnosti primárne dáta/individuálne hodnoty väčšinou nepublikujú a pri porovnávaní musíme vychádzať zo sekundárnych dát). Veľmi jednoducho tak môžeme otestovať, či sa proporcie zistené v našom výskume (napr. pomer jedincov vnímajúcich a nevnímajúcich chuť fenylthiokarbamidu v českej populácii) a pomer zistený v publikovanej porovnávacej štúdii líši.

Nech $X \sim Bin(N, p)$. Majme dvojice hypotéz $H_{01} : p = p_0$ oproti $H_{11} : p \neq p_0$, $H_{02} : p \leq p_0$ oproti $H_{12} : p > p_0$, $H_{03} : p \geq p_0$ oproti $H_{13} : p < p_0$, ktoré chceme testovať.

Ak H_0 platí, potom

$$Z_W = \frac{X/N - p_0}{\sqrt{X/N\,(1 - X/N)/\,N}} \overset{\mathcal{D}}{\sim} N(0,1).$$

Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **jednovýberový Z-test o pravdepodobnosti** p.

Minimálny rozsah výberu pre normálnu aproximáciu stanovíme na základe *Haldovej podmienky* $Np(1-p) > 9$ (pozri tabuľku 7.4).

Tabuľka 6.3: Minimálne rozsahy N pre rôzne p v súvislosti s Haldovou podmienkou

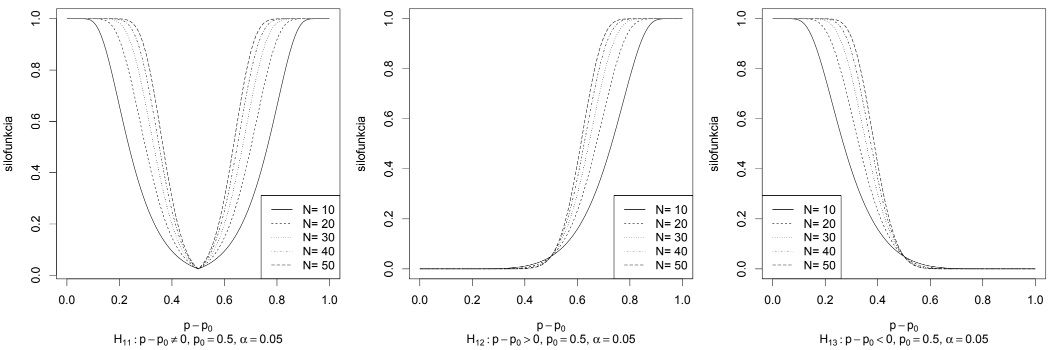
p	0.01	0.02	0.05	0.10	0.15	0.20	0.30	0.40	0.50
$1-p$	0.99	0.98	0.95	0.90	0.85	0.80	0.70	0.60	0.50
N	910	460	190	100	71	57	43	38	36

Definícia 61 (Kritický obor a silofunkcia Z_W testu o p) *Kritický obor a silofunkcia sú definované nasledovne (pozri obrázok 6.8):*

H_0	H_1	$\mathcal{W}$	$1-\beta(p)$
$p = p_0$	$p \neq p_0$	$\mathcal{W}_1 = \{Z_W; \lvert Z_W\rvert \geq u(\alpha/2)\}$	$1 - \Phi\left(u_{\alpha/2} - \frac{\lvert p_0 - p\rvert}{\sqrt{p(1-p)/N}}\right)$
$p \leq p_0$	$p > p_0$	$\mathcal{W}_2 = \{Z_W; Z_W \geq u(\alpha)\}$	$1 - \Phi\left(u_{\alpha} + \frac{p_0 - p}{\sqrt{p(1-p)/N}}\right)$
$p \geq p_0$	$p < p_0$	$\mathcal{W}_3 = \{Z_W; Z_W \leq -u(\alpha)\}$	$1 - \Phi\left(u_{\alpha} - \frac{p_0 - p}{\sqrt{p(1-p)/N}}\right)$

Definícia 62 (p-hodnota Z_W testu o p) *Nech Z_W je **nejaká testovacia štatistika** a $z_W = \frac{\widehat{p} - p_0}{\sqrt{\widehat{p}(1-\widehat{p})/N}}$ je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W \geq |z_W| \,|H_{01}), \ ak & H_{11} : p \neq p_0 \\ Pr(Z_W \geq z_W | H_{02}), \ ak & H_{12} : p > p_0 \\ Pr(Z_W \leq z_W | H_{03}), \ ak & H_{13} : p < p_0 \end{cases}$$



Obr. 6.8: Silofunkcie asymptotického testu o p pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 63 (Waldove $100 \times (1-\alpha)\%$ empirické IS pre p) *Waldove $100\times(1-\alpha)\%$ empirické IS pre p pre všetky tri typy hypotéz majú nasledovný tvar:*

H_0	H_1	*hranice (d,h) pre $100(1-\alpha)\%$ empirický IS*
$p = p_0$	$p \neq p_0$	$\mathcal{CS}_{1-\alpha} = \left\{p_0 : p_0 \in \left(\widehat{p} - u_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})/N}, \widehat{p} + u_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})/N}\right)\right\}$
$p \leq p_0$	$p > p_0$	$\mathcal{CS}_{1-\alpha} = \left\{p_0 : p_0 \in \left(\widehat{p} - u_{\alpha}\sqrt{\widehat{p}(1-\widehat{p})/N}, 1\right)\right\}$
$p \geq p_0$	$p < p_0$	$\mathcal{CS}_{1-\alpha} = \left\{p_0 : p_0 \in \left(0, \widehat{p} + u_{\alpha}\sqrt{\widehat{p}(1-\widehat{p})/N}\right)\right\}$

Majme H_{11}. Ak H_{01} neplatí, chceme túto skutočnosť odhaliť s pravdepodobnosťou (silou) aspoň $1-\beta$. Ak $c = \frac{p-p_0}{\sigma}$, potom **minimálny rozsah** N definujeme nasledovne

$$N \geq \left(\frac{u_{\alpha/2} + u_\beta}{c}\right)^2 = \left(\frac{u_{\alpha/2} + u_\beta}{p - p_0}\right)^2 p\,(1-p)\,.$$

Pre H_{12} a H_{13} zameníme $u_{\alpha/2}$ za u_α.

Majme testovaciu štatistiku $Z_W^{(\mathrm{alt})} = \frac{g(\widehat{p})-g(p_0)}{\widehat{SE}[g(\widehat{p})]}$ a Waldov $100 \times (1-\alpha)\%$ empirický IS rovný (d,h) pre tri tranformácie nájdeme spätnou transformáciou (d_g, h_g) nasledovne:

1. ak $g(p) = \frac{p}{1-p}$, potom $\widehat{SE[g(\widehat{p})]} = \sqrt{\frac{\widehat{p}}{N(1-\widehat{p})^3}}$, $d = \frac{d_g}{1+d_g}$ a $h = \frac{h_g}{1+h_g}$
2. ak $g(p) = \ln(\frac{p}{1-p})$, potom $\widehat{SE[g(\widehat{p})]} = \sqrt{\frac{1}{N\widehat{p}(1-\widehat{p})}}$, $d = \frac{\exp(d_g)}{1+\exp(d_g)}$ a $h = \frac{\exp(h_g)}{1+\exp(h_g)}$.
3. ak $g(p) = \arcsin(\sqrt{p})$, potom $\widehat{SE[g(\widehat{p})]} = \frac{1}{2\sqrt{N}}$, $d = \sin^2(d_g)$ a $h = \sin^2(h_g)$.

Príklad 209 (vylepšená vierohodnosť pomocou $g(\theta)$) *Nakreslite (a) logaritmus funkcie vierohodnosti parametra $g(p) = \frac{p}{1-p}$ binomického rozdelenia $Bin(N,p)$, kde $N = 10$ a $n = 8$, superponovaný jeho kvadratickou aproximáciou. Nakreslite (b) logaritmus funkcie vierohodnosti $g(p) = \ln\frac{p}{1-p}$ (pri rovnakom zadaní N a n ako v (a)), superponovaný jeho kvadratickou aproximáciou. Nakreslite (c) logaritmus funkcie vierohodnosti $g(p) = \arcsin\left(\sqrt{p}\right) = \sin^{-1}\sqrt{p}$ (pri rovnakom zadaní N a n ako v (a); táto transformácia stabilizuje rozptyl p), superponovaný jeho kvadratickou aproximáciou. Je funkcia vierohodnosti $g(p)$ regulárnejšia ako funkcia vierohodnosti pre p? (d) Vypočítajte vierohodnostný $100\times(1-\alpha)\%$ empirický IS pre $g(p)$ z (a), (b) a (c). (e) Vypočítajte Waldov $100\times(1-\alpha)$ % empirický IS pre $g(p)$ a transformujte ho späť do originálnej škály. (f) Ukážte, že vierohodnostný IS pre p v škále p je identický s vierohodnostným IS v škále $g(p)$ po jeho spätnej transformácii do originálnej škály.*

Riešenie (čiastočné; odvodenie odhadu rozptylu $g(p)=\frac{p}{1-p}$)

$$\widehat{Var\left(g(\widehat{p})\right)} = \frac{\left[\frac{\partial}{\partial p}\left(\frac{p}{1-p}\right)\right]^2}{-\frac{\partial^2}{\partial p^2}\ln L\left(p|\mathbf{x}\right)} = \frac{\left[\frac{(1-p)+p}{(1-p)^2}\right]^2}{\frac{N}{p(1-p)}} \overset{p=\widehat{p}}{=} \frac{\widehat{p}}{N\left(1-\widehat{p}\right)^3}.$$

Skóre testovaciu štatistiku U_S odvodíme nasledovne

$$U_\mathrm{S} = \frac{\left[\frac{\partial}{\partial\theta}\ln L\left(p|\mathbf{x}\right)\right]^2}{\mathcal{I}(p)} = \frac{\left(\frac{X}{p}-\frac{N-X}{1-p}\right)^2}{\frac{N}{p(1-p)}} \overset{p=p_0}{=} \frac{\left(X/N-p_0\right)^2}{p_0\left(1-p_0\right)/N} \overset{\mathcal{D}}{\sim} \chi^2_1.$$

Skóre $100\times(1-\alpha)\%$ **empirický DIS pre** p má tvar $\mathcal{CS}_{1-\alpha}=\left\{|Z_S|\le u_{\alpha/2}\right\}$, kde $Z_S=\frac{X/N-p_0}{\sqrt{p_0(1-p_0)/N}}$, čo je identické $\mathcal{CS}_{1-\alpha}=\left\{U_\mathrm{S}\le\chi^2_{\alpha/2}\right\}$. Po niekoľkých algebraických úpravách môžeme $\mathcal{CS}_{1-\alpha}$ prepísať do kvadratickej nerovnice

$$\mathcal{CS}_{1-\alpha}=\left\{p_0:\left(1+\frac{u^2_{\alpha/2}}{N}\right)p_0^2-\left(2\widehat{p}+\frac{u^2_{\alpha/2}}{N}\right)p_0+\widehat{p}^2\le 0\right\}.$$

Korene tejto nerovnice definujú hranice DIS (d,h) nasledovne (Wilson, 1927)

$$\frac{2\widehat{p}+u^2_{\alpha/2}/N\pm\sqrt{\left(2\widehat{p}+u^2_{\alpha/2}/N\right)^2-4\widehat{p}^2\left(1+u^2_{\alpha/2}/N\right)}}{2\left(1+u^2_{\alpha/2}/N\right)}.$$

Po ďalších úpravách dostaneme DIS v tvare

$$\widehat{p}\left(\frac{N}{N+u^2_{\alpha/2}}\right)+\frac{1}{2}\left(\frac{u^2_{\alpha/2}}{N+u^2_{\alpha/2}}\right)\pm u_{\alpha/2}\sqrt{\frac{1}{N+u^2_{\alpha/2}}\left[\widehat{p}\left(1-\widehat{p}\right)\left(\frac{N}{N+u^2_{\alpha/2}}\right)+\frac{1}{4}\frac{u^2_{\alpha/2}}{N+u^2_{\alpha/2}}\right]}.$$

Stred tohoto IS je váženým priemerom dvoch členov a patrí intervalu s hranicami $\widehat{p}$ a $\frac{1}{2}$ s váhou pre $\widehat{p}$ kovergujúcou asymptoticky do jednotky. Tento stred posúva $\widehat{p}$ smerom ku 0.5, pričom s rastúcim N sa tento posun zmenšuje. Rozptyl je tiež váženým priemerom rozptylu $\widehat{p}$ a 0.5 použitím $N+u^2_{\alpha/2}$ namiesto rozsahu N (Agresti a Coull, 1998).

Jednovýberový skóre test o pravdepodobnosti p v R (funkcia **prop.test()**).
Argumenty (vstupy) funkcie:

1. početnosť (úspechy n) x;
2. rozsah (počet úspechov a neúspechov N) n;
3. pravdepodobnosť p_0 za platnosti H_0 p ($p_0\in(0,1)$);
4. formulácia alternatívy alternative="two.sided" je prednastavené, ďalšie voľby sú "greater", "less";
5. spoľahlivosť conf.level=p, prednastavené je conf.level=.95;
6. Yatesova korekcia na spojitosť (Newcombe, 1998), prednastavené correct = TRUE.

Výstupy funkcie:

1. testovacia štatistika statistic ($\chi^2=Z_S^2$);
2. stupne voľnosti df (parameter; $df=1$);

3. p-hodnota p.value;
4. odhad pravdepodobnosti $\widehat{p}$ estimate;
5. interval spoľahlivosti conf.int (skóre (Wilsonov) IS);
6. hodnota p_0 null value;
7. alternatívna hypotéza alternative hypothesis;
8. názov použitého testu method.

Funkcia binconf() s knižnice Hmisc počíta intervaly spoľahlivosti pre p. Argument method="wilson" znamená skóre interval spoľahlivosti a method="asymptotic" Waldov interval spoľahlivosti.

Test pomerom vierohodnosti pre p.

Majme H_{01} oproti H_{11}. Nech $X \sim Bin(N,p)$, kde $\theta = p$, potom logaritmus vierohodnostnej funkcie bude mať tvar

$$l(p|\mathbf{x}) = x \ln p + (N-x)\ln(1-p).$$

MLE θ je rovný $\widehat{p} = x/N$, t.j. $\Theta_1 = \{\theta : p \neq p_0\}$. Za platnosti H_{01} je $\theta_0 = p_0$, t.j. $\Theta_0 = \{\theta : p = p_0\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{x})) &= l(\widehat{p}|\mathbf{x}) - l(p_0|\mathbf{x}) = x\ln\widehat{p} + (N-x)\ln(1-\widehat{p}) - x\ln p_0 - (N-x)\ln(1-p_0) \\ &= x\ln\frac{\widehat{p}}{p_0} + (N-x)\ln\frac{1-\widehat{p}}{1-p_0} = x\ln\frac{x}{Np_0} + (N-x)\ln\frac{N-x}{N-Np_0}. \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln(\lambda(\mathbf{X})) \overset{\mathcal{D}}{\sim} \chi_1^2$. Preznačme p na p_1 a $q = 1-p$ na p_2, x na n_1 a $N-x$ na n_2. Potom $\boldsymbol{\theta} = \mathbf{p} = (p_1, p_2)^T$, $\widehat{\boldsymbol{\theta}} = \widehat{\mathbf{p}} = (\widehat{p}_1, \widehat{p}_2)^T$, kde $\widehat{p}_1 = \frac{n_1}{N}$, $\widehat{p}_2 = \frac{n_2}{N}$, $p_1 + p_2 = 1$ a $n_1 + n_2 = N$; $\boldsymbol{\theta}_0 = \mathbf{p}_0 = (p_{01}, p_{02})^T$, kde $p_{01} = p_0$ a $p_{02} = 1 - p_0$. Potom dostaneme

$$u_{\mathrm{LR}} = 2\sum_{i=1}^{2} n_i \ln\frac{\widehat{p}_i}{p_{0i}} = 2\sum_i \text{pozorované}_i \times \ln\frac{\text{pozorované}_i}{\text{očakávané}_i}.$$

Vierohodnostný $100 \times (1-\alpha)\%$ **empirický DIS pre** p_1 bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{p_0 : U_{\mathrm{LR}} \leq \chi_1^2(\alpha)\right\}.$$

Príklad 210 (porovnanie troch DIS v extrémnej situácii) *Nech* $N = 25$ *študentov, ktorým sme položili otázku, či sú vegetariáni. Z nich* $x = 0$ *odpovedalo „áno“. Vypočítajte empirické* $100 \times (1-\alpha)\%$ *DIS (a) Waldov DIS, (b) skóre DIS a (c) vierohodnostný DIS pre* p $(1-\alpha = 0.95)$.

Riešenie (aj v R)
Ak $x = 0, \widehat{p} = 0/25 = 0$.
(a) Waldov 95% empirický DIS pre p má tvar $0 \pm 1.96\sqrt{(0\times 1)/25}$, kde $(d,h) = (0,0)$. Z toho vyplýva, že ak máme pravdepodobnosť na hraniciach výberového priestoru, Waldov DIS nie je vhodný.

```
N <- 25
phat <- 0/N
var.phat <- phat*(1-phat)/N
ISw <- phat+c(-1,1)*qnorm(0.975)*sqrt(var.phat) #  0 0
```

alebo

```
library(Hmisc)
binconf(x=0,n=25,method="asymptotic")
```

(b) Skóre 95% empirický DIS pre p má tvar

$$\frac{1}{2}\left(\frac{1.96^2}{25+1.96^2}\right) \pm 1.96\sqrt{\frac{1.96^2}{4\left(25+1.96^2\right)^2}},$$

kde $(d,h) = (0.000, 0.133)$.

```
skoreIS <- prop.test(x=0,n=25,conf.level=0.95,correct=FALSE)$conf.int
# 0 0.1331923
```

alebo

```
binconf(x=0,n=25,method="wilson")
```

(c) Jadro funkcie vierohodnosti má tvar $L(p|\mathbf{x}) = p^0 (1-p)^{25}$, potom $l(p|\mathbf{x}) = 25\ln(1-p)$. Keďže $L(\widehat{p}|\mathbf{x}) = 0$, bude U_{LR} rovná $-2\left[l(p_0|\mathbf{x}) - l(\widehat{p}|\mathbf{x})\right] = -50\ln(1-p_0)$ a 95% vierohodnostný IS je definovaný ako

$$\mathcal{CS}_{1-\alpha} = \left\{p_0 : U_{\mathrm{LR}} \leq \chi_1^2(\alpha)\right\}, \text{ kde } \chi_1^2(\alpha) = 3.84.$$

Potom $h = 1 - \exp(-3.84/50) = 0.074$ a nakoniec 95% empirický DIS má tvar $(0.000, 0.074)$.

```
"LR.p" <- function(pi.0, x,N, alpha) {
-2*(x*log(pi.0)+(N-x)*log(1-pi.0))-qchisq(1-alpha,df=1)
}
uniroot(f=LR,interval=c(0.000001,.999999),x=0,N=25,alpha=.05)$root # 0.07395093
```

Spomenuté tri metódy dávaju rôzne výsledky. Ak je p blízko nuly, rozdelenie $\widehat{p}$ je vysoko zošikmené doprava pre malé N. V takejto situácii sa odporúča používať skóre alebo vierohodnostný DIS.

Príklad 211 (funkcia vierohodnosti v extrémnej situácii) *Nakreslite funkciu vierohodnosti pre situáciu v predchádzajúcom príklade.*

Príklad 212 (test o pravdepodobnosti p) *Majme dáta* one-sample-probability-sexratio.txt *a premennú pohlavie novorodenca* sex, *kde početnosť chlapcov $n_m = 729$ a početnosť dievčat $n_f = 674$. Nech početnosť chlapcov $X \sim Bin(N,p)$. (a) Otestujte $H_0 : p_m = p_0 = 0.5$ oproti $H_1 : p_m \neq 0.5$ a vypočítajte $100 \times (1-\alpha)$% empirický DIS pre pravdepodobnosť narodenia chlapca p_m, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku $U_W = Z_W^2$, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR}, (3) skóre testovaciu štatistiku $U_S = Z_S^2$ a DIS prislúchajúce (1), (2) a (3). Výsledky skóre testu skontrolujte pomocou funkcie* prop.test(). *(b) Otestujte $H_0 : g(p_m) = g(p_0) = 1$ oproti $H_1 : g(p_m) \neq 1$ pomocou Z_W^{alt}, kde $g(p_m) = \frac{p_m}{1-p_m} = \frac{p_m}{p_f}$, $\widehat{SE[g(\widehat{p}_m)]} = \sqrt{\frac{\widehat{p}_m}{N(1-\widehat{p}_m)^3}}$ a vypočítajte DIS pre* ***pomer pohlaví (=šanca narodenia chlapca)*** *a použitím spätnej trasfomácie DIS pre p_m.*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia (a)** – testujeme H_0: pravdepodobnosť narodenia chlapca je identická s pravdepodobnosťou narodenia dievčaťa oproti H_1: pravdepodobnosť narodenia chlapca nie je identická s pravdepodobnosťou narodenia dievčaťa.
   - **slovná formulácia (b)** – testujeme H_0: pomer pohlaví (šanca narodenia chlapca) je rovný jednej oproti H_1: pomer pohlaví nie je rovný jednej.
   - **matematická formulácia (a)** – $H_0 : p_{\mathrm{m}} = p_0$ oproti $H_1 : p_{\mathrm{m}} \neq p_0$, kde $p_0 = 0.5$.
   - **matematická formulácia (b)** – $H_0 : g(p_{\mathrm{m}}) = g(p_0)$ oproti $H_1 : g(p_{\mathrm{m}}) \neq g(p_0)$, kde $g(p_{\mathrm{m}}) = \frac{p_{\mathrm{m}}}{1-p_{\mathrm{m}}} = \frac{p_{\mathrm{m}}}{p_{\mathrm{f}}}$ a $g(p_0) = 1$.

2. **Testovacia štatistika** – $\widehat{p}_{\mathrm{m}} = \frac{n_{\mathrm{m}}}{n_{\mathrm{m}}+n_{\mathrm{f}}} = \frac{729}{729+674} = 0.52$, $\widehat{p}_{\mathrm{f}} = 0.48$, $\mathrm{SR_m} = \frac{\widehat{p}_{\mathrm{m}}}{\widehat{p}_{\mathrm{f}}} = \frac{n_{\mathrm{m}}}{n_{\mathrm{f}}} = 1.08$.

```
DATA <- read.table("one-sample-probability-sexratio.txt",header=TRUE)
n <- as.numeric(table(DATA))
n.m <- n[2]
n.f <- n[1]
N <- sum(n) # 1403
p.m <- n.m/N # 0.5196009
p.f <- n.f/N # 0.4803991
p <- c(p.f,p.m)
SRm <- p.m/p.f # 1.081602
```

Pozorované testovacie štatistiky:
Waldova testovacia štatistika $z_W \doteq 1.469$ (pre $H_0 : p_{\mathrm{m}} = p_0$).
Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} = 2.157$ (pre $H_0 : p_{\mathrm{m}} = p_0$).
Skóre testovacia štatistika $z_S \doteq 1.468$ (pre $H_0 : p_{\mathrm{m}} = p_0$).
Waldova testovacia štatistika $z_W \doteq 1.412$ (pre $H_0 : g(p_{\mathrm{m}}) = g(p_0)$).

```
p0 <- 0.5
p0.vec <- c(p0,1-p0)
SR0 <- 1
zW.obs <- (p.m-p0)/sqrt(p.m*p.f/N) # 1.469494
uLR.obs <-  2*sum(n*log(p/p0.vec)) # 2.156647
zS.obs <- (p.m-p0)/sqrt(p0*(1-p0)/N) # 1.468364
zS.obs <- sqrt(prop.test(n.m,N,p=0.5,correct=FALSE)$stat) # 1.468364
zWg.obs <- (SRm-SR0)/sqrt(p.m/(N*(1-p.m)^3)) # 1.411887
```

3. **Zamietacia oblasť** –
   Waldov a skóre test:
   kritické hodnoty $u(1-\alpha/2) \doteq -1.960$ a $u(\alpha/2) \doteq 1.960$;
   kritický obor $\mathcal{W}_1 = (z_{\min}, u(1-\alpha/2)) \cup (u(\alpha/2), z_{\max}) \doteq (-\infty, -1.960) \cup (1.960, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi^2_1(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi^2_1(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
z.krit.d <- qnorm(0.025) # -1.959964
z.krit.h <- qnorm(0.975) # 1.959964
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100(1-\alpha)\%$ empirický DIS = 95% empirický DIS pre p_{m}:
   $(d,h) \doteq (0.493, 0.546)$.
   Vierohodnostný 95% empirický DIS pre p_{m}:
   $\mathcal{CS}_{0.95} = \left\{p_0 : u_{\mathrm{LR}}(p_0) < \chi^2_1(\alpha)\right\} \doteq (0.493, 0.546)$.
   Skóre 95% empirický DIS pre p_{m}: $(d,h) \doteq (0.493, 0.546)$.
   Waldov 95% empirický DIS pre $g(p_{\mathrm{m}})$: $(d,h) \doteq (0.968, 1.195)$.
   Waldov 95% empirický DIS pre p_{m}: $(d,h) \doteq (0.492, 0.544)$.

```
ISzW <- p.m+c(-1,1)*qnorm(0.975)*sqrt(p.m*p.f/N) # 0.4934579 0.5457438
p0i <- seq(0.000001,0.999999,by=0.000001)
uLR.i <- 2*(n[2]*log(p[2]/p0i)+n[1]*log(p[1]/(1-p0i)))
IS.LR <- range(p0i[which(uLR.i<lr.krit.hodn)]) # 0.49344 0.545696
ISzS <- prop.test(n.m,N,p=0.5,correct=FALSE)$conf.int # 0.4934400 0.5456547
ISzWg <- SRm+c(-1,1)*qnorm(0.975)*sqrt(p.m/(N*(1-p.m)^3))
# 0.968323 1.194882
ISzWtrans <- c(ISzWg[1]/(1+ISzWg[1]),ISzWg[2]/(1+ISzWg[2]))
# 0.4919533 0.5443946
```

5. **Štatistický záver** –
   Waldov test: p-hodnota=$2\Pr(Z_W \geq |1.469||H_0) \doteq 0.142$ (pre $H_0 : p_{\mathrm{m}} = p_0$).
   Test pomerom vierohodnosti: p-hodnota=$\Pr(U_{\mathrm{LR}} \geq 2.157|H_0) \doteq 0.142$ (pre $H_0 : p_{\mathrm{m}} = p_0$).
   Skóre test: p-hodnota=$2\Pr(Z_S \geq |1.469||H_0) \doteq 0.142$ (pre $H_0 : p_{\mathrm{m}} = p_0$).
   Waldov test: p-hodnota=$2\Pr(Z_W \geq |1.412||H_0) \doteq 0.158$ (pre $H_0 : g(p_{\mathrm{m}}) = g(p_0)$).

```
p.hodn.zW <- 2*(1-pnorm(abs(zW.obs))) # 0.1416988
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # 0.1419542
p.hodn.zS <- 2*(1-pnorm(abs(zW.obs))) # 0.1416988
p.hodn.zS <- prop.test(n.m,N,p=0.5,correct=FALSE)$p.val
p.hodn.zWg <- 2*(1-pnorm(abs(zWg.obs))) # 0.1579831
```

   H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $p_0 = 0.5$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver (a)** – Nezamietame nulovú hypotézu o tom, že pravdepodobnosť narodenia chlapca je identická s pravdepodobnosťou narodenia dievčaťa.

   Slovný záver (b) – Nezamietame nulovú hypotézu o tom, že pomer pohlaví je rovný jednej (šanca narodenia chlapca je rovná jednej).

7. **Antropologický záver** – Prenatálne faktory, ktoré ovplyvňujú vznik a udržanie mužských a ženských embryí a plodov, pôsobili v rovnováhe a neposunuli pomer mimo vyrovnaný (1.0). Vzhľadom na to, že interval spoľahlivosti zahŕňa aj celosvetovú strednú hodnotu pomeru pohlaví novorodencov (1.05) a strednú hodnotu pre českú populáciu (1.06), nemožno v empiricky zistenom pomere pohlaví v sledovanej geografickej oblasti vidieť žiadny vplyv špecifických lokálnych faktorov, ktoré by miestne hodnoty vychyľovali mimo normu.

Príklad 213 (test o pravdepodobnosti p; malý vs. veľký rozsah) *Ako by sa zmenili výsledky predchádzajúceho príkladu, keby sme mali $N = 50$, $n_m = 26$ a $n_f = 24$?*

Riešenie
Pozorované testovacie štatistiky sú rovné:
Waldova testovacia štatistika $z_W \doteq 0.283$ (pre $H_0 : p_\mathrm{m} = p_0$).
Testovacia štatistika pomerom vierohodnosti $u_\mathrm{LR} \doteq 0.080$ (pre $H_0 : p_\mathrm{m} = p_0$).
Skóre testovacia štatistika $z_S \doteq 0.283$ (pre $H_0 : p_\mathrm{m} = p_0$).
Waldova testovacia štatistika $z_W \doteq 0.272$ (pre $H_0 : g(p_\mathrm{m}) = g(p_0)$).
p-hodnoty:
Waldov test: p-hodnota $\doteq 2\Pr(Z_W \geq |0.283||H_0) \doteq 0.777$ (pre $H_0 : p_\mathrm{m} = p_0$).
Test pomerom vierohodnosti: p-hodnota $\doteq \Pr(U_\mathrm{LR} \geq 0.080|H_0) \doteq 0.777$ (pre $H_0 : p_\mathrm{m} = p_0$).
Skóre test: p-hodnota $\doteq 2\Pr(Z_S \geq |0.283||H_0) \doteq 0.777$ (pre $H_0 : p_\mathrm{m} = p_0$).
Waldov test: p-hodnota $\doteq 2\Pr(Z_W \geq |0.272||H_0) \doteq 0.786$ (pre $H_0 : g(p_\mathrm{m}) = g(p_0)$).
IS:
Waldov 95% empirický DIS pre p_m: $(d, h) \doteq (0.382, 0.658)$.
Vierohodnostný 95% empirický DIS pre p:
$\mathcal{CS}_{0.95} = \left\{p_0 : u_\mathrm{LR}(p_0) < \chi^2_1(\alpha)\right\} \doteq (0.383, 0.655)$.
Skóre 95% empirický DIS pre p_m: $(d, h) \doteq (0.358, 0.652)$.
Waldov 95% empirický DIS pre $g(p_\mathrm{m})$: $(d, h) \doteq (0.482, 1.684)$.
Waldov 95% empirický DIS pre p_m: $(d, h) \doteq (0.325, 0.627)$.

Príklad 214 (test o pravdepodobnosti p; minimálny rozsah) *Aký minimálny rozsah N by sme potrebovali na zamietnutie H_0 z predchádzajúceho príkladu, ak $\alpha = 0.05$, $1 - \beta = 0.8$?*

Riešenie v R

```
N <- ceiling(((qnorm(0.975)+qnorm(0.8))/(p.m-p0))^2*p.m*p.f) # 5100
```

Príklad 215 (test o pravdepodobnosti p) *Majme dáta* one-sample-probability-sexratio.txt *a premennú pohlavie novorodenca* sex*, kde početnosť chlapcov $n_m = 729$ a početnosť dievčat $n_f = 674$. Nech početnosť chlapcov $X \sim Bin(N, p_m)$. Otestujte $H_0 : g(p_m) = g(p_0)$ oproti $H_1 : g(p_m) \neq g(p_0)$ pomocou Z_W^{alt}, kde $g(p_m) = \frac{p_m}{1-p_m}$, $\widehat{SE[g(\widehat{p}_m)]} = \sqrt{\frac{\widehat{p}_m}{N(1-\widehat{p}_m)^3}}$. (a) $g(p_0) = 1.06$ a (b) $g(p_0) = 1.06$.*

Príklad 216 (Fisherova miera informácie) *Vypočítajte Fisherovu mieru informácie $\mathcal{I}(g(\widehat{p}))$, kde (a) $g(p) = \ln\frac{p}{1-p}$ a (b) $g(p) = \arcsin(\sqrt{p})$.*

Príklad 217 (Fisherova miera informácie) *Porovnajte DIS pre p použitím (a) Waldovho DIS, (b) vierohodnostného DIS, (c) skóre DIS, (d) transformovaného Waldovho DIS pre $g(p) = \frac{p}{1-p}$, (f) transformovaného Waldovho DIS pre $g(p) = \ln\frac{p}{1-p}$, (g) transformovaného Waldovho DIS pre $g(p) = \arcsin(\sqrt{p})$, kde (1) $N = 1400$ a $p = 0.05$, (2) $N = 40$ a $p = 0.05$, (3) $N = 1400$ a $p = 0.25$, (4) $N = 40$ a $p = 0.25$, (5) $N = 1400$ a $p = 0.5$, (6) $N = 40$ a $p = 0.5$.*

Príklad 218 (minimálny rozsah N) *Vypočítajte minimálny rozsah n pre $p = 0.1, 0.2, \ldots, 0.9$, $p_0 = 0$ pri $\alpha = 0.05$, $\beta = 0.8$ a obojstrannej alternatíve H_{11}. Skontrolujte, či je splnená Haldova podmienka. Ak nie je, doplňte minimálne N, ktoré túto podmienku spĺňa.*

Riešenie (pozri tabuľku 6.4)

Tabuľka 6.4: Minimálne rozsahy N pre rôzne rozdiely p a p_0 (kde $p_0 = 0$) v porovnaní s minimálnym rozsahom vypočítaným pomocou Haldovej podmienky (tučne sú zvýraznené tie N, ktoré spĺňajú obe kritériá)

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
N	71	32	19	12	8	6	4	2	1
$Np(1-p)$	6.39	5.12	3.99	2.88	2.00	1.44	0.84	0.32	0.09
$9/(p(1-p))$	**100**	**57**	**43**	**38**	**36**	**38**	**43**	**57**	**100**

Príklad 219 (minimálny rozsah N) *Vypočítajte minimálny rozsah N pre $p = 0.1, 0.2, \ldots, 0.9$, p_0 vždy o 0.1 menšie ako p, pri $\alpha = 0.05$, $\beta = 0.8$ a obojstrannej alternatíve H_{11}. Skontrolujte, či je splnená Haldova podmienka. Ak nie je, doplňte minimálne N, ktoré túto podmienku spĺňa.*

Riešenie (pozri tabuľku 6.5)

Tabuľka 6.5: Minimálne rozsahy N pre rozdiel $p - p_0 = 0.1$ pri rôznych p a p_0 v porovnaní s minimálnym rozsahom vypočítaným pomocou Haldovej podmienky (tučne sú zvýraznené tie N, ktoré spĺňajú obe kritériá)

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
p_0	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$p(1-p)$	0.09	0.16	0.21	0.24	0.25	0.24	0.21	0.16	0.09
N	71	**126**	**165**	**189**	**197**	**189**	**165**	**126**	71
$Np(1-p)$	6.39	20.16	34.65	45.36	49.25	45.36	34.65	20.16	20.16
$9/(p(1-p))$	**100**	57	43	38	36	38	43	57	**100**

Príklad 220 (pravdepodobnost pokrytia) *Nech $X \sim Bin(N, p)$, kde $N = 30$ a $p = 0.8$ a pravdepodobnosť úspechu $\widehat{p} = \frac{24}{30} = 0.8$, kde $x = 24$ a $N = 30$. Waldov 95% empirický DIS pre p je rovný $(d, h) = (0.657, 0.943)$. Vypočítajte pravdepodobnosť pokrytia tohoto intervalu. Pozn.: pravdepodobnosť pokrytia Waldovho 95% DIS pre p vypočítame nasledovne*

$$\Pr(pokrytie) = \sum_j \Pr(X = Np_j : p \in Waldov\ 95\%\ DIS\ pre\ p_j),$$

kde $p_j \in \mathcal{M}_J = \left\{\frac{1}{30}, \frac{2}{30}, \ldots, 1 - \frac{1}{30}\right\}$, t.j. ide o súčet takých funkčných hodnôt pravdepodobnostnej funkcie v bodoch Np_j, kde $p \in$ Waldovmu 95% DIS pre p_j. Výsledky usporiadajte do tabuľky, ktorej stĺpce budú x_j, p_j, d_j (dolná hranica Waldovho 95% DIS pre p_j), h_j (horná hranica Waldovho 95% DIS pre p_j), $\Pr$(pokrytie) a pokrytie (indikácia toho, či p patrí alebo nepatrí Waldovmu 95% DIS pre p_j).

Riešenie (aj v R)
Nech $X \sim Bin(30, 0.8)$. Pri $N = 30$ môže byť $x = 0, 1, \ldots, 30$ (spolu 31 možností), a teda máme aj 31 možností aj pre p. Na zodpovedanie otázky ohľadom pravdepodobnosti pokrytia je potrebné vypočítať 95% empirický DIS pre každé p a z pravdepodobnostnej funkcie pre každé X vybrať také pravdepodobnosti korešpondujúce x, ktorých 95% empirický DIS pokrýva hodnotu 0.8.

```
N <- 30
x <- 0:N
p.hat.j <- x/N
DH <- p.hat.j-qnorm(0.975)*sqrt(p.hat.j*(1-p.hat.j)/N)
HH <- p.hat.j+qnorm(0.975)*sqrt(p.hat.j*(1-p.hat.j)/N)
p.hat <- 0.8
prav.fcia <- dbinom(x,N,p.hat)
pokrytie <- (p.hat>/DH)&(p.hat</HH)  # vektor 0 a 1
round(cbind(x, p.hat.j,DH,HH,prav.fcia,pokrytie),4)
sum(dbinom(x[pokrytie],N,p.hat)) # 0.9463279
```

Pravdepodobnosť pokrytia Waldoho 95% empirický DIS pre $p = 0.8$ je rovná $\sum_{x=19}^{27} \Pr(X = x) = 0.946$. V prípade, že $p = 0.79$ je už pravdepodobnosť pokrytia len 0.888.

Príklad 221 (pravdepodobnost pokrytia) *Nech $X_i \sim Bin(N, p_i)$. Vypočítajte pravdepodobnosti pokrytia Waldovho 95% DIS pre každé p_i, kde p_i patria množine $\mathcal{M}_I = \left\langle \frac{1}{N}, 1 - \frac{1}{N} \right\rangle$, sú ekvidištantne vzdialené medzi $\frac{1}{N}$ a $1 - \frac{1}{N}$ a ich počet $M = 5000$. Nakreslite obrázok, kde na osi x budú p_i a na osi y pravdepodobnosť pokrytia $\Pr_i$(pokrytie). Zvoľte (a) $N = 30$, (b) $N = 100$ a (c) $N = 1000$. Pozn.: pravdepodobnosti pokrytia Waldovho 95% DIS pre p_i vypočítame nasledovne*

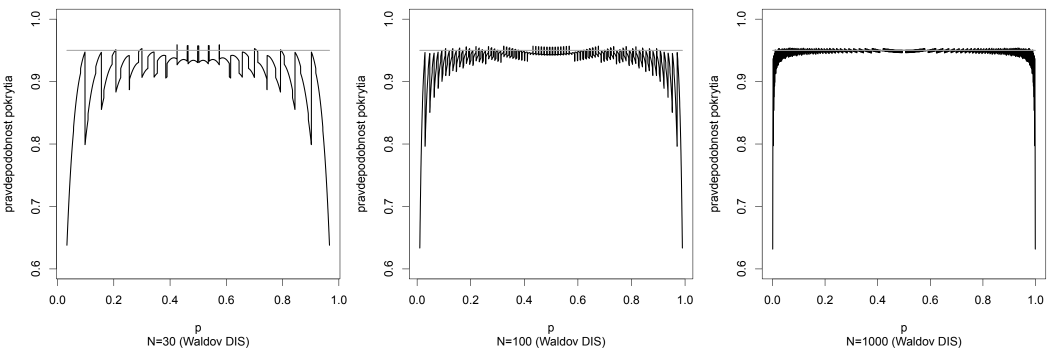
$$\Pr_i(pokrytie) = \sum_j \Pr(X = Np_j : p_i \in Waldov\ 95\%\ DIS\ pre\ p_j),$$

kde $p_j \in \mathcal{M}_J = \left\{\frac{1}{N}, \frac{2}{N}, \ldots, 1 - \frac{1}{N}\right\}$, *t.j. ide o súčet takých funkčných hodnôt pravdepodobnostnej funkcie v bodoch* Np_j, *kde* $p_i \in$ *Waldovmu* 95% *DIS pre* p_j.

Riešenie (aj v R; pozri obrázok 6.9)

```
N <- 30
x <- 0:N
p.hat.j <- x/N
sd.p.j <- sqrt(p.hat.j*(1-p.hat.j)/N)
DH <- p.hat.j-qnorm(0.975)*sd.p.j
HH <- p.hat.j+qnorm(0.975)*sd.p.j
M = 5000
p.hat.i <- seq(1/N,1-1/N,length=M)
p.pokrytie.i <- numeric(M)
for (i in 1:M)
{
   pokrytie.i <- (p.hat.i[i]>/DH)&(p.hat.i[i]</HH)
   prav.fcia <- dbinom(x[pokrytie.i],N,p.hat.i[i])
   p.pokrytie.i[i] <- sum(prav.fcia)
}
windows(8,8)
plot(p.hat.i,p.pokrytie.i,type="n",xlab="p",ylab="pravdepodobnost pokrytia",
      sub="N=30",ylim=c(0.6,1))
lines(c(1/N,1-1/N), c(0.95,0.95),lwd=2,col="gray")
lines(p.hat.i,p.pokrytie.i,lwd=2)
```

Zo simulačnej štúdie Waldovho DIS pre p vyplýva (pozri obrázok 6.9), že aktuálne hodnoty pravdepodobnosti pokrytia sú závislé na hodnote p a sú dostatočne blízko nominálnej hodnote $1-\alpha = 0.95$ len pre p dostatočne ďaleko od nuly a jednotky (t.j. n a p nie príliž malé alebo veľké). Ak je N dostatočne veľké, aktuálne hodnoty pravdepodobnosti pokrytia sú dostatočne blízko nominálnej hodnote pre p takmer na celom intervale $(0, 1)$. Tento DIS je príliž liberálny pre malé N a je liberálny pre stredne veľké N.



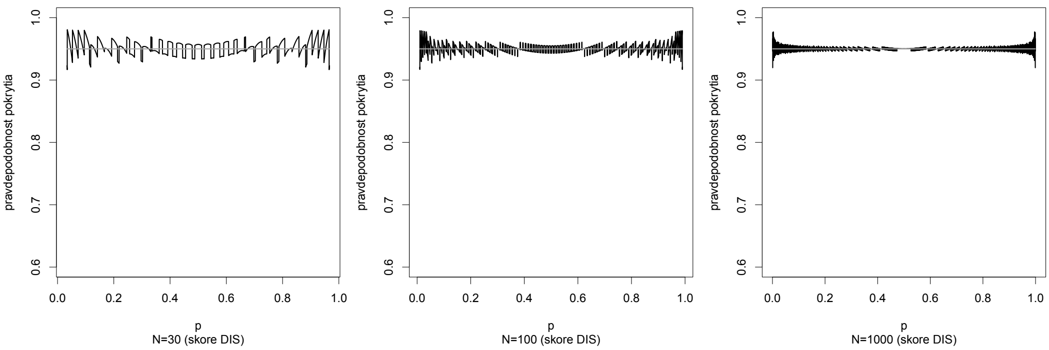
Obr. 6.9: Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p; $N = 30$ (vľavo), 100 (uprostred) a $N = 1000$ (vpravo)

Príklad 222 (pravdepodobnost pokrytia) *Nech* $X_i \sim Bin(N, p_i)$. *Vypočítajte pravdepodobnosti pokrytia:*
(a) vierohodnostného 95% *DIS,*
(b) skóre 95% *DIS,*
(c) spätne tranformovaného Waldovho 95% *DIS pre* $g(p_i)$ *s hranicami* $(d_g^{(i)}, h_g^{(i)})$ *na Waldov* 95% *DIS pre* p_i *s hranicami* $\left((g(d_g^{(i)}))^{-1}, (g(h_g^{(i)}))^{-1}\right)$, *kde (1)* $g(p_i) = \frac{p_i}{1-p_i}$, *(2)* $g(p_i) = \ln \frac{p_i}{1-p_i}$ *a (3)* $g(p_i) = \arcsin(\sqrt{p_i})$
pre každé p_i, *kde* p_i *patria množine* $\mathcal{M}_I = \left\langle \frac{1}{N}, 1 - \frac{1}{N} \right\rangle$, *sú ekvidištantne vzdialené medzi* $\frac{1}{N}$ *a* $1 - \frac{1}{N}$ *a ich počet* $M = 5000$. *Nakreslite obrázok, kde na x-ovej osi budú* p_i *a na y-ovej osi pravdepodobnosť pokrytia* $\Pr_i(pokrytie)$. *Zvoľte (a)* $N = 30$, *(b)* $N = 100$ *a (c)* $N = 1000$ *(pozri obrázky 6.10 až 6.14).*
Pozn.: pravdepodobnosti pokrytia 95% *DIS pre* p_i *vypočítame nasledovne*

$$\Pr_i(pokrytie) = \sum_j \Pr(X = Np_j : p_i \in 95\%\ DIS\ pre\ p_j),$$

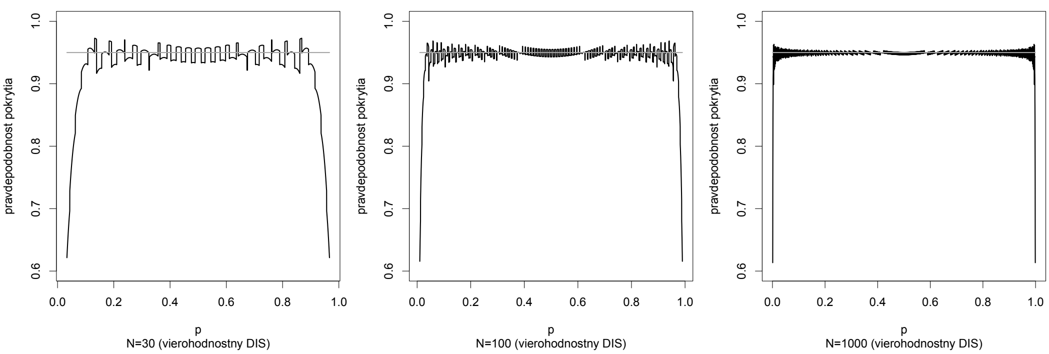
kde $p_j \in \mathcal{M}_J = \left\{\frac{1}{N}, \frac{2}{N}, \ldots, 1 - \frac{1}{N}\right\}$*, t.j. ide o súčet takých funkčných hodnôt pravdepodobnostnej funkcie v bodoch* Np_j*, kde* $p_i \in 95\%$ *DIS pre* p_j*. Pre tie DIS, ktoré majú pre* $p = 0$ *a* $p = 1$ *nenulovú šírku, môžeme použiť* $\mathcal{M}_I = \left\langle \frac{0}{N}, \frac{N}{N} \right\rangle$*.*

Zo simulačnej štúdie vierohodnostného DIS pre p vyplýva (pozri obrázok 6.10), že aktuálne hodnoty pravdepodobnosti pokrytia sú závislé na hodnote p a sú dostatočne blízko nominálnej hodnote $1 - \alpha = 0.95$. Pre veľmi malé a veľké n (resp. veľmi malé a veľké p) vznikajú okrajové efekty výrazného zníženia pravdepodobnosti pokrytia.



Obr. 6.10: Pravdepodobnosť pokrytia skóre 95% empirického DIS pre p pri rôznych N

Zo simulačnej štúdie skóre DIS pre p vyplýva (pozri obrázok 6.11), že aktuálne hodnoty pravdepodobnosti pokrytia sú závislé na hodnote p, ale aj pre malé a veľké n (resp. malé a veľké p) sú dostatočne blízko nominálnej hodnote $1 - \alpha = 0.95$ (t.j. nevznikajú okrajové efekty výrazného zníženia pravdepodobnosti pokrytia).



Obr. 6.11: Pravdepodobnosť pokrytia vierohodnostného 95% empirického DIS pre p pri rôznych N

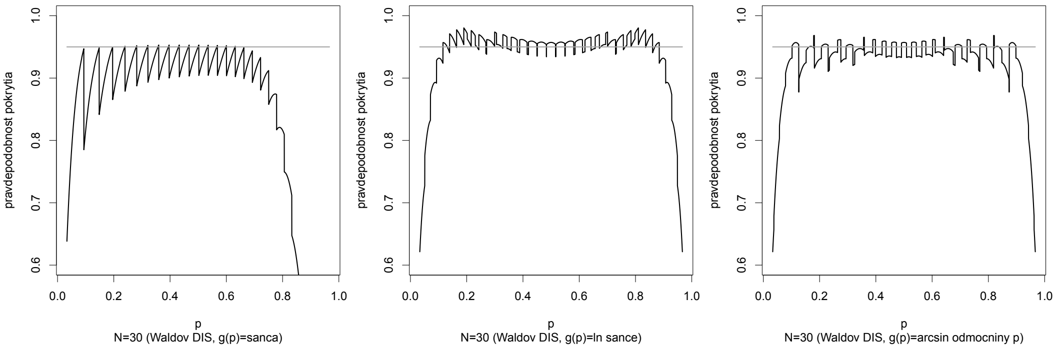
Zo simulačnej štúdie spätne transformovaných Waldových DIS pre p vyplýva (pozri obrázok 6.12 až 6.14), že aktuálne hodnoty pravdepodobnosti pokrytia závisia od tranformácie nasledovne:

1. spätná transformovaného DIS pre šancu – tento DIS je príliš liberálny pre malé N; je liberálny pre stredne veľké N; tento DIS nepoužiteľný pre veľké p, ktoré nám dáva veľkú šancu („použiteľnosť" veľkého p sa zväčšuje s rastúcim N);

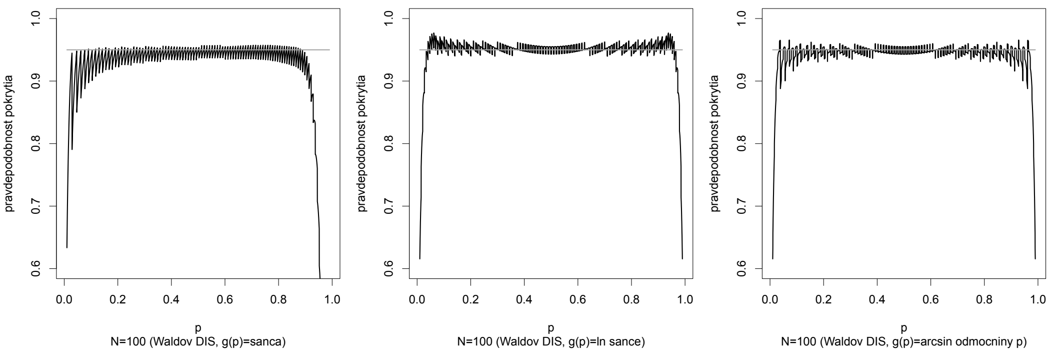
2. spätná transformovaného DIS pre logaritmus šance – aktuálne hodnoty pravdepodobnosti pokrytia tohoto DIS sú dostatočne blízko nominálnej hodnote pre p ďalej od hraníc intervalu $(0, 1)$ už pre malé a stredne veľké N; pre p blízko hraníc intervalu 0 a 1 je mierne konzervatívny nezávisle na N;

3. spätná transformovaného DIS pre arcsin $\sqrt{p}$ – tento DIS je liberálny pre malé N; aktuálne hodnoty pravdepodobnosti pokrytia tohoto DIS sú dostatočne blízko nominálnej hodnote pre p ďalej od hraníc intervalu $(0, 1)$ už pre stredne veľké N; pre p blízko hraníc intervalu 0 a 1 je mierne liberálny nezávisle na N.

Ak je N dostatočne veľké, aktuálne hodnoty pravdepodobnosti pokrytia všetkých troch DIS sú dostatočne blízko nominálnej hodnote pre p takmer na celom intervale $(0, 1)$.



Obr. 6.12: Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p pri $N = 30$; spätne transformovaný DIS pre šancu (vľavo), spätne transformovaný DIS pre logaritmus šance (uprostred) a spätne transformovaný DIS pre arcsin $\sqrt{p}$ (vpravo)



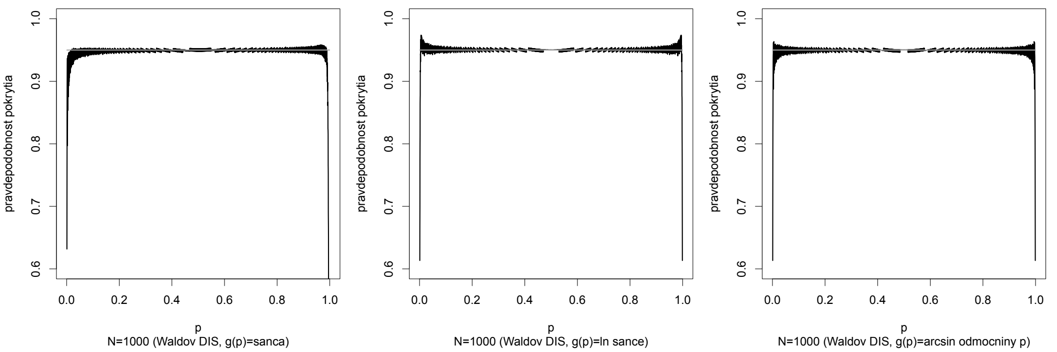
Obr. 6.13: Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p pri $N = 100$; spätne transformovaný DIS pre šancu (vľavo), spätne transformovaný DIS pre logaritmus šance (uprostred) a spätne transformovaný DIS pre arcsin $\sqrt{p}$ (vpravo)

Asymptotické testy o parametri λ Poissonovho rozdelenia. Nech $X \sim Poiss(\lambda)$. Majme dvojice hypotéz $H_{01} : \lambda = \lambda_0$ oproti $H_{11} : \lambda \neq \lambda_0$, $H_{02} : \lambda \leq \lambda_0$ oproti $H_{12} : \lambda > \lambda_0$, $H_{03} : \lambda \geq \lambda_0$ oproti $H_{13} : \lambda < \lambda_0$, ktoré chceme testovať.

Ak H_0 platí, potom

$$Z_W = \frac{\overline{X} - \lambda_0}{\sqrt{\overline{X}/N}} \overset{\mathcal{D}}{\approx} N(0, 1),$$

Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **jednovýberový Z-test o λ**.



Obr. 6.14: Pravdepodobnosť pokrytia Waldovho 95% empirického DIS pre p pri $N = 1000$; spätne transformovaný DIS pre šancu (vľavo), spätne transformovaný DIS pre logaritmus šance (uprostred) a spätne transformovaný DIS pre arcsin $\sqrt{p}$ (vpravo)

Definícia 64 (Kritický obor a silofunkcia pre Z_W test o λ) *Kritický obor a silofunkcia sú definované nasledovne:*

$$\begin{array}{llll} H_0 & H_1 & \mathcal{W} & 1-\beta(\lambda) \\ \lambda=\lambda_0 & \lambda\neq\lambda_0 & \mathcal{W}_1=\{Z_W;|Z_W|\geq u(\alpha/2)\} & 1-\Phi\left(u_{\alpha/2}-\frac{|\lambda_0-\lambda|}{\sqrt{\lambda/N}}\right) \\ \lambda\leq\lambda_0 & \lambda>\lambda_0 & \mathcal{W}_2=\{Z_W;Z_W\geq u(\alpha)\} & 1-\Phi\left(u_{\alpha}+\frac{\lambda_0-\lambda}{\sqrt{\lambda/N}}\right) \\ \lambda\geq\lambda_0 & \lambda<\lambda_0 & \mathcal{W}_3=\{Z_W;Z_W\leq -u(\alpha)\} & 1-\Phi\left(u_{\alpha}-\frac{\lambda_0-\lambda}{\sqrt{\lambda/N}}\right) \end{array}$$

Definícia 65 (p-hodnota pre Z_W test o λ) *Nech Z_W je **nejaká testovacia štatistika** a $z_W = \frac{\overline{x}-\lambda_0}{\sqrt{\overline{x}/N}}$ je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-hodnota} = \begin{cases} 2Pr(Z_W \geq |z_W||H_{01}), \ ak & H_{11}: \lambda\neq\lambda_0 \\ Pr(Z_W \geq z_W|H_{02}), \ ak & H_{12}: \lambda>\lambda_0 \\ Pr(Z_W \leq z_W|H_{03}), \ ak & H_{13}: \lambda<\lambda_0 \end{cases}$$

Definícia 66 (Waldove $100\times(1-\alpha)\%$ empirické IS pre λ) *Waldove $100\times(1-\alpha)\%$ empirické IS pre λ pre všetky tri typy hypotéz majú nasledovný tvar:*

$$\begin{array}{lll} H_0 & H_1 & \textit{hranice } (d,h) \textit{ pre } 100(1-\alpha)\% \textit{ empirický IS} \\ \lambda=\lambda_0 & \lambda\neq\lambda_0 & \mathcal{CS}_{1-\alpha}=\left\{\lambda_0:\lambda_0\in\left(\overline{x}-u_{\alpha/2}\sqrt{\overline{x}/N},\overline{x}+u_{\alpha/2}\sqrt{\overline{x}/N}\right)\right\} \\ \lambda\leq\lambda_0 & \lambda>\lambda_0 & \mathcal{CS}_{1-\alpha}=\left\{\lambda_0:\lambda_0\in\left(\overline{x}-u_{\alpha}\sqrt{\overline{x}/N},\infty\right)\right\} \\ \lambda\geq\lambda_0 & \lambda<\lambda_0 & \mathcal{CS}_{1-\alpha}=\left\{\lambda_0:\lambda_0\in\left(-\infty,\overline{x}+u_{\alpha}\sqrt{\overline{x}/N}\right)\right\} \end{array}$$

Majme H_{11}. **Minimálny rozsah** je definovaný ako

$$N \geq \left(\frac{u_{\alpha/2}+u_\beta}{\lambda-\lambda_0}\right)^2\lambda.$$

Pre H_{12} a H_{13} zameníme v oboch prípadoch $u_{\alpha/2}$ za u_α.

Majme H_{01} oproti H_{11}. Nech $X \sim Poiss(\lambda)$, kde $\theta = \lambda$, potom logaritmus vierohodnostnej funkcie bude mať tvar $l(\lambda|\mathbf{x}) = \sum_{i=1}^{N} x_i \ln\lambda - N\lambda$.

Skóre testovaciu štatistiku U_{S} odvodíme nasledovne

$$U_{\mathrm{S}} = \frac{\left[\frac{\partial}{\partial\lambda}\ln L(\lambda|\mathbf{X})\right]^2}{\mathcal{I}(\lambda)} = \frac{\left(\frac{N\overline{X}}{\lambda} - N\right)^2}{\frac{N\lambda}{\lambda^2}} \overset{\lambda=\lambda_0}{=} \frac{\left(\overline{X} - \lambda_0\right)^2}{\lambda_0/N} \overset{\mathcal{D}}{\sim} \chi_1^2.$$

Skóre $100 \times (1-\alpha)\%$ **empirický DIS pre** λ má tvar $\mathcal{CS}_{1-\alpha} = \left\{|Z_S| \le u_{\alpha/2}\right\}$, kde $Z_S = \frac{\overline{X}-\lambda_0}{\sqrt{\lambda_0/N}}$. Po niekoľkých algebraických úpravách môžeme $\mathcal{CS}_{1-\alpha}$ prepísať do kvadratickej nerovnice

$$\mathcal{CS}_{1-\alpha} = \left\{\lambda_0 : \lambda_0^2 - (2\overline{x} + u_{\alpha/2}^2/N)\lambda_0 + \overline{x}^2 \le 0\right\}.$$

Korene tejto nerovnice definujú hranice DIS nasledovne

$$\frac{2\overline{x} + u_{\alpha/2}^2/N \pm \sqrt{(2\overline{x} + u_{\alpha/2}^2/N)^2 - 4\overline{x}^2}}{2}.$$

Po ďalších úpravách dostaneme DIS v tvare

$$\overline{x} + \frac{1}{2}\frac{u_{\alpha/2}^2}{N} \pm u_{\alpha/2}\sqrt{\frac{1}{N}\left(\overline{x} + \frac{u_{\alpha/2}^2}{4N}\right)}.$$

Test pomerom vierohodnosti pre λ.

MLE θ je rovný $\widehat{\lambda} = \overline{x}$, t.j. $\Theta_1 = \{\theta : \lambda \ne \lambda_0\}$. Za platnosti H_{01} je $\theta_0 = \lambda_0$, t.j. $\Theta_0 = \{\theta : \lambda = \lambda_0\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$-\ln(\lambda(\mathbf{x})) \;=\; l\left(\widehat{\lambda}|\mathbf{x}\right) - l(\lambda_0|\mathbf{x}) = (N\overline{x}\ln\overline{x} - N\overline{x}) - (N\overline{x}\ln\lambda_0 - N\lambda_0) = N\left(\overline{x}\ln\frac{\overline{x}}{\lambda_0} - \overline{x} + \lambda_0\right).$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln(\lambda(\mathbf{X})) \overset{\mathcal{D}}{\sim} \chi_1^2$.

Vierohodnostný $100 \times (1-\alpha)\%$ **empirický DIS pre** λ bude mať tvar $\mathcal{CS}_{1-\alpha} = \left\{\lambda_0 : U_{\mathrm{LR}} \le \chi_1^2(\alpha)\right\}$.

Príklad 223 (pruské armádne jednotky) *Majme $X \sim Poiss(\lambda)$; pozri príklad 78, tabuľka 2.7. Vypočítajte (a) Waldov 95% DIS pre λ, (b) skóre 95% DIS pre λ a (c) vierohodnostný 95% DIS pre λ.*

Riešenie v R

Výpočet $\widehat{\lambda}$ a $\widehat{Var[\widehat{\lambda}]}$ pozri v príkladoch 131 a 134. Waldov 95% DIS pre λ: $(d,h) = (0.502, 0.718)$. Skóre 95% DIS pre λ: $(d,h) = (0.511, 0.728)$. Vierohodnostný 95% DIS pre λ: $(d,h) = (0.508, 0.725)$.

```
m <- c(109,65,22,3,1,0); n <- c(0,1,2,3,4,5); N <- sum(m)
lambda.hat <- sum(n*m)/N  # 0.61
var.lambda.hat <- lambda.hat/N  # 0.003057
IS.W <- lambda.hat+c(-1,1)*qnorm(0.975)*sqrt(var.lambda.hat)
IS.S <- lambda.hat+1/2*qnorm(0.975)^2/N+c(-1,1)*qnorm(0.975)*
        sqrt(1/N*(lambda.hat+qnorm(0.975)^2/(4*N)))
IS.W # 0.5017575 0.7182425
IS.S # 0.5109359 0.7282714
min.lambda <- lambda.hat-3*sqrt(var.lambda.hat)
max.lambda <- lambda.hat+3*sqrt(var.lambda.hat)
lambda.0i <- seq(min.lambda,max.lambda,by=0.0001)
uLR.i <- 2*N*(lambda.hat*log(lambda.hat/lambda.0i)-lambda.hat+lambda.0i)
plot(lambda.0i,uLR.i)
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
IS.LR <- range(lambda.0i[which(uLR.i<lr.krit.hodn)])
IS.LR # 0.5081196 0.7247196
```

7 Testovanie hypotéz o dvoch parametroch

V tejto kapitole sa budeme venovať testovaniu hypotéz o rozdiele stredných hodnôt $\mu_1 - \mu_2$ za predpokladov normality, t.j. $X_j \sim N(\mu_j, \sigma_j^2),\ j = 1, 2,\ \overline{X}_1 - \overline{X}_2 \sim N\left(\mu_1 - \mu_2, \sigma_D^2\right)$, o podiele rozptylov σ_1^2/σ_2^2 za predpokladov normality, t.j. $X_j \sim N(\mu_j, \sigma_j^2),\ j = 1, 2$, o rozdiele korelačných koeficientov $\rho_1 - \rho_2$, za predpokladov normality, t.j. $(X_j, Y_j)^T \sim N_2\left(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j\right),\ j = 1, 2$, a o rozdiele pravdepodobností $p_1 - p_2$, za prepokladov $X \sim Bin\,(N_1, p_1)$ a $Y \sim Bin\,(N_2, p_2)$.

7.1 Asymptotické testy o rozdiele stredných hodnôt

Testovanie rozdielov dvoch stredných hodnôt spojitej premennej patrí medzi najčastejšie úlohy, s ktorými sa biologický antropológ stretáva. Základné zdroje dát totiž v antropológii často tvoria normálne rozdelené metrické premenné (rozmery) na kontinuálnej (spojitej) škále, ktorých variabilita je multifaktoriálne podmienená a z genetického hľadiska predstavujú polygénne znaky. Nemusí ísť zďaleka len o rozmery tela, ale o rôzne ukazovatele fyzického výkonu (sila stisku ruky), psychických vlastností (skóre v psychologickom dotazníku), spôsobe správania (reakčná doba na určitý podnet) i iných vlastností. Typickou úlohou je porovnanie stredných hodnôt výberov dvoch prirodzených alebo zámerne definovaných skupín danej populácie: muži vs. ženy (*sexuálny dimorfizmus*), Poliaci vs. Nemci (*etnický rozdiel*), dvadsaťroční v 60. rokoch 20. storočia vs. dvadsaťroční v 80. rokoch 20. storočia (*rozdiel vekových kohort*, *sekulárny trend*), nižšia sociálna vrstva vs. stredná sociálna vrstva (*sociálne rozdiely*), vysokoškoláci vs. stredoškoláci (*rozdiely v stupni vzdelania*), prvorodičky vs. druhorodičky (*rozdiely v parite*), kresťania vs. moslimovia (*rozdiely vo vyznaní*), sila pravej paže vs. sila paže ľavej (*stranová asymetria sily*) atď. Vzhľadom na štatistickú kvalitu tohto typu premenných (najväčšie množstvo informácií) a súčasne univerzálnosť, prirodzenosť a jednoduchosť tohto typu porovnávania/uvažovania (jedna skupina voči druhej z hľadiska strednej hodnoty veľkosti daného znaku) je tento typ testov vôbec jeden z najdôležitejších vo fyzickej antropológii.

Množstvo foriem/úrovní rôznych dvojíc viac či menej prirodzených skupín, ktorých stredné hodnoty možno porovnávať a rozdiel testovať, otvára priestor pre celú škálu situácií, ktoré môžu nastať. Dané dve skupiny sa nemusia zhodovať v rozptyle, nemusia mať zhodné rozdelenie hodnôt, môžu byť zaťažené chybou merania v rôznej miere, môžu pochádzať z častí populácie s veľmi odlišnou celkovou veľkosťou (praváci vs. ľaváci; homosexuáli vs. heterosexuáli; matky s jedným dieťaťom vs. matky s desiatimi deťmi). Tieto rozdiely v ďalších ohľadoch problematizujú symetriu medzi testovanými skupinami z hľadiska vzorkovania a ekvivalencie porovnávania po stránke rovnocennosti oboch vzoriek.

Nech $X_j \sim N(\mu_j, \sigma_j^2), j = 1, 2$, sú dve nezávislé náhodné premenné. Potom

$$\overline{X}_1 - \overline{X}_2 \stackrel{\mathcal{D}}{\sim} N\left(\mu_1 - \mu_2, \sigma_D^2\right).$$

Majme dvojice hypotéz $H_{01} : \mu_1 - \mu_2 = \mu_0$ oproti $H_{11} : \mu_1 - \mu_2 \neq \mu_0$, $H_{02} : \mu_1 - \mu_2 \leq \mu_0$ oproti $H_{12} : \mu_1 - \mu_2 > \mu_0$, $H_{03} : \mu_1 - \mu_2 \geq \mu_0$ oproti $H_{13} : \mu_1 - \mu_2 < \mu_0$.

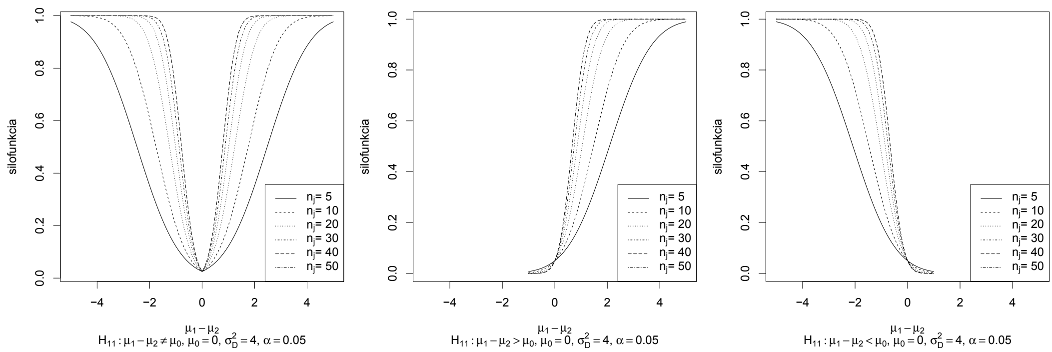
Nech σ_j^2 sú známe. Ak H_0 platí, potom

$$Z_W = \frac{\overline{X}_1 - \overline{X}_2 - \mu_0}{\sigma_D} \stackrel{\mathcal{D}}{\sim} N\left(0, 1\right),$$

kde $\sigma_D^2 = \sigma_1^2/n_1 + \sigma_2^2/n_2$. V špeciálnom prípade $\sigma^2 = \sigma_1^2 = \sigma_2^2$ a môžeme písať $\sigma_D^2 = \sigma^2/n_1 + \sigma^2/n_2 = \frac{n_1+n_2}{n_1 n_2}\sigma^2$. Navyše v niektorých praktických situáciách môže nastať $n = n_1 = n_2$, kde $\sigma_D^2 = \frac{1}{n}\left(\sigma_1^2 + \sigma_2^2\right)$, keď nepredpokladáme rovnosť rozptylov, resp. $\sigma_D^2 = \frac{2n}{n^2}\sigma^2 = \frac{2}{n}\sigma^2$, keď rovnosť rozptylov predpokladáme. Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **dvojvýberový** Z**-test o rozdiele stredných hodnôt** $\mu_1 - \mu_2$.

Definícia 67 (Kritický obor a silofunkcia Z_W testu o $\mu_1 - \mu_2$) *Kritický obor a silofunkcia sú definované (za predpokladu $\sigma_1 = \sigma_2$ a $n_1 = n_2$) nasledovne (pozri obrázok 7.1):*

H_0	H_1	$\mathcal{W}$	$1-\beta(\mu_1,\mu_2)$
$\mu_1-\mu_2=\mu_0$	$\mu_1-\mu_2\neq\mu_0$	$\mathcal{W}_1=\{Z_W;\lvert Z_W\rvert\geq u(\alpha/2)\}$	$1-\Phi\left(u_{\alpha/2}-\frac{\lvert\mu_0-(\mu_1-\mu_2)\rvert}{\sigma_D}\right)$
$\mu_1-\mu_2\leq\mu_0$	$\mu_1-\mu_2>\mu_0$	$\mathcal{W}_2=\{Z_W;Z_W\geq u(\alpha)\}$	$1-\Phi\left(u_{\alpha}+\frac{\mu_0-(\mu_1-\mu_2)}{\sigma_D}\right)$
$\mu_1-\mu_2\geq\mu_0$	$\mu_1-\mu_2<\mu_0$	$\mathcal{W}_3=\{Z_W;Z_W\leq -u(\alpha)\}$	$1-\Phi\left(u_{\alpha}-\frac{\mu_0-(\mu_1-\mu_2)}{\sigma_D}\right)$



Obr. 7.1: Silofunkcie Waldovho testu o $\mu_1 - \mu_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 68 (p-hodnota Z_W testu o $\mu_1 - \mu_2$) *Nech Z_W je* ***nejaká testovacia štatistika*** *a $z_W = \frac{\overline{x}_1-\overline{x}_2-\mu_0}{\sigma_D}$ je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$\textit{p-hodnota} = \begin{cases} 2Pr(Z_W \geq |z_W||H_{01}),\ ak & H_{11}: \mu_1-\mu_2 \neq \mu_0 \\ Pr(Z_W \geq z_W|H_{02}),\ ak & H_{12}: \mu_1-\mu_2 > \mu_0 \\ Pr(Z_W \leq z_W|H_{03}),\ ak & H_{13}: \mu_1-\mu_2 < \mu_0 \end{cases}$$

Definícia 69 (Waldove $100\times(1-\alpha)\%$ empirické IS pre $\mu_1 - \mu_2$) *Waldove $100\times(1-\alpha)\%$ empirické IS pre $\mu_1 - \mu_2$ pre všetky tri typy hypotéz majú nasledovný tvar:*

H_0	H_1	*hranice (d,h) pre $100(1-\alpha)\%$ empirický IS*
$\mu=\mu_0$	$\mu\neq\mu_0$	$\mathcal{CS}_{1-\alpha}=\{\mu_0:\mu_0\in\left(\overline{x}_1-\overline{x}_2-u_{\alpha/2}\sigma_D,\overline{x}_1-\overline{x}_2+u_{\alpha/2}\sigma_D\right)\}$
$\mu\leq\mu_0$	$\mu>\mu_0$	$\mathcal{CS}_{1-\alpha}=\{\mu_0:\mu_0\in(\overline{x}_1-\overline{x}_2-u_{\alpha}\sigma_D,\infty)\}$
$\mu\geq\mu_0$	$\mu<\mu_0$	$\mathcal{CS}_{1-\alpha}=\{\mu_0:\mu_0\in(-\infty,\overline{x}_1-\overline{x}_2+u_{\alpha}\sigma_D)\}$

Majme H_{11}. Zvoľme rozdiel $|\mu_1 - \mu_2| = c\sigma_D$. **Minimálny rozsah** n_1 definujeme nasledovne:

$$n_1 \geq \left(\frac{u_{\alpha/2}+u_\beta}{c}\right)^2 = \left(\frac{u_{\alpha/2}+u_\beta}{\mu_1-\mu_2-\mu_0}\right)^2\left(\sigma_1^2+\frac{\sigma_2^2}{k}\right), n_2 = kn_1.$$

Pre H_{12} a H_{13} zameníme v oboch prípadoch $u_{\alpha/2}$ za u_α.

Príklad 224 (minimálny rozsah súboru) *Predpokladajme, že v roku 1986 sme zistili priemernú výšku 10-ročných dievčat ako $\overline{x}_{1986} = 136.1$ cm. V roku 1996 už bol $\overline{x}_{1996} = 139.225$ cm. Ďalej predpokladajme, že rozptyl $\sigma^2 = 6.25^2$. Pokiaľ by sme v roku 2006 v porovnaní s rokom 1996 chceli rozdiel*

$139.225 - 136.100 = 3.125$ *odhaliť s pravdepodobnosťou* 0.90, *teda s* $\beta = 0.1$ *na hladine významnosti* $\alpha = 0.05$, *aký minimálny rozsah n musí mať náhodný výber?*

Riešenie v

```
k <- 1
n1 <- ceiling(((qnorm(0.975)+qnorm(0.9))/(3.125))^2*(6.25^2+6.25^2/k)) # 85
```

Nech σ_j^2 sú neznáme. Ak H_0 platí, potom

$$T_W = \frac{\overline{X}_1 - \overline{X}_2 - \mu_0}{S_D} \overset{\mathcal{D}}{\sim} t_{df},$$

a rozdelenie t_{df} sa nazýva **centrálne t-rozdelenie s *df* stupňami voľnosti**. Pre $\sigma_1^2 = \sigma_2^2$ platí $df = n_1 + n_2 - 2$ a pre $\sigma_1^2 \neq \sigma_2^2$ platí

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}},$$

$S_D^2 = S_1^2/n_1 + S_2^2/n_2$, S_1^2 a S_2^2 sú výberové rozptyly. V špeciálnom prípade predpokladáme $S_1^2 = S_2^2$ a môžeme písať $S_D^2 = \frac{n_1+n_2}{n_1 n_2} S^2$, kde $S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$. Navyše v niektorých praktických situáciách môže nastať $n = n_1 = n_2$, kde $S_D^2 = \frac{1}{n}\left(S_1^2 + S_2^2\right)$, keď nepredpokladáme rovnosť rozptylov, resp. $S_D^2 = \frac{2}{n} S^2$, keď rovnosť rozptylov predpokladáme. T_W sa nazýva **Waldova testovacia štatistika** (alebo **t-štatistika**) a test **dvojvýberový Studentov t-test o rozdiele stredných hodnôt $\mu_1 - \mu_2$**. Za predpokladu rovnosti rozptylov hovoríme o **klasickom dvojvýberovom t-teste**. Za prepokladu nerovnosti rozptylov, hovoríme o **Welchovom dvojvýberovom t-teste** (alebo **dvojvýberovom Studentovom t-teste s Welchovou aproximáciou stupňov voľnosti**) a problému nerovnosti rozptylov hovoríme **Behrens-Fisherov problém** (Welch, 1947). Ak $n_1 = n_2$ a $\sigma_1^2 = \sigma_2^2 = \sigma^2$ sú neznáme, potom Welchova aproximácia *df* je rovná $2n - 2$ a Welchov t-test je ekvivalentný Studentovmu t-testu.

Ak H_0 neplatí, náhodné výbery sú vyvážené ($n_1 = n_2 = n$) a rozptyly rovnaké ($\sigma_1^2 = \sigma_2^2 = \sigma^2$), táto situácia vedie k necentrálnemu t-rozdeleniu s parametrom necentrality $\lambda = \delta / \left(\sqrt{2}\sigma\right)\sqrt{n}$, kde $\delta = \mu_1 - \mu_2 - \mu_0$ je minimálne detegovateľná vzdialenosť medzi μ_1 a μ_2. Pri $n_1 \neq n_2$ a $\sigma_1^2 \neq \sigma_2^2$, táto situácia vedie tiež k necentrálnemu t-rozdeleniu, kde parameter necentrality $\lambda = \delta / \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$.

Ak predpokladáme $n_1 \neq n_2$ ($n_1 + n_2 = n$), predefinujme T_W nasledovne

$$T_W = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\overline{X}_1 - \overline{X}_2 - \mu_0}{S}\right) = \sqrt{n \times w_1 w_2} \left(\frac{\overline{X}_1 - \overline{X}_2 - \mu_0}{S}\right) \overset{\mathcal{D}}{\sim} t_{df},$$

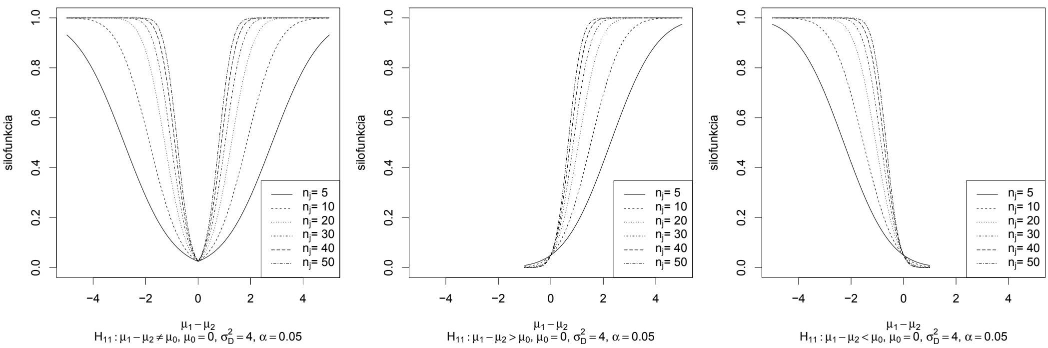
kde $\frac{n_1 n_2}{n_1 + n_2} = \frac{n w_1 \times n w_2}{n} = n \times w_1 w_2$, $w_j = n_j / n, j = 1, 2$. Ak H_0 neplatí, táto situácia vedie k **necentrálnemu t-rozdeleniu** s *df* stupňami voľnosti a parametrom necentrality λ, ozn. $t_{df,\lambda}$, kde

$$T_{W,\lambda} = \sqrt{n \times w_1 w_2} \left(\frac{\overline{X}_1 - \overline{X}_2 - \mu}{S}\right) \overset{\mathcal{D}}{\sim} t_{df,\lambda},$$

kde parameter necentrality $\lambda = \sqrt{n \times w_1 w_2}\,(\delta/\sigma)$ a $\delta = \mu_1 - \mu_2 - \mu_0$ je **minimálne detegovateľná vzdialenosť** medzi rozdielom $\mu_1 - \mu_2$ a μ_0.

Definícia 70 (Kritický obor a silofukcia T_W testu o $\mu_1 - \mu_2$) *Kritický obor a silofukcia sú definované (za platnosti $\sigma_1 = \sigma_2$ a $n_1 = n_2$) nasledovne (pozri obrázok 7.2):*

H_0	H_1	$\mathcal{W}$	$1 - \beta\,(\mu_1, \mu_2)$
$\mu_1 - \mu_2 = \mu_0$	$\mu_1 - \mu_2 \neq \mu_0$	$\mathcal{W}_1 = \{T_W; \lvert T_W \rvert \geq t_{df}\,(\alpha/2)\}$	$\Pr\left(F_{1,df,\lambda^2} \geq F_{1,df}\,(\alpha)\right)$
$\mu_1 - \mu_2 \leq \mu_0$	$\mu_1 - \mu_2 > \mu_0$	$\mathcal{W}_2 = \{T_W; T_W \geq t_{df}\,(\alpha)\}$	$\Pr\left(t_{df,\lambda} \geq t_{df}\,(\alpha)\right)$
$\mu_1 - \mu_2 \geq \mu_0$	$\mu_1 - \mu_2 < \mu_0$	$\mathcal{W}_3 = \{T_W; T_W \leq -t_{df}\,(\alpha)\}$	$\Pr\left(t_{df,\lambda} \leq t_{df}\,(1 - \alpha)\right)$



Obr. 7.2: Silofunkcie Waldovho testu o $\mu_1 - \mu_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 71 (p-hodnota T_W testu o $\mu_1 - \mu_2$) *Nech T_W je* ***nejaká testovacia štatistika*** *a $t_W = \frac{\overline{x}_1 - \overline{x}_2 - \mu_0}{s_D}$ je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$\textit{p-hodnota} = \begin{cases} 2Pr(T_W \geq |t_W| | H_{01}), \ ak & H_{11}: \mu_1 - \mu_2 \neq \mu_0 \\ Pr(T_W \geq t_W | H_{02}), \ ak & H_{12}: \mu_1 - \mu_2 > \mu_0 \\ Pr(T_W \leq t_W | H_{03}), \ ak & H_{13}: \mu_1 - \mu_2 < \mu_0 \end{cases}$$

Definícia 72 ($100 \times (1-\alpha)\%$ empirické IS pre $\mu_1 - \mu_2$) $100 \times (1-\alpha)\%$ *empirické IS pre $\mu_1 - \mu_2$ pre všetky tri typy hypotéz majú nasledovný tvar*

H_0	H_1	*hranice* (d, h) *pre* $100(1-\alpha)\%$ *empirický IS*
$\mu = \mu_0$	$\mu \neq \mu_0$	$\mathcal{CS}_{1-\alpha} = \{\mu_0 : \mu_0 \in (\overline{x}_1 - \overline{x}_2 - t_{df}(\alpha/2)s_D, \overline{x}_1 - \overline{x}_2 + t_{df}(\alpha/2)s_D)\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\mathcal{CS}_{1-\alpha} = \{\mu_0 : \mu_0 \in (\overline{x}_1 - \overline{x}_2 - t_{df}(\alpha)s_D, \infty)\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\mathcal{CS}_{1-\alpha} = \{\mu_0 : \mu_0 \in (-\infty, \overline{x}_1 - \overline{x}_2 + t_{df}(\alpha)s_D)\}$

Dvojvýberový Studentov t-test o rozdiele stredných hodnôt $\mu_1 - \mu_2$ v R (funkcia **t.test()**). Argumenty (vstupy) funkcie:

1. vektory dát x a y;
2. alternatíva alternative="two.sided" je prednastavená, ďalšie voľby sú "greater", "less";
3. rovnosť rozptylov var.equal=TRUE alebo var.equal=FALSE;
4. stredná hodnota za platnosti nulovej hypotézy mu, t.j. μ_0;
5. spoľahlivosť conf.level=p, prednastavené je p =.95.

Výstupy funkcie:

1. názov použitého testu method;
2. testovacia štatistika statistic;
3. počet stupňov voľnosti df parameter;
4. p-hodnota p.value;
5. alternatívna hypotéza alternative hypothesis;

6. interval spoľahlivosti rozdielu $\mu_1 - \mu_2$ `conf.int`;

7. bodové odhady (aritmetické priemery $\overline{x}_1$ a $\overline{x}_2$) `sample estimates`.

Príklad 225 (pravdepodobnosť pokrytia ak $\sigma_1^2 = \sigma_2^2$ sú neznáme) *Nech $X_j \sim N(\mu_j, \sigma^2)$, kde $j = 1, 2$, $\mu_1 = 20$, $\mu_2 = 35$ a $\sigma^2 = 100$. Pomocou simulačnej štúdie ($M = 100000$) vypočítajte pravdepodobnosť pokrytia 95% DIS pre $\mu_1 - \mu_2$ ako podiel $\sum_{m=1}^{M} I\left(t_{W,m} < t_{df}(\alpha/2)\right)/M$, kde $t_{W,m}$ sú testovacie štatistiky klasického dvojvýberového t-testu. Zvoľte (a) $n = 5$, (b) $n = 50$ a (c) $n = 100$.*

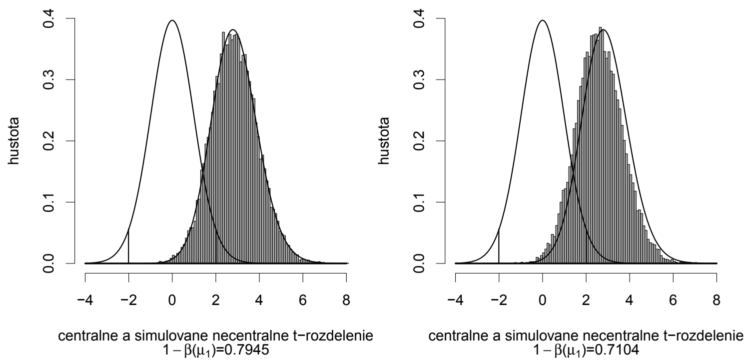
Príklad 226 (pravdepodobnosť pokrytia ak $\sigma_1^2 \neq \sigma_2^2$ sú neznáme) *Nech $X_j \sim N(\mu_j, \sigma_j^2)$, kde $j = 1, 2$, $\mu_1 = 20$, $\mu_2 = 35$, $\sigma_1^2 = 100$ a $\sigma_2^2 = 150$. Pomocou simulačnej štúdie ($M = 100000$) vypočítajte pravdepodobnosť pokrytia 95% DIS pre $\mu_1 - \mu_2$ ako podiel $\sum_{m=1}^{M} I\left(t_{W,m} < t_{df}(\alpha/2)\right)/M$, kde $t_{W,m}$ sú (1) testovacie štatistiky klasického dvojvýberového t-testu a (2) testovacie štatistiky Welchovho dvojvýberového t-testu. Zvoľte (a) $n = 5$, (b) $n = 50$ a (c) $n = 100$.*

Príklad 227 (pravdepodobnosť pokrytia ak $\sigma_1^2 = \sigma_2^2$ sú neznáme) *Nech $X_j \sim [pN(\mu_j, \sigma^2) + (1-p)N(\mu_j, \sigma_a^2)]$, kde $p = 0.9$, $j = 1, 2$, $\mu_1 = 20$, $\mu_2 = 35$, $\sigma^2 = 100$ a $\sigma_a^2 = 400$. Pomocou simulačnej štúdie ($M = 100000$) vypočítajte pravdepodobnosť pokrytia 95% DIS pre $\mu_1 - \mu_2$ ako podiel $\sum_{m=1}^{M} I\left(t_{W,m} < t_{df}(\alpha/2)\right)/M$, kde $t_{W,m}$ sú testovacie štatistiky klasického dvojvýberového t-testu. Zvoľte (a) $n = 5$, (b) $n = 50$ a (c) $n = 100$.*

Príklad 228 (pravdepodobnosť pokrytia ak $\sigma_1^2 \neq \sigma_2^2$ sú neznáme) *Nech $X_j \sim [pN(\mu_j, \sigma_j^2) + (1-p)N(\mu_j, \sigma_{ja}^2)]$, kde $p = 0.9$, $j = 1, 2$, $\mu_1 = 20$, $\mu_2 = 35$, $\sigma_1^2 = 100$, $\sigma_2^2 = 150$, $\sigma_{1a}^2 = 400$ a $\sigma_{2a}^2 = 450$. Pomocou simulačnej štúdie ($M = 100000$) vypočítajte pravdepodobnosť pokrytia 95% DIS pre $\mu_1 - \mu_2$ ako podiel $\sum_{m=1}^{M} I\left(t_{W,m} < t_{df}(\alpha/2)\right)/M$, kde $t_{W,m}$ sú (1) testovacie štatistiky klasického dvojvýberového t-testu a (2) testovacie štatistiky Welchovho dvojvýberového t-testu. Zvoľte (a) $n = 5$, (b) $n = 50$ a (c) $n = 100$.*

Príklad 229 (sila a silofunkcia testu rozdielu stredných hodnôt) *Použite R na simuláciu hustoty rozdelenia testovacej štatistiky $T_{df,\lambda}$ dvojvýberového testu rozdielu stredných hodnôt $\mu_1 - \mu_2$ za platnosti dvojstrannej alternatívy H_{11} pri $M = 20000$ opakovaniach. Túto hustotu v podobe histogramu v relatívnej škále zakreslite do obrázka a superponujte ju s teoretickou hustotou. Vypočítajte silu za platnosti alternatívy $H_{11} : \mu_1 - \mu_2 = 2$, kde $n_1 = n_2 = 25$. Použite (a) klasický dvojvýberový t-test a (b) Welchov dvojvýberový t-test. (1) $X_1 \sim N(4, 2.5^2)$, $X_2 \sim N(2, 2.5^2)$ a (2) $X_1 \sim [pN(4, 2.5^2) + (1-p)N(4, 4.5^2)]$, $X_2(2, 4.5^2)$, kde $p = 0.9$.*

Riešenie (a) pozri na obrázku 7.3.



Obr. 7.3: Hustota centrálneho a necentrálneho t-rozdelenia; vľavo – hustota necentrálneho t-rozdelenia je superponovaná histogramom simulácií pre $X_1 \sim N(4, 2.5^2)$, $X_2 \sim N(2, 2.5^2)$ a vpravo – pre $X_1 \sim [pN(4, 2.5^2) + (1 - p)N(4, 4.5^2)]$, $X_2 \sim N(2, 4.5^2)$, kde $p = 0.9$

Test pomerom vierohodnosti pre $\mu_1 - \mu_2$.
Budeme rozlišovať tri základné situácie – (1) σ_1^2, σ_2^2 sú rôzne a známe, (2) σ_1^2, σ_2^2 sú rovnaké a neznáme a (3) σ_1^2, σ_2^2 sú rôzne a neznáme.

Nech $X_1 \sim N(\mu_1, \sigma_1^2)$ a $X_2 \sim N(\mu_2, \sigma_2^2)$, kde $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)^T$ a σ_1^2, σ_2^2 sú rôzne a známe. Odhad $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \sigma_1^2, \sigma_2^2)^T$, kde $\widehat{\mu}_1 = \overline{x}_1, \widehat{\mu}_2 = \overline{x}_2$, t.j. $\Theta_1 = \{\boldsymbol{\theta} : \mu_1 \neq \mu_2\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\widehat{\mu}, \widehat{\mu}, \sigma_1^2, \sigma_2^2)^T$, t.j. $\Theta_0 = \{\boldsymbol{\theta} : \mu_1 = \mu_2\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$
\begin{aligned}
-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{x}_2) - l(\boldsymbol{\theta}_0|\mathbf{x}_1, \mathbf{x}_2) \\
&= \left(-\frac{1}{2\sigma_1^2} \sum_{i=1}^{n_1} (x_{1i} - \overline{x}_1)^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^{n_2} (x_{2j} - \overline{x}_2)^2 \right) \\
&- \left(-\frac{1}{2\sigma_1^2} \sum_{i=1}^{n_1} (x_{1i} - \widehat{\mu})^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^{n_2} (x_{2j} - \widehat{\mu})^2 \right) \\
&= \frac{1}{2\sigma_1^2} \left(\sum_{i=1}^{n_1} (x_{1i} - \widehat{\mu})^2 - \sum_{i=1}^{n_1} (x_{1i} - \overline{x}_1)^2 \right) \\
&+ \frac{1}{2\sigma_2^2} \left(\sum_{j=1}^{n_2} (x_{2j} - \widehat{\mu})^2 - \sum_{j=1}^{n_2} (x_{2j} - \overline{x}_2)^2 \right),
\end{aligned}
$$

kde

$$\widehat{\mu} = \frac{\sum_{i=1}^{n_1} x_{1i} + \gamma \sum_{j=1}^{n_2} x_{2j}}{n_1 + \gamma n_2}, \text{kde } \gamma = \frac{\sigma_1^2}{\sigma_2^2}.$$

Použitím rovnosti

$$\sum_{i=1}^{n_1} (x_{1i} - \widehat{\mu})^2 = \sum_{i=1}^{n_1} (x_{1i} - \overline{x}_1)^2 + n_1(\overline{x}_1 - \widehat{\mu})^2 \text{ a } \sum_{j=1}^{n_2} (x_{2j} - \widehat{\mu})^2 = \sum_{j=1}^{n_2} (x_{2j} - \overline{x}_2)^2 + n_2(\overline{x}_2 - \widehat{\mu})^2$$

dostaneme

$$-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) = \frac{n_1}{2\sigma_1^2} (\widehat{\mu} - \overline{x}_1)^2 + \frac{n_2}{2\sigma_2^2} (\widehat{\mu} - \overline{x}_2)^2.$$

Potom

$$\widehat{\mu} - \overline{x}_1 = \frac{n_1 \sum_{i=1}^{n_1} x_{1i} + n_1 \gamma \sum_{j=1}^{n_2} x_{2j} - n_2 \sum_{i=1}^{n_1} x_{1i} - \gamma n_2 \sum_{j=1}^{n_2} x_{2j}}{n_1(n_1 + \gamma n_2)} = \frac{\gamma n_2 (\overline{x}_2 - \overline{x}_1)}{n_1 + \gamma n_2}$$

a

$$\widehat{\mu} - \overline{x}_2 = \frac{n_1(\overline{x}_1 - \overline{x}_2)}{n_1 + \gamma n_2}.$$

Nakoniec

$$
\begin{aligned}
-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) &= \frac{n_1}{2\sigma_1^2} \frac{\gamma^2 n_2^2 (\overline{x}_2 - \overline{x}_1)^2}{(n_1 + \gamma n_2)^2} + \frac{n_2}{2\sigma_2^2} \frac{n_1^2 (\overline{x}_2 - \overline{x}_1)^2}{(n_1 + \gamma n_2)^2} \\
&= \frac{n_1 n_2 (\overline{x}_2 - \overline{x}_1)^2}{2(n_1 + \gamma n_2)^2} \left(\frac{\gamma^2 n_2}{\sigma_1^2} + \frac{n_1}{\sigma_2^2} \right).
\end{aligned}
$$

Ak položíme $\gamma = 1$, dostaneme

$$-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) = \frac{n_1 n_2 (\overline{x}_2 - \overline{x}_1)^2}{2n^2} \left(\frac{n_2}{\sigma_1^2} + \frac{n_1}{\sigma_2^2} \right).$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2 \ln \left(\lambda \left(\mathbf{X}_1, \mathbf{X}_2 \right) \right) \overset{\mathcal{D}}{\sim} \chi_1^2$, kde H_0 bude zamietnutá pre veľké hodnoty rozdielu $|\overline{x}_2 - \overline{x}_1|$. Dá sa ukázať, že u_{LR} je rastúcou funkciou $|z_W|$.

Nech $X_1 \sim N(\mu_1, \sigma^2)$ a $X_2 \sim N(\mu_2, \sigma^2)$, kde $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma^2)^T$ a σ^2 je neznáma. Logaritmus funkcie vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n_1}(x_{1i} - \mu_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \mu_2)^2\right),$$

kde $n = n_1 + n_2$. MLE $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \widetilde{\sigma}^2)^T$, kde

$$\widehat{\mu}_1 = \overline{x}_1, \widehat{\mu}_2 = \overline{x}_2, \widetilde{\sigma}^2 = \frac{1}{n}\left(\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \overline{x}_2)^2\right),$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \mu_1 \neq \mu_2\}$. Za platnosti H_{01} je $\widehat{\boldsymbol{\theta}}_0 = (\widehat{\mu}, \widehat{\mu}, \widetilde{\sigma}_0^2)^T$ platí

$$\widehat{\mu} = \frac{1}{n}\left(\sum_{i=1}^{n_1} x_{1i} + \sum_{j=1}^{n_2} x_{2j}\right) = \frac{n_1\overline{x}_1 + n_2\overline{x}_2}{n}$$

a

$$\begin{aligned}\widetilde{\sigma}_0^2 &= \frac{1}{n}\left(\sum_{i=1}^{n_1}(x_{1i} - \widehat{\mu})^2 + \sum_{j=1}^{n_2}(x_{2j} - \widehat{\mu})^2\right) \\ &= \frac{1}{n}\left(\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \overline{x}_2)^2 + \frac{n_1 n_2}{n}\left(\overline{x}_1 - \overline{x}_2\right)^2\right),\end{aligned}$$

kde

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n_1}(x_{1i} - \widehat{\mu})^2 &= \frac{1}{n}\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2 + \frac{n_1}{n}(\overline{x}_1 - \widehat{\mu})^2, \\ \frac{1}{n}\sum_{j=1}^{n_2}(x_{2j} - \widehat{\mu})^2 &= \frac{1}{n}\sum_{j=1}^{n_2}(x_{2j} - \overline{x}_2)^2 + \frac{n_2}{n}(\overline{x}_2 - \widehat{\mu})^2\end{aligned}$$

získame z rovnosti $(x_{1i} - \widehat{\mu})^2 = ((x_{1i} - \overline{x}_1) + (\overline{x}_1 - \widehat{\mu}))^2$ (a podobne pre x_{2j}). Tiež platí $\Theta_0 = \{\boldsymbol{\theta} : \mu_1 = \mu_2 = \mu\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) = \frac{n}{2}\ln\left(1 + \frac{\frac{n_1 n_2}{n}\left(\overline{x}_1 - \overline{x}_2\right)^2}{\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \overline{x}_2)^2}\right) = \frac{n}{2}\ln\left(\frac{\widetilde{\sigma}_0^2}{\widetilde{\sigma}^2}\right).$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{X}_1, \mathbf{X}_2\right)\right) \overset{\mathcal{D}}{\sim} \chi_1^2$, kde H_0 bude zamietnutá pre veľké hodnoty podielu $\frac{\widetilde{\sigma}_0^2}{\widetilde{\sigma}^2}$.

Modifikovaním u_{LR}, kde $\widetilde{\sigma}^2$ substituujeme

$$s^2 = \frac{n}{n-2}\widetilde{\sigma}^2 = \frac{1}{n-2}\left(\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2i} - \overline{x}_2)^2\right),$$

dostaneme u_{LR}, ktorá je rastúcou funkciou $|t_W|$ a $|\overline{x}_1 - \overline{x}_2|$. Potom

$$u_{\mathrm{LR}} = n\ln\left(1 + \frac{t_W^2}{n-2}\right).$$

Nech $X \sim N(\mu_1, \sigma_1^2)$ a $Y \sim N(\mu_2, \sigma_2^2)$, kde $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)^T$ a σ_1^2, σ_2^2 sú rôzne a neznáme. Jadro logaritmu funkcie vierohodnosti je rovné

$$l(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = -\frac{1}{2\sigma_1^2}\sum_{i=1}^{n_1}(x_{1i} - \mu_1)^2 - \frac{1}{2\sigma_2^2}\sum_{j=1}^{n_2}(x_{2j} - \mu_2)^2.$$

Nech σ_1^2, σ_2^2 sú neznáme. Odhad $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2)^T$, kde

$$\widehat{\mu}_1 = \overline{x}_1, \widehat{\mu}_2 = \overline{x}_2, \widehat{\sigma}_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2, \widehat{\sigma}_2^2 = \frac{1}{n_2}\sum_{j=1}^{n_2}(x_{2j} - \overline{x}_2)^2,$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \mu_1 \neq \mu_2\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\widehat{\mu}, \widehat{\mu}, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2)^T$, t.j. $\Theta_0 = \{\boldsymbol{\theta} : \mu_1 = \mu_2\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned}
-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{x}_2) - l(\boldsymbol{\theta}_0|\mathbf{x}_1, \mathbf{x}_2) \\
&= \left(-\frac{1}{2\widehat{\sigma}_1^2}\sum_{i=1}^{n_1}(x_{1i} - \overline{x}_1)^2 - \frac{1}{2\widehat{\sigma}_2^2}\sum_{j=1}^{n_2}(x_{2j} - \overline{x}_2)^2\right) \\
&- \left(-\frac{1}{2\widehat{\sigma}_1^2}\sum_{i=1}^{n_1}(x_{1i} - \widehat{\mu})^2 - \frac{1}{2\widehat{\sigma}_2^2}\sum_{j=1}^{n_2}(x_{2j} - \widehat{\mu})^2\right) \\
&= \left(-\frac{n_1\widehat{\sigma}_1^2}{2\widehat{\sigma}_1^2} - \frac{n_2\widehat{\sigma}_2^2}{2\widehat{\sigma}_2^2}\right) - \left(-\frac{n_1\widehat{\sigma}_1^2}{2\widehat{\sigma}_1^2} - \frac{n_2\widehat{\sigma}_2^2}{2\widehat{\sigma}_2^2} - \frac{n_1(\overline{x}_1 - \widehat{\mu})^2}{2\widehat{\sigma}_1^2} - \frac{n_2(\overline{x}_2 - \widehat{\mu})^2}{2\widehat{\sigma}_2^2}\right) \\
&= \frac{n_1(\overline{x}_1 - \widehat{\mu})^2}{2\widehat{\sigma}_1^2} + \frac{n_2(\overline{x}_2 - \widehat{\mu})^2}{2\widehat{\sigma}_2^2},
\end{aligned}$$

kde potom dosadením $\gamma = \widehat{\sigma}_1^2/\widehat{\sigma}_2^2$ a po použití identického postupu ako v prípade, kde σ_1^2 a σ_2^2 sú známe, dostaneme

$$-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) = \frac{n_1 n_2(\overline{x}_2 - \overline{x}_1)^2}{2(n_1 + \gamma n_2)^2}\left(\frac{\gamma^2 n_2}{\widehat{\sigma}_1^2} + \frac{n_1}{\widehat{\sigma}_2^2}\right).$$

Ak položíme $\gamma = 1$, dostaneme

$$-\ln(\lambda(\mathbf{x}_1, \mathbf{x}_2)) = \frac{n_1 n_2(\overline{x}_2 - \overline{x}_1)^2}{2n^2}\left(\frac{n_2}{\widehat{\sigma}_1^2} + \frac{n_1}{\widehat{\sigma}_2^2}\right).$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{X}_1, \mathbf{X}_2\right)\right) \overset{\mathcal{D}}{\sim} \chi_1^2$, kde H_0 bude zamietnutá pre veľké hodnoty rozdielu $|\overline{x}_2 - \overline{x}_1|$. Dá sa ukázať, že u_{LR} je rastúcou funkciou $|t_W|$.

Vierohodnostný $100 \times (1-\alpha)\%$ **empirický DIS pre** $\mu_1 - \mu_2$ bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{\mu_0 : U_{\mathrm{LR}} < \chi_1^2\left(\alpha\right)\right\}.$$

Na výpočet tohoto DIS by sme museli reparametrizovať funkciu vierohodnosti na parameter záujmu $(\mu_1 - \mu_2)$ a rušivý parameter[1]. Alternatívne by bolo možné vypočítať u_{LR} pre rôzne kombinácie μ_1 a μ_2 (na vhodne zvolených intervaloch) a pre tieto $\mu_j,\ j = 1, 2$, vypočítať rozdiely $\mu_1 - \mu_2$. Potom $\mathcal{CS}_{1-\alpha} = \left\{\mu_0 : U_{\mathrm{LR}}(\mu_1 - \mu_2) < \chi_1^2\left(\alpha\right)\right\}$.

[1] Namiesto reparametrizácie môžeme použiť upravenú testovaciu štatistiku u_{LR}, kde $\overline{x}_1 - \overline{x}_2$ substituujeme $\overline{x}_1 - \overline{x}_2 - \mu_0$; μ_0 patrí vhodne zvolenému intervalu, na ktorom hľadáme μ_0 vyhovujúce nerovnosti $U_{\mathrm{LR}}(\mu_0) < \chi_1^2\left(\alpha\right)$.

Príklad 230 (test o rozdiele stredných hodnôt μ_1 a μ_2) *Majme dáta* two-samples-means-birth.txt, *premennú pôrodná hmotnosť novorodencov (chlapcov, narodených v krajskej nemocnici v priebehu jedného roka)* birth.W *v gramoch a premennú počet starších súrodencov* o.sib.N, *ktorá nadobúda hodnoty* 0 *(žiadny) a* 1 *(jeden). Predpokladáme, že premenná* birth.W *chlapcov bez staršieho súrodenca má normálne rozdelenie $N(\mu_1, \sigma_1^2)$ a* birth.W *chlapcov s jedným starším súrodencom má normálne rozdelenie $N(\mu_2, \sigma_2^2)$. (a) Otestujte hypotézu o zhode stredných hodnôt μ_1 a μ_2 na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre rozdiel stredných hodnôt $\mu_1 - \mu_2$, kde koeficient spoľahlivosti $1-\alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku T_W za predpokladu (1.1) rovnosti a (1.2) nerovnosti rozptylov, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} za predpokladu (2.1) rovnosti a (2.2) nerovnosti rozptylov a DIS prislúchajúce (1.1), (1.2), (2.1) a (2.2).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: stredná hodnota pôrodnej hmotnosti chlapcov bez staršieho súrodenca je zhodná so strednou hodnotou pôrodnej hmotnosti chlapcov s jedným starším súrodencom oproti H_1: stredná hodnota pôrodnej hmotnosti chlapcov bez staršieho súrodenca nie je zhodná so strednou hodnotou pôrodnej hmotnosti chlapcov s jedným starším súrodencom.
   - **matematická formulácia** – $H_0: \mu_1 - \mu_2 = \mu_0$ oproti $H_1: \mu_1 - \mu_2 \neq \mu_0$, kde $\mu_0 = 0$.

2. **Testovacia štatistika** – Stredné hodnoty μ_j a rozptyly σ_j^2, kde $j = 1, 2$, sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú aritmetické priemery $\overline{x}_1 \doteq 3127.68$, $\overline{x}_2 \doteq 3194.20$ a rozptyly $s_1^2 \doteq 440261.80$ (smerodajná odchýlka $s_1 \doteq 663.52$), $s_2^2 \doteq 516335.7$ (smerodajná odchýlka $s_2 \doteq 718.565$). Rozdiel $\overline{x}_1 - \overline{x}_2 \doteq -66.53$.

```
DATA <- read.table("two-samples-means-birth.txt",header=TRUE)
names(DATA)
## "o.sib.N" "birth.W"
attach(DATA)
birth.W.0 <- birth.W[o.sib.N==0]
birth.W.1 <- birth.W[o.sib.N==1]
n.0 <- length(birth.W.0) # 297
n.1 <- length(birth.W.1) # 276
n <- n.0+n.1 # 573
priemer.0 <- mean(birth.W.0) # 3127.677
rozptyl.0 <- var(birth.W.0) # 440261.8
smerodch.0 <- sd(birth.W.0) # 663.5223
priemer.1 <- mean(birth.W.1) # 3194.203
rozptyl.1 <- var(birth.W.1) # 516335.7
smerodch.1 <- sd(birth.W.1) # 718.565
priemer.diff <- priemer.0-priemer.1 # -66.52613
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $t_W \doteq -1.152$ (ak $\sigma_1^2 = \sigma_2^2$).
   Waldova testovacia štatistika $t_W \doteq -1.149$ (ak $\sigma_1^2 \neq \sigma_2^2$).
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} \doteq 1.330$ (ak $\sigma_1^2 = \sigma_2^2$).
   Testovacia štatistika pomerom vierohodnosti $u_{\mathrm{LR}} \doteq 1.320$ (ak $\sigma_1^2 \neq \sigma_2^2$).

```
W.test <- t.test(birth.W.0,birth.W.1,var=TRUE)
W.test.welch <- t.test(birth.W.0,birth.W.1,var=FALSE)
tW.obs <- W.test$stat # -1.152217
tW.obs.welch <- W.test.welch$stat # -1.148857
mu.hat <- (sum(birth.W.0)+sum(birth.W.1))/n # 3159.721
sigma.sq.tilde <- (sum((birth.W.0-priemer.0)^2)+
                   sum((birth.W.1-priemer.1)^2))/n # 475235.3
sigma0.sq.tilde <- (sum((birth.W.0-mu.hat)^2)+sum((birth.W.1-mu.hat)^2))/n # 476340.2
uLR.obs <- n*log(sigma0.sq.tilde/sigma.sq.tilde) # 1.330707
sigmad.sq <- (sum((birth.W.0-priemer.0)^2)+sum((birth.W.1-priemer.1)^2))/(n-2)
tWt.obs <- (priemer.diff-0)/sqrt(sigmad.sq)*sqrt(n.0*n.1/(n.0+n.1))
uW.obs <- tWt.obs^2
uLRt.obs <- n*log(1+uW.obs/(n-2)) # 1.330707
gama <- (n.0-1)/n.0*rozptyl.0/((n.1-1)/n.1*rozptyl.1) # 0.852885
uLR.obs.2 <- n.0*n.1*(priemer.1-priemer.0)^2/((n.0+gama*n.1)^2)*((gama^2*n.1)/((n.0-1)/n.0*rozptyl.0)+
             n.0/((n.1-1)/n.1*rozptyl.1)) # 1.324522
```

3. **Zamietacia oblasť** –
   Waldov test (ak $\sigma_1^2 = \sigma_2^2$):

kritické hodnoty $t_{n-2}(1-\alpha/2) = t_{571}(1-0.025) \doteq -1.9641$ a $t_{n-2}(\alpha/2) = t_{571}(0.025) \doteq 1.9641$;
kritický obor $\mathcal{W}_1 = (t_{\min}, t_{n-2}(1-\alpha/2)) \cup (t_{n-2}(\alpha/2), t_{\max}) = (-\infty, -1.9641) \cup (1.9641, \infty)$.
Waldov test (ak $\sigma_1^2 \neq \sigma_2^2$):
kritické hodnoty $t_{df}(1-\alpha/2) = t_{df}(1-0.025) \doteq -1.9642$ a $t_{df}(\alpha/2) = t_{df}(0.025) \doteq 1.9642$;
kritický obor $\mathcal{W}_1 = (t_{\min}, t_{df}(1-\alpha/2)) \cup (t_{df}(\alpha/2), t_{\max}) \doteq (-\infty, -1.9642) \cup (1.9642, \infty)$,
kde $df \doteq 557.986$.
Test pomerom vierohodnosti:
kritická hodnota $\chi_1^2(\alpha) \doteq 3.841$;
kritický obor $\mathcal{W} = (\chi_1^2(\alpha), t_{\max}) \doteq (3.841, \infty)$.

```
W.test$parameter # 571
t.krit.d <- qt(0.025,df=W.test$parameter) # -1.964127
t.krit.h <- qt(0.975,df=W.test$parameter) # 1.964127
W.test.welch$parameter # 557.9862
t.krit.d.welch <- qt(0.025,df=W.test.welch$parameter) # -1.964225
t.krit.h.welch <- qt(0.975,df=W.test.welch$parameter) # -1.964225
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
Waldov $100 \times (1-\alpha)\%$ empirický DIS $\doteq$ 95% empirický DIS pre μ (ak $\sigma_1^2 = \sigma_2^2$):
$(d, h) \doteq (-179.93, 46.88)$.
Waldov 95% empirický DIS pre μ (ak $\sigma_1^2 \neq \sigma_2^2$):
$(d, h) \doteq (-180.27, 47.21)$.
Vierohodnostný 95% empirický DIS pre μ (ak $\sigma_1^2 = \sigma_2^2$):
$\mathcal{CS}_{0.95} = \left\{\mu_0 : u_{\mathrm{LR}}(\mu_0) < \chi_1^2(\alpha)\right\} \doteq (-179.68, 46.62)$.
Vierohodnostný 95% empirický DIS pre μ (ak $\sigma_1^2 \neq \sigma_2^2$):
$\mathcal{CS}_{0.95} = \left\{\mu_0 : u_{\mathrm{LR}}(\mu_0) < \chi_1^2(\alpha)\right\} \doteq (-179.82, 46.76)$.

```
IS.W <- W.test$conf.int # -179.92997 46.87771
IS.W.welch <- W.test.welch$conf.int # -180.26719 47.21493
min.mu <- -3*sqrt(sigmad.sq)
max.mu <- 3*sqrt(sigmad.sq)
mu.0i <- seq(min.mu,max.mu,by = 0.01)
tW.obs.i <- (priemer.diff-mu.0i)/sqrt(sigmad.sq)*sqrt(n.0*n.1/(n.0+n.1))
uW.obs.i <- tW.obs.i^2
uLR.i <- n*log(1+uW.obs.i/(n-2))
IS.LR <- range(mu.0i[which(uLR.i<lr.krit.hodn)]) # -179.67937 46.62063
uLR.i.2 <- n.0*n.1*(priemer.diff-mu.0i)^2/((n.0+gama*n.1)^2)*
                   ((gama^2*n.1)/((n.0-1)/n.0*rozptyl.0)+
                   n.0/((n.1-1)/n.1*rozptyl.1))
IS.LR2 <- range(mu.0i[which(uLR.i.2<lr.krit.hodn)]) # -179.8181 46.7619
```

5. **Štatistický záver** –
Waldov test (ak $\sigma_1^2 = \sigma_2^2$): p-hodnota $\doteq 2\Pr(T_W \geq |-1.152| | H_0) = 0.250$.
Waldov test (ak $\sigma_1^2 \neq \sigma_2^2$): p-hodnota $\doteq 2\Pr(T_W \geq |-1.149| | H_0) = 0.251$.
Test pomerom vierohodnosti (ak $\sigma_1^2 = \sigma_2^2$): p-hodnota $\doteq \Pr(U_{\mathrm{LR}} \geq 1.330 | H_0) = 0.249$.
Test pomerom vierohodnosti (ak $\sigma_1^2 \neq \sigma_2^2$): p-hodnota $\doteq \Pr(U_{\mathrm{LR}} \geq 1.320 | H_0) = 0.251$.

```
p.hodn.W <- W.test$p.val # 0.2497143
p.hodn.W.welch <- W.test.welch$p.val # 0.251107
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # 0.2486795
p.hodn.LR.2 <- 1-pchisq(uLR.obs.2,df=1) # 0.2497821
```

H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\mu_0 = 0$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že stredná hodnota pôrodnej hmotnosti chlapcov bez staršieho súrodenca je zhodná so strednou hodnotou pôrodnej hmotnosti chlapcov s jedným starším súrodencom.

7. **Antropologický slovný záver** – Napriek tomu, že pôrodná hmotnosť prvorodených chlapcov bez staršieho súrodenca je v priemere číselne nižšia než u novorodených chlapcov s jedným starším súrodencom, rozdiel nie je štatisticky významný. V našej vzorke sa teda nepodarilo dokázať obvykle nachádzaný stav, ktorý vyplýva z rozdielov medzi matkami v kapacite reprodukčnej sústavy. Malý rozdiel môže byť daný kvalitou ako prenatálnej starostlivosti, tak výživy matiek a ďalšími civilizačnými faktormi, ktoré môžu obe skupiny zbližovať. Ďalšou príčinou môže byť skutočnosť, že ide o porovnanie dvoch nezávislých vzoriek, ktoré sa môžu líšiť genetickým pozadím (sociálne

a etnickým zložením) i podmienkami prostredia. Nie sú tiež k dispozícii údaje o proporcii prirodzených pôrodov a pôrodov cisárskym rezom. Vhodným spôsobom realizácie výskumu by bolo párové porovnanie staršieho a mladšieho syna z jedného páru súrodencov tých istých rodičov, kde by bol vplyv nekontrolovaných faktorov obmedzenejší.

Príklad 231 (test o rozdiele stredných hodnôt μ_1 a μ_2) *Majme dáta* two-samples-means-skull.txt, *premennú výška lebky* skull.H *v mm starovekej egyptskej populácii a premennú pohlavie* sex, *ktorá nadobúda hodnoty* m *(muž) a* f *(žena). Predpokladáme, že premenná* skull.H *mužov má normálne rozdelenie $N(\mu_1, \sigma_1^2)$ a* skull.H *žien má normálne rozdelenie $N(\mu_2, \sigma_2^2)$. (a) Otestujte hypotézu o zhode stredných hodnôt μ_1 a μ_2 na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)$% empirický DIS pre rozdiel stredných hodnôt $\mu_1 - \mu_2$, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku T_W za predpokladu (1.1) rovnosti a (1.2) nerovnosti rozptylov, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} za predpokladu (2.1) rovnosti a (2.2) nerovnosti rozptylov a DIS prislúchajúce (1.1), (1.2), (2.1) a (2.2).*

7.2 Asymptotické testy o podiele rozptylov

Ako bolo uvedené v úvodnom texte kapitoly 6.2, rozptyl hodnôt a jeho porovnanie medzi rôznymi výbermi má veľký význam pri odhaľovaní procesov, ktoré sa v kontexte daného znaku v populácii odohrávajú, resp. odohrávali v priebehu evolúcie. Testovanie rozdielov v rozptyle je preto dôležitým prostriedkom pri zisťovaní biologických rozdielov medzi dvoma populáciami. Na rozdiel od testovania rozptylu jednej vzorky s porovnávacou (tabelovou, publikovanou) hodnotou majú testy rozptylov dvoch výberov spracovaných v jednej štúdii tú výhodu, že obe vzorky boli získané rovnakou metodikou a rozdiely medzi nimi vypovedajú o porovnateľných zdrojoch variability (a snáď obdobnej miere vplyvu skutočných biologických rozdielov).

Nech $X_j \sim N(\mu_j, \sigma_j^2), j = 1, 2$, sú dve nezávislé náhodné premenné. Majme dvojice hypotéz $H_{01} : \sigma_1^2/\sigma_2^2 = \sigma_0$ oproti $H_{11} : \sigma_1^2/\sigma_2^2 \neq \sigma_0$, $\mathrm{H}_{02} : \sigma_1^2/\sigma_2^2 \leq \sigma_0$ oproti $H_{21} : \sigma_1^2/\sigma_2^2 > \sigma_0$ a $\mathrm{H}_{03} : \sigma_1^2/\sigma_2^2 \geq \sigma_0$ oproti $H_{13} : \sigma_1^2/\sigma_2^2 < \sigma_0$.

Ak H_0 platí a $\sigma_0 = 1$, potom

$$F_W = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \overset{\mathcal{D}}{\sim} F_{n_1-1,n_2-1},$$

kde $S_j^2 = \frac{1}{n_j-1}\sum_{i=1}^{n_j}(X_{ji} - \overline{X}_j)^2, j = 1, 2$, a rozdelenie F_{df_1,df_2} sa nazýva **centrálne F-rozdelenie s $df_1 = n_1 - 1$ a $df_2 = n_2 - 1$ stupňami voľnosti**, n_1 a n_2 sú rozsahy náhodných výberov. F_W sa nazýva **Waldova testovacia štatistika** (alebo **F-štatistika**) a test **dvojvýberový F-test o podiele rozptylov σ_1^2/σ_2^2**.

Chceme nájsť kritickú hodnotu d_S^* za platnosti H_{02}, tak aby platilo $\Pr\left(S_1^2/S_2^2 < d_S^* | H_{02}\right) = \alpha$. Teda

$$\alpha = \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < d_S^* \frac{\sigma_2^2}{\sigma_1^2} | H_{02}\right) = \Pr\left(F_W < d_S^* | H_{02}\right)$$

a z toho vyplýva, že $d_S^* = F_{n_1-1,n_2-1}(1-\alpha)$. Ak H_{02} neplatí, t.j. skutočný rozptyl je σ_1^2/σ_2^2, potom chceme nájsť pravdepodobnosť zamietnutia H_{02} (silu testu). To je pravdepodobnosť, že S_1^2/S_2^2 neprekročí d_S^*, teda

$$\begin{aligned}\Pr\left(S_1^2/S_2^2 < d_S^* | H_{12}\right) &= \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < d_S^* \frac{\sigma_2^2}{\sigma_1^2} | H_{12}\right) = \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < d_S^* \frac{\sigma_2^2}{\sigma_1^2} | H_{12}\right) \\ &= \Pr\left(F_W^{(\mathrm{alt})} < F_{n_1-1,n_2-1}(1-\alpha) \frac{\sigma_2^2}{\sigma_1^2} | H_{12}\right),\end{aligned}$$

kde $F_W^{(\mathrm{alt})} = \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2}$.

Chceme nájsť kritickú hodnotu h_S^* za platnosti H_{03}, tak aby platilo $\Pr\left(S_1^2/S_2^2 < h_S^*|H_{03}\right) = 1-\alpha$. Teda

$$1-\alpha \quad = \quad \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < h_S^*\frac{\sigma_2^2}{\sigma_1^2}|H_{03}\right) = \Pr\left(F_W < h_S^*|H_{03}\right)$$

a z toho vyplýva, že $h_S^* = F_{n_1-1,n_2-1}(\alpha)$. Ak H_{03} neplatí, t.j. skutočný rozptyl je σ_1^2/σ_2^2, potom chceme nájsť pravdepodobnosť zamietnutia H_{03} (silu testu). To je pravdepodobnosť, že S_1^2/S_2^2 prekročí h_S^*, teda

$$\begin{aligned}\Pr\left(S_1^2/S_2^2 > h_S^*|H_{13}\right) &= \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} > h_S^*\frac{\sigma_2^2}{\sigma_1^2}|H_{13}\right) = \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} > h_S^*\frac{\sigma_2^2}{\sigma_1^2}|H_{13}\right)\\ &= \Pr\left(F_W^{(\mathrm{alt})} > F_{n_1-1,n_2-1}(\alpha)\frac{\sigma_2^2}{\sigma_1^2}|H_{13}\right).\end{aligned}$$

Ak by sme chceli nájsť kritické hodnoty d_S a h_S pre test H_{01} oproti H_{11}, použili by sme podobnú úvahu ako vyššie.

Príklad 232 *Odvoďte vzorec na výpočet kritických hodnôt d_S a h_S pre H_{01}.*

Definícia 73 (Kritický obor a silofunkcia F_W testu o σ_1^2/σ_2^2) *Kritický obor a silofunkcia sú definované nasledovne (ozn. $d_{\alpha/2} = F_{n_1-1,n_2-1}(1-\alpha/2)$, $h_{\alpha/2} = F_{n_1-1,n_2-1}(\alpha/2)$, $d_\alpha = F_{n_1-1,n_2-1}(1-\alpha)$ a $h_\alpha = F_{n_1-1,n_2-1}(\alpha)$):*

H_0	H_1	$\mathcal{W}$	$1-\beta\left(\sigma_1^2/\sigma_2^2\right)$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$\mathcal{W}_1 = \left\{F_W; F_W \notin \left(d_{\alpha/2}, h_{\alpha/2}\right)\right\}$	$1-\Pr\left(\frac{\sigma_2^2}{\sigma_1^2}d_{\alpha/2} < F_W^{(alt)} < \frac{\sigma_2^2}{\sigma_1^2}h_{\alpha/2}\right)$
$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$\mathcal{W}_2 = \{F_W; F_W \geq h_\alpha\}$	$\Pr\left(F_W^{(alt)} \geq \frac{\sigma_2^2}{\sigma_1^2}h\right)$
$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	$\mathcal{W}_3 = \{F_W; F_W \leq d_\alpha\}$	$\Pr\left(F_W^{(alt)} \leq \frac{\sigma_2^2}{\sigma_1^2}d_\alpha\right)$

Definícia 74 (p-hodnota F_W testu o σ_1^2/σ_2^2) *Nech F_W je **nejaká testovacia štatistika** a $F_{obs} = \frac{s_1^2}{s_2^2}$ je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2\min(\Pr(F_W \leq F_{obs}|H_{01}), \Pr(F_W \geq F_{obs}|H_0)),\ ak & H_{11}: \sigma_1^2 \neq \sigma_2^2\\ Pr(F_W \geq F_{obs}|H_{02}),\ ak & H_{12}: \sigma_1^2 > \sigma_2^2\\ Pr(F_W \leq F_{obs}|H_{03}),\ ak & H_{13}: \sigma_1^2 < \sigma_2^2\end{cases}$$

Definícia 75 (Waldove $100\times(1-\alpha)\%$ empirické IS pre σ_1^2/σ_2^2) *Waldove $100\times(1-\alpha)\%$ empirické IS pre σ_1^2/σ_2^2 pre všetky tri typy hypotéz majú nasledovný tvar (vieme, že platí $\frac{1}{F_{n_2-1,n_1-1}(1-\alpha/2)} = F_{n_1-1,n_2-1}(\alpha/2)$):*

H_0	H_1	*hranice (d,h) pre $100(1-\alpha)\%$ empirický IS*
$\sigma_1^2/\sigma_2^2 = \sigma_0^2$	$\sigma_1^2/\sigma_2^2 \neq \sigma_0^2$	$\mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : \sigma_0^2 \in \left(\frac{s_1^2}{s_2^2 h_{\alpha/2}}, \frac{s_1^2}{s_2^2 d_{\alpha/2}}\right)\right\}$
$\sigma_1^2/\sigma_2^2 \leq \sigma_0^2$	$\sigma_1^2/\sigma_2^2 > \sigma_0^2$	$\mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : \sigma_0^2 \in \left(\frac{s_1^2}{s_2^2 h_\alpha}, \infty\right)\right\}$
$\sigma_1^2/\sigma_2^2 \geq \sigma_0^2$	$\sigma_1^2/\sigma_2^2 < \sigma_0^2$	$\mathcal{CS}_{1-\alpha} = \left\{\sigma_0^2 : \sigma_0^2 \in \left(0, \frac{s_1^2}{s_2^2 d_\alpha}\right)\right\}$

Minimálne rozsahy n_1, n_2 pre nejaký podiel σ_1^2/σ_2^2 (pri nejakom α a β) vypočítame iteračne pomocou nasledovných rovností:

$$\frac{F_{n_1-1,n_2-1}(1-\beta)}{F_{n_1-1,n_2-1}(\alpha/2)} \approx \frac{\sigma_2^2}{\sigma_1^2} \text{ za platnosti } H_{11};$$
$$\frac{F_{n_1-1,n_2-1}(\beta)}{F_{n_1-1,n_2-1}(1-\alpha)} = \frac{\sigma_2^2}{\sigma_1^2} \text{ za platnosti } H_{12};$$
$$\frac{F_{n_1-1,n_2-1}(1-\beta)}{F_{n_1-1,n_2-1}(\alpha)} = \frac{\sigma_2^2}{\sigma_1^2} \text{ za platnosti } H_{13}.$$

Pri obojstrannej alternatíve ide o približný výpočet, kedy sa zanedbáva časť sily, ktorá predstavuje obsah pod dolnou kritickou hodnotou.

Dvojvýberový F-test o podiele rozptylov σ_1^2/σ_2^2 v R (funkcia **var.test()**).
Argumenty (vstupy) funkcie:

1. vektory dát x a y;
2. alternatíva alternative="two.sided" je prednastavená, ďalšie voľby sú "greater", "less";
3. podiel rozptylov za platnosti nulovej hypotézy ratio, t.j. σ_0^2;
4. spoľahlivosť conf.level=p, prednastavené je p =.95.

Výstupy funkcie:

1. názov použitého testu method;
2. testovacia štatistika statistic;
3. počet stupňov voľnosti df parameter (num df a denom df);
4. p-hodnota p.value;
5. alternatívna hypotéza alternative hypothesis;
6. interval spoľahlivosti podielu σ_1^2/σ_2^2 conf.int;
7. bodový odhad (podiel $\widehat{\sigma}_1^2/\widehat{\sigma}_2^2$) estimate.

Test pomerom vierohodnosti pre σ_1^2/σ_2^2.

Nech $X_1 \sim N(\mu_1, \sigma_1^2)$ a $X_2 \sim N(\mu_2, \sigma_2^2)$, kde $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)^T$. Logaritmus funkcie vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = -\sum_{j=1}^{2} \frac{n_j}{2} \ln(2\pi\sigma_j^2) - \sum_{j=1}^{2} \frac{1}{2\sigma_j^2} \left(\sum_{i=1}^{n_j} (x_{ji} - \mu_j)^2 \right).$$

MLE $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2)^T$, kde

$$\widehat{\mu}_j = \overline{x}_j, \widehat{\sigma}_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} (x_{ji} - \overline{x}_j)^2,$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \sigma_1^2 \neq \sigma_2^2 \neq \sigma^2\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\widehat{\mu}_1, \widehat{\mu}_2, \widetilde{\sigma}^2, \widetilde{\sigma}^2)^T$, kde (pozri Mood a kol., 1987)

$$\widetilde{\sigma}^2 = \frac{\sum_{j=1}^{2} n_j \widehat{\sigma}_j^2}{\sum_{j=1}^{2} n_j},$$

t.j. $\Theta_0 = \{\boldsymbol{\theta} : \sigma_1^2 = \sigma_2^2 = \sigma^2\}$. Logaritmus funkcie vierohodnosti pre $\widehat{\boldsymbol{\theta}}$ bude mať tvar

$$\begin{aligned} l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{x}_2) &= -\sum_{j=1}^{2} \frac{n_j}{2} \ln(2\pi\widehat{\sigma}_j^2) - \sum_{j=1}^{2} \frac{1}{2\widehat{\sigma}_j^2} \left(\sum_{i=1}^{n_j} (x_{ji} - \mu_j)^2 \right) \\ &= -\sum_{j=1}^{2} \frac{n_j}{2} \ln(2\pi\widehat{\sigma}_j^2) - \sum_{j=1}^{2} \frac{n_j\widehat{\sigma}_j^2}{2\widehat{\sigma}_j^2}. \end{aligned}$$

Logaritmus funkcie vierohodnosti pre $\boldsymbol{\theta}_0$ bude mať tvar

$$\begin{aligned} l(\boldsymbol{\theta}_0|\mathbf{x}_1,\mathbf{x}_2) &= -\sum_{j=1}^{2}\frac{n_j}{2}\ln(2\pi\widetilde{\sigma}^2)-\sum_{j=1}^{2}\frac{1}{2\widetilde{\sigma}^2}\left(\sum_{i=1}^{n_j}(x_{ji}-\mu_j)^2\right) \\ &= -\sum_{j=1}^{2}\frac{n_j}{2}\ln(2\pi\widetilde{\sigma}^2)-\frac{1}{\widetilde{\sigma}^2}\sum_{j=1}^{2}\frac{n_j\widehat{\sigma}_j^2}{2} \end{aligned}$$

Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{x}_1,\mathbf{x}_2)) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1,\mathbf{x}_2)-l\left(\boldsymbol{\theta}_0|\mathbf{x}_1,\mathbf{x}_2\right) \\ &= -\sum_{j=1}^{2}\frac{n_j}{2}\ln\widehat{\sigma}_j^2-\sum_{j=1}^{2}\frac{n_j}{2}+\sum_{j=1}^{2}\frac{n_j}{2}\ln\left(\frac{\sum_{i=1}^{2}n_i\widehat{\sigma}_i^2}{\sum_{i=1}^{2}n_i}\right)+\sum_{j=1}^{2}\frac{n_j}{2} \\ &= \frac{1}{2}\sum_{j=1}^{2}n_j\ln\left(\frac{\sum_{i=1}^{2}n_i\widehat{\sigma}_i^2}{\widehat{\sigma}_j^2\sum_{i=1}^{2}n_i}\right). \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}}=-2\ln\left(\lambda\left(\mathbf{X}_1,\mathbf{X}_2\right)\right)\overset{\mathcal{D}}{\sim}\chi_1^2$.

Vierohodnostný $100\times(1-\alpha)\%$ **empirický DIS pre** σ_1^2/σ_2^2 bude mať tvar

$$\mathcal{CS}_{1-\alpha}=\left\{\sigma_0^2: U_{\mathrm{LR}}<\chi_1^2\left(\alpha\right)\right\}.$$

Na výpočet tohoto DIS by sme museli reparametrizovať funkciu vierohodnosti na parameter záujmu (σ_1^2/σ_2^2) a rušivý parameter. Alternatívne by bolo možné vypočítať u_{LR} pre rôzne kombinácie σ_1^2 a σ_2^2 (na vhodne zvolených intervaloch) a pre tieto σ_j^2, $j=1,2$, vypočítať podiely σ_1^2/σ_2^2. Potom $\mathcal{CS}_{1-\alpha}=\left\{\sigma_0^2: U_{\mathrm{LR}}(\sigma_1^2/\sigma_2^2)<\chi_1^2\left(\alpha\right)\right\}$.

Príklad 233 (test o podiele rozptylov σ_1^2 a σ_2^2) *Majme dáta two-samples-variances-skull.txt, premennú výška lebky skull.H v milimetroch u mužov a žien v starovekej egypskej populácii. Predpokladáme, že premenná skull.H mužov má normálne rozdelenie $N(\mu_1,\sigma_1^2)$ a skull.H žien má normálne rozdelenie $N(\mu_2,\sigma_2^2)$. (a) Otestujte hypotézu o podiele rozptylov σ_1^2 a σ_2^2 na hladine významnosti $\alpha=0.05$. (b) Vypočítajte $100\times(1-\alpha)\%$ empirický DIS pre podiel rozptylov σ_1^2/σ_2^2, kde koeficient spoľahlivosti $1-\alpha=0.95$. Použite (1) Waldovu testovaciu štatistiku F_W, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} a DIS prislúchajúce (1).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: podiel rozptylov výšky lebky u mužov a u žien je rovný jednej (t.j. rozptyly výšky lebky u mužov a u žien sa rovnajú) oproti H_1: podiel rozptylov výšky lebky u mužov a u žien nie je rovný jednej (t.j. rozptyly výšky lebky u mužov a u žien sa nerovnajú).
   - **matematická formulácia** – $H_0:\sigma_1^2/\sigma_2^2=\sigma_0^2$ oproti $H_1:\sigma_1^2/\sigma_2^2\neq\sigma_0^2$, kde $\sigma_0^2=1$.
2. **Testovacia štatistika** – Rozptyly σ_j^2, kde $j=1,2$, sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú $s_1^2\doteq 23.38$ a $s_2^2\doteq 21.65$. Podiel $s_1^2/s_2^2=1.08$.

```
DATA <- read.table("two-samples-variances-skull.txt",header=TRUE)
names(DATA)
##    "id"         "pop"         "sex"         "skull.H"
attach(DATA)
skull.H.m <- na.omit(skull.H[sex=="m"])
skull.H.f <- na.omit(skull.H[sex=="f"])
n.m <- length(skull.H.m) # 215
n.f <- length(skull.H.f) # 107
n <- n.m+n.f # 322
```

```
rozptyl.m <- var(skull.H.m) # 23.382
rozptyl.f <- var(skull.H.f) # 21.65279
rozptyl.ratio <- rozptyl.m/rozptyl.f # 1.079861
rozptyl.tilde <- (n.m*rozptyl.m+n.f*rozptyl.f)/n # 22.80739
```

Pozorované testovacie štatistiky:
Waldova testovacia štatistika $F_{\text{obs}} \doteq 1.08$.
Testovacia štatistika pomerom vierohodnosti $u_{\text{LR}} \doteq 0.21$.

```
F.test <- var.test(skull.H.m, skull.H.f)
F.obs <- F.test$stat # 1.079861
uLR.obs <- n.m*log((n.m*rozptyl.m+n.f*rozptyl.f)/(rozptyl.m*n))+
            n.f*log((n.m*rozptyl.m+n.f*rozptyl.f)/(rozptyl.f*n)) # 0.2090286
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $F_{n_{\text{m}-1,\text{f}-1}}(1-\alpha/2) = F_{214,106}(1-\alpha/2) \doteq 1.40$ a $F_{n_{\text{m}-1,\text{f}-1}}(\alpha/2) = F_{214,106}(\alpha/2) \doteq 0.73$;
   kritický obor $\mathcal{W}_1 \doteq (0, 0.73) \cup (1.40, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi^2_1(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi^2_1(\alpha), \infty) = (3.841, \infty)$.

```
df <- F.test$param # 214 106
F.krit.d <- qf(0.025, df1=df[1], df2=df[2]) # 0.7251934
F.krit.h <- qf(0.975, df1=df[1], df2=df[2]) # 1.404928
lr.krit.hodn <- qchisq(0.95, df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100 \times (1-\alpha)\%$ empirický DIS = 95% empirický DIS pre σ_1^2/σ_2^2:
   $(d, h) \doteq (0.769, 1.489)$.

```
IS.W <- F.test$conf.int # 0.7686237 1.4890660
```

5. **Štatistický záver** –
   Waldov test: p-hodnota $\doteq 2\Pr(F_W \geq 1.08|H_0) = 0.663$.
   Test pomerom vierohodnosti: p-hodnota $\doteq \Pr(U_{\text{LR}} \geq 0.209|H_0) = 0.648$.

```
p.hodn.F <- F.test$p.val # 0.6630295
p.hodn.LR <- 1-pchisq(uLR.obs, df=1) # 0.6475298
```

   H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\sigma_0^2 = 0$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že podiel rozptylov výšky lebky u mužov a u žien je rovný jednej.

7. **Antropologický slovný záver** – Nemôžeme zamietnuť zhodu rozptylu výšky lebky mužov a žien v stredovekej egyptskej populácii. Faktory ovplyvňujúce rozptyl hodnôt tohto znaku boli u oboch pohlaví podobné.

7.3 Asymptotické testy o rozdiele korelačných koeficientov

Obvyklé typy testov v antropologickom výskume riešia veľkosť a znamienko závislosti (súvislosti) medzi dvoma premennými, ale zriedka sa stretneme so snahou porovnávať dve súvislosti tých istých premenných testovaním rozdielov medzi dvoma korelačnými koeficientmi. Pritom práve takýto test nás môže posunúť v pohľade na študovaný problém oveľa ďalej než testy rozdielov stredných hodnôt. Ak súvislosť medzi dvoma premennými odráža nejaký biologický proces, a pokiaľ je v jednej skupine súvislosť systematicky odlišná (korelačné koeficienty sa štatisticky významne líšia), vypovedá to o zásadných rozdieloch v biologických procesoch, ktoré za danú súvislosť zodpovedajú. Typickou oblasťou, v ktorej možno uplatniť testy rozdielov korelačných koeficientov, je sexuálny dimorfizmus. Napríklad hladiny steroidných hormónov môžu ovplyvňovať určitú časť tela u každého pohlavia odlišne alebo v rôznej miere.

Nech $(X_j, Y_j)^T \sim N_2(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, kde $j = 1, 2$. Nech ρ_j je korelačný koeficient X_j a Y_j, R_j je výberový korelačný koeficient a $\widehat{\rho}_j = r_j$ je odhad ρ_j. Korelačný koeficient má asymptoticky normálne rozdelenie, t.j. $R_j \sim N\left(\rho_j, \left(1-\rho_j^2\right)^2 / (n_j - 1)\right)$, kde n_j sú rozsahy náhodných výberov. Zaujíma nás rozdiel $R_1 - R_2$, avšak jeho konvergencia k normalite je pomalá, preto sa namieto R_j používajú **Fisherove** Z**-premenné** Z_1 a Z_2. Potom

$$Z_1 - Z_2 \sim N\left(0, \frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}\right), Z_j = \frac{1}{2}\ln\frac{1+R_j}{1-R_j}, j = 1, 2.$$

Majme dvojice hypotéz $H_{01}: \rho_1 - \rho_2 = \rho_0$ oproti $H_{11}: \rho_1 - \rho_2 \neq \rho_0$, $H_{02}: \rho_1 - \rho_2 \leq \rho_0$ oproti $H_{12}: \rho_1 - \rho_2 > \rho_0$, $H_{03}: \rho_1 - \rho_2 \geq \rho_0$ oproti $H_{13}: \rho_1 - \rho_2 < \rho_0$.

Ak H_0 platí, potom

$$Z_W = \frac{Z_1 - Z_2 - \xi_0}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \overset{\mathcal{D}}{\sim} N(0,1).$$

kde $\xi_0 = \frac{1}{2}\ln\frac{1+\rho_0}{1-\rho_0}$. Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **dvojvýberový** Z**-test o rozdiele korelačných koeficientov** ρ_1 **a** ρ_2.

Definícia 76 (Kritický obor a silofunkcia Z_W testu o $\rho_1 - \rho_2$) *Kritický obor a silofunkcia sú definované nasledovne (ozn. $\xi = \frac{1}{2}\ln\frac{1+\rho}{1-\rho}$, $\delta_\xi = \xi_0 - (\xi_1 - \xi_2)$ a $n_1 = n_2 = n$):*

H_0	H_1	$\mathcal{W}$	$1-\beta(\xi_1, \xi_2)$
$\rho_1 - \rho_2 = \rho_0$	$\rho_1 - \rho_2 \neq \rho_0$	$\mathcal{W}_1 = \{Z_W; \lvert Z_W\rvert \geq u(\alpha/2)\}$	$1 - \Phi\left(u_{\alpha/2} - \sqrt{(n-3)/2}\,\lvert\delta_\xi\rvert\right)$
$\rho_1 - \rho_2 \leq \rho_0$	$\rho_1 - \rho_2 > \rho_0$	$\mathcal{W}_2 = \{Z_W; Z_W \geq u(\alpha)\}$	$1 - \Phi\left(u_\alpha + \sqrt{(n-3)/2}\,(\delta_\xi)\right)$
$\rho_1 - \rho_2 \geq \rho_0$	$\rho_1 - \rho_2 < \rho_0$	$\mathcal{W}_3 = \{Z_W; Z_W \leq -u(\alpha)\}$	$1 - \Phi\left(u_\alpha - \sqrt{(n-3)/2}\,(\delta_\xi)\right)$

Definícia 77 (p-hodnota Z_W testu o $\rho_1 - \rho_2$) *Nech Z_W je* ***nejaká testovacia štatistika*** *a z_W je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W \geq |z_W| \,|\, H_{01}),\ ak & H_{11}: \rho_1 - \rho_2 \neq \rho_0 \\ Pr(Z_W \geq z_W | H_{02}),\ ak & H_{12}: \rho_1 - \rho_2 > \rho_0 \\ Pr(Z_W \leq z_W | H_{03}),\ ak & H_{13}: \rho_1 - \rho_2 < \rho_0 \end{cases}$$

Definícia 78 (Waldov $100 \times (1-\alpha)\%$ empirický DIS pre ξ a IS pre $\rho_1 - \rho_2$) *Keďže* $\Pr\left(\frac{1}{\sigma_g}|Z_1 - Z_2 - \xi| \leq u(\alpha/2)\right) = 1 - \alpha$, *kde* $\sigma_g^2 = \frac{1}{n_1-3} + \frac{1}{n_2-3}$, *potom Waldov* $100 \times (1-\alpha)\%$ *empirický DIS pre ξ má tvar*

$$\mathcal{CS}_{1-\alpha} = \left\{\xi_0; \xi_0 \in \left(z_1 - z_2 - u_{\alpha/2}\widehat{\sigma}_g, z_1 - z_2 + u_{\alpha/2}\widehat{\sigma}_g\right)\right\},$$

kde $\widehat{\sigma}_g^2 = \frac{1}{n_1-3} + \frac{1}{n_2-3}$. *Vieme, že* $\rho = \tanh(\xi)$ *je monotónna funkcia ξ. Potom môžeme písať spätne transformované* ***Waldove*** $100 \times (1-\alpha)\%$ ***empirické IS pre*** $\rho_1 - \rho_2$ *ako:*

H_0	H_1	*hranice (d, h) pre $100(1-\alpha)\%$ empirický IS*
$\rho_1 - \rho_2 = \rho_0$	$\rho_1 - \rho_2 \neq \rho_0$	$\mathcal{CS}_{1-\alpha} = \{\rho_0 : \rho_0 \in (\tanh[d], \tanh[h])\}$
$\rho_1 - \rho_2 \leq \rho_0$	$\rho_1 - \rho_2 > \rho_0$	$\mathcal{CS}_{1-\alpha} = \{\rho_0 : \rho_0 \in (\tanh[z_1 - z_2 - u_\alpha\widehat{\sigma}_g], 2)\}$
$\rho_1 - \rho_2 \geq \rho_0$	$\rho_1 - \rho_2 < \rho_0$	$\mathcal{CS}_{1-\alpha} = \{\rho_0 : \rho_0 \in (-2, \tanh[z_1 - z_2 + u_\alpha\widehat{\sigma}_g])\}$

kde $d = z_1 - z_2 - u_{\alpha/2}\widehat{\sigma}_g$ *a* $h = z_1 - z_2 + u_{\alpha/2}\widehat{\sigma}_g$.

Príklad 234 (DIS pre $\rho_1 - \rho_2$ použitím Fisherovej Z-premennej) *Naprogramujte DIS pre $\rho_1 - \rho_2$.*

Riešenie v R

```
"ISkor.rozdiel" <- function(x1,y1,x2,y2,conf.level=0.95){
z1 <- atanh(cor(x1,y1)); z2 <- atanh(cor(x2,y2))
n1 <- length(x1); n2 <- length(x2)
a <- qnorm(1-(1-conf.level)/2)*(1/(n1-3)+1/(n2-3))^0.5
IS.xi <- c(z1-z2-a,z1-z2+a)
IS <- tanh(IS.xi)
return(IS)
}
```

Majme H_{11}. Použitím Fisherovej Z-premennej pre nejaký rozdiel $\xi_1 - \xi_2 - \xi_0$, α a β dostaneme **minimálny rozsah** $n = n_1 = n_2$ ako

$$n \geq 3 + 2\left(\frac{u_{\alpha/2} + u_\beta}{\xi_1 - \xi_2 - \xi_0}\right)^2.$$

Pre H_{12} a H_{13} zameníme $u_{\alpha/2}$ za u_α.

Príklad 235 *Sledujeme hladinu LDL cholesterolu a krvného tlaku u slovenských žien a mužov. Lekári stanovili klinicky dôležitý rozdiel korelačných koeficientov ako $\rho_1 - \rho_2 = 0.2$, kde $\rho_1 = 0.1, 0.3, \ldots, 0.9$. Ďalej nech sila testu je $1-\beta = 0.90$ (pravdepodobnosť odhalenia tohoto rozdielu) a hladina významnosti $\alpha = 0.05$. Aspoň koľko žien a mužov ($n_1 = n_2 = n$) musíme sledovať pri danej α a $1-\beta$?*

Riešenie v R (pozri tabuľku 7.1)

```
"min.rozsah.n.kor" <- function(r1,r2,alfa,sila){
xi1 <- atanh(r1); xi2 <- atanh(r2)
xi.diff <- xi1-xi2
n <- ceiling(3+2*((qnorm(1-alfa/2)+qnorm(sila))/(xi1-xi2))^2)
TAB <- rbind(round(r1,1),round(r2,1),round(xi.diff,4),n)
dimnames(TAB)[[1]] <- c("r1","r2","xi.diff","n")
return(TAB)
}
r1 <- seq(0.1,0.9,by=0.1)
r2 <- r1-0.2
min.rozsah.n.kor(r1,r2,0.05,0.9)
```

Tabuľka 7.1: Minimálne rozsahy n pre rozdiel $\rho_1 - \rho_2 = 0.2$ pri rôznych ρ a ρ_0 spolu s rozdielom $\xi_1 - \xi_2$, ktorý je funkciou ρ_1 a ρ_2

ρ_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
ρ_2	−0.1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$\xi_1 - \xi_2$	0.2007	0.2027	0.2092	0.2209	0.2398	0.2695	0.3180	0.4055	0.6049
n	525	515	484	434	369	293	211	131	61

Test pomerom vierohodnosti pre $\rho_1 - \rho_2$.

Majme H_{01} oproti H_{11}. Nech $(X_j, Y_j)^T \sim N_2(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, kde $(X_j, Y_j)^T$ sú nezávislé, vektor $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)^T$, $\boldsymbol{\theta}_j = (\mu_{j1}, \mu_{j2}, \sigma^2_{j1}, \sigma^2_{j2}, \rho_j)^T$ a $j = 1, 2$. Logaritmus funkcie vierohodnosti bude mať tvar

$$\begin{aligned} l(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2) &= -n\ln(2\pi) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(\sigma^2_{j1}\sigma^2_{j2}\right) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1-\rho_j^2\right) \\ &- \sum_{j=1}^{2}\frac{1}{2(1-\rho_j^2)}\left(\frac{\sum_{i=1}^{n_j}(x_{ji}-\mu_{j1})^2}{\sigma^2_{j1}} - 2\rho_j\frac{\sum_{i=1}^{n_j}(x_{ji}-\mu_{j1})(y_{ji}-\mu_{j2})}{\sigma_{j1}\sigma_{j2}}\right. \\ &+ \left.\frac{\sum_{i=1}^{n_j}(y_{ji}-\mu_{j2})^2}{\sigma^2_{j2}}\right), \end{aligned}$$

kde $n = \sum_{j=1}^{2} n_j$. MLE $\widehat{\boldsymbol{\theta}}_j = (\widehat{\mu}_{j1}, \widehat{\mu}_{j2}, \widehat{\sigma}_{j1}^2, \widehat{\sigma}_{j2}^2, \widehat{\rho}_j)^T$, kde

$$\widehat{\mu}_{j1} = \overline{x}_j, \widehat{\mu}_{j2} = \overline{y}_j, \widehat{\sigma}_{j1}^2 = \frac{1}{n_j}\sum_{i=1}^{n_j}(x_{ji} - \overline{x}_j)^2, \widehat{\sigma}_{j2}^2 = \frac{1}{n_j}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j)^2,$$

$$\widehat{\rho}_j = \frac{\sum_{i=1}^{n_j}(x_{ji} - \overline{x}_j)(y_{ji} - \overline{y}_j)}{n_j \widehat{\sigma}_{j1}\widehat{\sigma}_{j2}}$$

a $\Theta_1 = \{\boldsymbol{\theta} : \rho_1 \neq \rho_2 \neq \rho\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_{j0} = (\widehat{\mu}_{j1}, \widehat{\mu}_{j2}, \widetilde{\sigma}_{j1}^2, \widetilde{\sigma}_{j2}^2, \widehat{\rho})^T$, kde

$$\widetilde{\sigma}_{j1}^2 = \frac{\widehat{\sigma}_{j1}^2(1 - \widehat{\rho}\widehat{\rho}_j)}{1 - \widehat{\rho}^2}, \widetilde{\sigma}_{j2}^2 = \frac{\widehat{\sigma}_{j2}^2(1 - \widehat{\rho}\widehat{\rho}_j)}{1 - \widehat{\rho}^2}$$

$\widehat{\rho}$ je iteračným riešením rovnice $\sum_{j=1}^{2} \frac{n_j(\widehat{\rho}_j - \widehat{\rho})}{1 - \widehat{\rho}_j\widehat{\rho}} = 0$ (pozri Pearson, 1933) a $\Theta_0 = \{\boldsymbol{\theta} : \rho_1 = \rho_2 = \rho\}$. Logaritmus funkcie vierohodnosti pre $\widehat{\boldsymbol{\theta}}$ bude mať tvar

$$\begin{aligned} l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2) &= -n\ln(2\pi) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(\widehat{\sigma}_{j1}^2\widehat{\sigma}_{j2}^2\right) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1 - \widehat{\rho}_j^2\right) \\ &- \sum_{j=1}^{2}\frac{1}{2(1 - \widehat{\rho}_j^2)}\left(\frac{n_j\widehat{\sigma}_{j1}^2}{\widehat{\sigma}_{j1}^2} - 2\widehat{\rho}_j\frac{n_j\widehat{\sigma}_{j1}^2\widehat{\sigma}_{j2}^2\widehat{\rho}_j}{\widehat{\sigma}_{j1}^2\widehat{\sigma}_{j2}^2} + \frac{n_j\widehat{\sigma}_{j2}^2}{\widehat{\sigma}_{j2}^2}\right) \\ &= -n\ln(2\pi) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(\widehat{\sigma}_{j1}^2\widehat{\sigma}_{j2}^2\right) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1 - \widehat{\rho}_j^2\right) - n. \end{aligned}$$

Logaritmus funkcie vierohodnosti pre $\boldsymbol{\theta}_0$ bude mať tvar

$$\begin{aligned} l(\boldsymbol{\theta}_0|\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2) &= -n\ln(2\pi) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(\widetilde{\sigma}_{j1}^2\widetilde{\sigma}_{j2}^2\right) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1 - \widehat{\rho}^2\right) \\ &- \sum_{j=1}^{2}\frac{1}{2(1 - \widehat{\rho}^2)}\left(\frac{n_j\widehat{\sigma}_{j1}^2}{\widetilde{\sigma}_{j1}^2} - 2\widehat{\rho}\frac{n_j\widehat{\sigma}_{j1}^2\widehat{\sigma}_{j2}^2\widehat{\rho}_j}{\widetilde{\sigma}_{j1}^2\widetilde{\sigma}_{j2}^2} + \frac{n_j\widehat{\sigma}_{j2}^2}{\widetilde{\sigma}_{j2}^2}\right) \\ &= -n\ln(2\pi) - \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(\widehat{\sigma}_{j1}^2\widehat{\sigma}_{j2}^2\right) - \sum_{j=1}^{2} n_j \ln\frac{(1 - \widehat{\rho}\widehat{\rho}_j)^2}{(1 - \widehat{\rho}^2)^2} \\ &- \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1 - \widehat{\rho}^2\right) - n \end{aligned}$$

Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2)) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2) - l\left(\boldsymbol{\theta}_0|\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2\right) \\ &= -\sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1 - \widehat{\rho}_j^2\right) + \sum_{j=1}^{2}\frac{n_j}{2}\ln\left(1 - \widehat{\rho}^2\right) + \sum_{j=1}^{2} n_j\ln\frac{(1 - \widehat{\rho}\widehat{\rho}_j)^2}{(1 - \widehat{\rho}^2)^2} \\ &= \frac{1}{2}\sum_{j=1}^{2} n_j \ln\frac{(1 - \widehat{\rho}\widehat{\rho}_j)^2}{(1 - \widehat{\rho}^2)(1 - \widehat{\rho}_j^2)}. \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{X}_1, \mathbf{Y}_1, \mathbf{X}_2, \mathbf{Y}_2\right)\right) \overset{\mathcal{D}}{\sim} \chi_1^2$.

Vierohodnostný $100 \times (1-\alpha)\%$ **empirický DIS pre** $\rho_1 - \rho_2$ bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{\rho_0 : U_{\text{LR}} < \chi_1^2(\alpha)\right\}.$$

Na výpočet tohoto DIS by sme museli reparametrizovať funkciu vierohodnosti na parameter záujmu $(\rho_1 - \rho_2)$ a rušivý parameter. Alternatívne by bolo možné vypočítať u_{LR} pre rôzne kombinácie ρ_1 a ρ_2 (na vhodne zvolených intervaloch) a pre tieto $\rho_j,\ j = 1, 2$, vypočítať rozdiely $\rho_1 - \rho_2$. Potom $\mathcal{CS}_{1-\alpha} = \left\{\rho_0 : U_{\text{LR}}(\rho_1 - \rho_2) < \chi_1^2(\alpha)\right\}$.

Príklad 236 (test o rozdiele ρ_1 a ρ_2) *Majme dáta* two-samples-correlations-trunk.txt, *premennú dĺžka dolnej končatiny* lowex.L *v milimetroch a dĺžka trupu* tru.L *v milimetroch u mužov a žien. Predpokladáme, že premenné* lowex.L *a* tru.L *majú $N_2(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, kde $j = 1, 2$ predstavujú mužov a ženy. (a) Otestujte hypotézu o rozdiele korelačných koeficientov ρ_1 a ρ_2 na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)\%$ empirický DIS pre rozdiel $\rho_1 - \rho_2$, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite (1) Waldovu testovaciu štatistiku Z_W, (2) testovaciu štatistiku pomerom vierohodnosti U_{LR} a DIS prislúchajúce (1).*

Riešenie (aj v R)

1. **Hypotézy**
   - **slovná formulácia** – testujeme H_0: rozdiel korelačných koeficientov premenných dĺžka dolnej končatiny a dĺžka trupu u mužov a u žien je rovný nule (t.j. korelačné koeficienty u mužov a u žien sa rovnajú) oproti H_1: rozdiel korelačných koeficientov premenných dĺžka dolnej končatiny a dĺžka trupu u mužov a u žien nie je rovný nule (t.j. korelačné koeficienty u mužov a u žien sa nerovnajú).
   - **matematická formulácia** – $H_0 : \rho_1 - \rho_2 = \rho_0$ oproti $H_1 : \rho_1 - \rho_2 \neq \rho_0$, kde $\rho_0 = 0$.

2. **Testovacia štatistika** – Pearsonove korelačné koeficienty ρ_j^2, kde $j = 1, 2$, sú neznáme, a preto ich musíme odhadnúť. Ich odhadmi sú $r_1 \doteq 0.06$ a $r_2 \doteq 0.29$. Rozdiel $r_1 - r_2 \doteq -0.23$.

```
DATA <- read.table("two-samples-correlations-trunk.txt",header=TRUE)
names(DATA)
##   "sex"       "lowex.L"  "tru.L"
attach(DATA)
kor.m <- cor(lowex.L[sex=="m"],tru.L[sex=="m"]) # 0.05975781
kor.f <- cor(lowex.L[sex=="f"],tru.L[sex=="f"]) # 0.285256
kor.diff <- kor.m-kor.f  # -0.2254981
```

   Pozorované testovacie štatistiky:
   Waldova testovacia štatistika $z_W \doteq -1.50$.
   Testovacia štatistika pomerom vierohodnosti $u_{\text{LR}} \doteq 2.33$.

```
xi.m <- atanh(kor.m) # 0.0598291
xi.f <- atanh(kor.f) # 0.285256
xi.diff <- xi.m-xi.f # -0.2335652
n.m <- length(lowex.L[sex=="m"]) # 75
n.f <- length(lowex.L[sex=="f"]) # 100
rozptyl.xi.diff <- 1/(n.m-3)+1/(n.f-3)
z.W <- xi.diff/sqrt(rozptyl.xi.diff) # -1.501471
"rho.hat" <- function(kor,kor.m,kor.f,n.m,n.f) {
    n.m*(kor.m-kor)/(1-kor.m*kor)+n.f*(kor.f-kor)/(1-kor.f*kor)
    }
kor <- uniroot(rho.hat,interval=c(-0.5,0.5),kor.m=kor.m,kor.f=kor.f,n.m=n.m, n.f=n.f)$root # 0.1910674
uLR.obs <- n.m*log((1-kor.m*kor)^2/((1-kor^2)*(1-kor.m^2)))+
           n.f*log((1-kor.f*kor)^2/((1-kor^2)*(1-kor.f^2))) # 2.332354
```

3. **Zamietacia oblasť** –
   Waldov test:
   kritické hodnoty $u(1-\alpha/2) \doteq -1.96$ a $u(\alpha/2) \doteq 1.96$;
   kritický obor $\mathcal{W}_1 = (-\infty, -1.96) \cup (1.96, \infty)$.
   Test pomerom vierohodnosti:
   kritická hodnota $\chi_1^2(\alpha) \doteq 3.841$;
   kritický obor $\mathcal{W} = (\chi_1^2(\alpha), \infty) \doteq (3.841, \infty)$.

```
u.krit.d <- qnorm(0.025) # -1.959964
u.krit.h <- qnorm(0.975) # 1.959964
lr.krit.hodn <- qchisq(0.95,df=1) # 3.841459
```

4. **Empirický dvojstranný interval spoľahlivosti** –
   Waldov $100 \times (1-\alpha)\%$ empirický DIS=95% empirický DIS pre $\rho_1 - \rho_2$:
   $(d, h) = (-0.49, 0.07)$.

```
IS.W <- tanh(xi.diff+c(-1,1)*u.krit.h*sqrt(rozptyl.xi.diff)) # -0.49181570  0.07120136
```

5. **Štatistický záver** –
   Waldov test: p-hodnota $\doteq 2\Pr(Z_W \geq 1.50|H_0) \doteq 0.133$.
   Test pomerom vierohodnosti: p-hodnota=$\Pr(U_{\mathrm{LR}} \geq 2.33|H_0) \doteq 0.127$.

```
p.hodn.Z <- 2*(1-pnorm(abs(z.W))) # 0.1332338
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # 0.1267102
```

   H_0 na hladine významnosti nezamietame, pretože (1) testovacia štatistika nepatrí kritickému oboru, (2) $\rho_0 = 0$ patrí DIS a nakoniec (3) p-hodnota je väčšia ako 0.05.

6. **Slovný záver** – Nezamietame nulovú hypotézu o tom, že rozdiel korelačných koeficientov premenných dĺžka dolnej končatiny a dĺžka trupu u mužov a u žien je rovný nule.

7. **Antropologický slovný záver** – Dĺžka trupu u žien, v oblasti brucha rešpektujúca nároky rastúceho plodu v dobe tehotenstva, by mala výraznejšie súvisieť s dĺžkou dolných končatín v porovnaní s mužmi. Podľa predpokladu vyplývajúceho z reprodukčných rozdielov medzi pohlaviami je korelačný koeficient u žien číselne väčší než u mužov (0.29 vs. 0.06). Na rozdiel od mužov, u žien je korelácia štatisticky významná (p-hodnota = 0.004 vs. 0.611) a oba rozmery spolu súvisia. I keď je medzi pohlaviami relatívne veľký rozdiel v korelačných koeficientoch medzi oboma rozmermi, rozdiel korelačných koeficientov v testovanej vzorke štatisticky významný nebol. Dôvodom môže byť celkovo nevyvážená vzorka (100 žien vs. 75 mužov) a nízky počet pozorovaní v oboch vzorkách. Pri $\alpha = 0.05$, $1 - \beta = 0.8$ a vyváženej vzorke ($n = n_1 = n_2$) by sme potrebovali minimálne $n = 291$, aby bol rozdiel korelačných koeficientov -0.23 štatisticky signifikantný. Iným dôvodom môže byť skladba oboch vzoriek, keďže ide o nepríbuzných mužov a ženy, vybraných z populácie mladých študentov pôvodom takmer z celej Českej a Slovenskej republiky. Najrôznejšie faktory (mätúce premenné: geografické, sociálne, genetické) tak môžu ovplyvňovať (zvyšovať) variabilitu oboch znakov u mužov a žien. Pokiaľ by sme tieto faktory obmedzili (napríklad na vzorky zložené z ľudí určitého geografického územia, obdobných rastových vzorcov alebo rovnaké vzorky zložené z párov brat – sestra), rozdiel korelačných koeficientov by pravdepodobne mohol byť väčší alebo jeho štatistická významnosť vyššia. Je tiež možné, že obmedzením definície rozmeru dĺžky trupu len na oblasť brucha (v našom prípade šlo o rozdiel výšky akromiálnej a výšky spinálnej, t.j. projekčná vzdialenosť *spina iliaca anterior superior* od laterálneho konca kľúčnej kosti) by korelácia u žien mohla byť ešte silnejšia a rozdiel medzi pohlaviami štatisticky významný.

7.4 Asymptotické testy o dvoch pravdepodobnostiach

Kontingenčné tabuľky 2×2 a testy dvoch pravdepodobností patria k častým úlohám v rôznych oblastiach antropológie a ich použitie je veľmi univerzálne. Porovnanie pravdepodobnosti výskytu nejakého javu medzi dvoma skupinami patrí medzi najbežnejšie štatistické úlohy vôbec. V kostrovej antropológii sa uplatňuje napr. pri testovaní rozdielov vo výskyte rôznych epigenetických znakov (pozri úvod ku kapitole 6.4), u živého človeka pri testovaní akýchkoľvek kvalitatívnych znakov, u ktorých zaznamenávame ich prítomnosť alebo neprítomnosť. V oblasti medicíny možno týmto spôsobom testovať rozdiel vo výskyte chorôb (*morbiditu*) i úmrtnosť (*fatalita*) na určitú chorobu.

Nech $X_j \sim Bin(N_j, p_j), j = 1, 2$, sú dve nezávislé náhodné premenné. Realizácie x_j môžeme usporiadať do nasledovnej **kontingenčnej tabuľky** 2×2 **pre početnosti**

	odpoveď 1 (má znak X)	odpoveď 0 (nemá znak X)	spolu
skupina 1	n_1	$N_1 - n_1$	N_1
skupina 2	n_2	$N_2 - n_2$	N_2

a **kontingenčnej tabuľky** 2×2 **pre pravdepodobnosti**

	odpoveď 1 (má znak X)	odpoveď 0 (nemá znak X)	spolu
skupina 1	p_1	$1 - p_1$	1
skupina 2	p_2	$1 - p_2$	1

Kontingenčnú tabuľku 2×2 pre početnosti je možné preznačiť nasledovne

	odpoveď 1 (má znak X)	odpoveď 0 (nemá znak X)	spolu
skupina 1	n_{11}	n_{12}	$N_{1\bullet}$
skupina 2	n_{21}	n_{22}	$N_{2\bullet}$
suma	$N_{\bullet 1}$	$N_{\bullet 2}$	N

Početnosti $n_{ij}, j = 1, 2$, sa nazývajú **združené početnosti**, početnosti $N_{\bullet j} = n_{1j} + n_{2j}$ **marginálne stĺpcové početnosti**, početnosti $N_{j\bullet} = n_{j1} + n_{j2}$ **marginálne riadkové početnosti** a početnosť $N = N_{\bullet 1} + N_{\bullet 2} = N_{1\bullet} + N_{2\bullet} = N_1 + N_2$ sa nazýva **celková (totálna) početnosť**. Kontingenčnú tabuľku 2×2 pre pravdepodobnosti je možné preznačiť nasledovne

	odpoveď 1 (má znak X)	odpoveď 0 (nemá znak X)	spolu
skupina 1	p_{11}	p_{12}	$p_{1\bullet}$
skupina 2	p_{21}	p_{22}	$p_{2\bullet}$
suma	$p_{\bullet 1}$	$p_{\bullet 2}$	1

Pravdepodobnosti $p_{ij} = \frac{n_{ij}}{N}, j = 1, 2$, sa nazývajú **združené pravdepodobnosti**, pravdepodobnosti $p_{\bullet j} = \frac{N_{\bullet j}}{N}$ **marginálne stĺpcové pravdepodobnosti**, pravdepodobnosti $p_{j\bullet} = \frac{N_{j\bullet}}{N}$ **marginálne riadkové pravdepodobnosti**. Alternatívnym preznačením dostaneme **kontingenčnú tabuľku** 2×2 **pre podmienené pravdepodobnosti**

	odpoveď 1 (má znak X)	odpoveď 0 (nemá znak X)	spolu
skupina 1	$p_{1\|1}$	$p_{2\|1}$	1
skupina 2	$p_{1\|2}$	$p_{2\|2}$	1

Ide o **podmienené pravdepodobnosti odpovede (výskytu znaku) za podmienky príslušnosti do skupiny**, kde $p_{1|1} = p_1$, $p_{1|2} = p_2$, $p_{2|1} = 1 - p_1$ a $p_{2|2} = 1 - p_2$.

Nech $\boldsymbol{\theta} = (p_1, p_2)^T$. Potom môžeme zhrnúť najčastejšie používané funkcie $g(\boldsymbol{\theta})$ nasledovne:

názov	označenie	$g(\boldsymbol{\theta})$	rozpätie	nulový bod
rozdiel rizík	RD	$p_1 - p_2$	$\langle -1, 1 \rangle$	0
relatívne riziko	RR	$\frac{p_1}{p_2}$	$\langle 0, \infty)$	1
logaritmus relatívneho rizika	$\ln$ RR	$\ln \frac{p_1}{p_2}$	$(-\infty, \infty)$	0
pomer šancí	OR	$\frac{p_1/(1-p_1)}{p_2/(1-p_2)}$	$\langle 0, \infty)$	1
logaritmus pomeru šancí	$\ln$ OR	$\ln \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$	$(-\infty, \infty)$	0

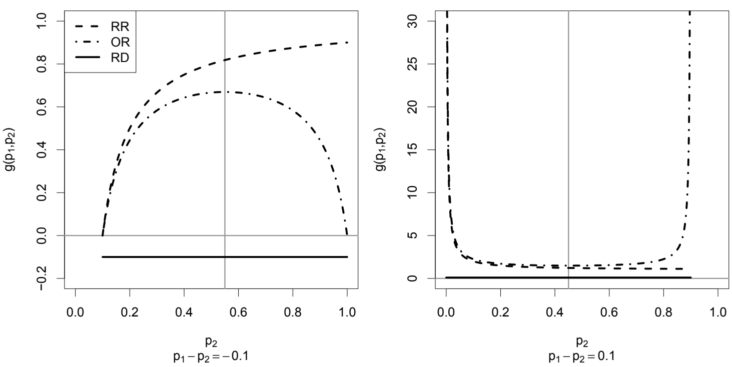
Rozdiel rizík predstavuje rozdiel pravdepodobností p_1 a p_2 výskytu sledovaného znaku X v dvoch populáciách. **Relatívne riziko** predstavuje podiel pravdepodobností p_1 a p_2 výskytu sledovaného znaku X v dvoch populáciách. **Pomer šancí** predstavuje podiel šancí $p_1/(1 - p_1)$ a $p_2/(1 - p_2)$ pre výskyt znaku X. Logaritmovanie podielu rizík a pomeru šancí spôsobuje zmenu nesymetrického rozdelenia $g(\boldsymbol{\theta})$ za platnosti H_0 ($g(\boldsymbol{\theta})$ je rovné svojmu nulovému bodu) na symetrické rozdelenie. Voľba $g(\boldsymbol{\theta})$ v praxi závisí na konkrétnej aplikácii. Avšak relatívne riziko nie je možné priamo odhadnúť (bez doplňujúcich informácií) v retrospektívnych štúdiách prípadov a kontrol. Ale pomer šancí priamo odhadnuteľný je a navyše za predpokladu, že choroba je zriedkavá, je aproximáciou relatívneho rizika. Lekári, biológovia a antropológovia považujú relatívne riziko za najvhodnejšiu $g(\boldsymbol{\theta})$. V matematickej štatistike je však najčastejšie používaným $g(\boldsymbol{\theta})$ pomer šancí, ktorý je prirodzeným parametrom funkcie vierohodnosti a parameter, ktorý sa odhaduje aj v logistickom regresnom modeli.

Medzi jednotlivými $g(\boldsymbol{\theta})$ je nelineárny vzťah. Vyberajme p_2 z intervalu $(0, 1)$ a ponechajme rozdiel rizík konštantný a rovný -0.1. Keď p_2 rastie, relatívne riziko je monotónnou funkciou rastúcou smerom k jeho nulovému bodu jedna, indikujúc proporčne menšiu redukciu rizika. Ak p_1 ide k nule a p_2 ide k jednotke, pomer šancí sa približne rovná nule. Ak p_2 rastie, pomer šancí rastie, dosiahne maximum približne rovné 0.669 v bode $\widetilde{p}_2 = 0.5(1 - \mathrm{RD}) = 0.55$ a potom klesá k nule (pozri obrázok 7.4 vľavo; pre RD=-0.5 pozri obrázok 7.5 vľavo).

Vyberajme p_2 opäť z intervalu $(0, 1)$ a ponechajme rozdiel rizík konštantný a rovný 0.1. Keď p_2 rastie, relatívne riziko je monotónnou funkciou klesajúcou smerom k jeho nulovému bodu jedna, indikujúc proporčne menšiu redukciu rizika. Ak p_1 ide k nule a p_2 ide k jednotke, pomer šancí sa rovná nejakému číslu oveľa väčšiemu ako jedna. Ak p_2 rastie, pomer šancí klesá, dosiahne minimum približne rovné 1.494 v bode $\widetilde{p}_2 = 0.5(1 - \mathrm{RD}) = 0.45$ a potom rastie (pozri obrázok 7.4 vpravo; pre RD=-0.5 pozri obrázok 7.5 vpravo).

Majme dve štúdie A a B s rovnakým záporným rozdielom rizík $\mathrm{RD} = -0.1$ (t.j. $p_1 < p_2$), ale pravdepodobnosti sú rozdielne, t.j. $\mathrm{RD}_A = \mathrm{RD}_B$, $p_{A1} < p_{B1}$ a $p_{A2} < p_{B2}$. Potom $\mathrm{RR}_A \neq \mathrm{RR}_B$ a štúdia s väčším p_2 bude mať RR bližsie k jednotke (RR bude väčšie). Pomer šancí sa tiež bude odlišovať, t.j. $\mathrm{OR}_A \neq \mathrm{OR}_B$, ak obe p_2 sú menšie ako $\widetilde{p}_2$ alebo obe p_2 sú väčšie ako $\widetilde{p}_2$. Ak $p_{A2} < \widetilde{p}_2 < p_{B2}$, môže nastať situácia, že $\mathrm{OR}_A = \mathrm{OR}_B$. Napr. ak $\widetilde{p}_2 = 0.5(1 - \mathrm{RD}) = 0.55$, potom $p_{A2} = 0.46$ a $p_{B2} = 0.64$, $\mathrm{OR}_A = \frac{p_{A1}/(1-p_{A1})}{p_{A2}/(1-_{A2})} = 0.6581197$ a $\mathrm{OR}_B = \frac{p_{B1}/(1-p_{B1})}{p_{B2}/(1-_{B2})} = 0.6581197$, pretože $|\widetilde{p}_2 - p_{A2}| = |p_{B2} - \widetilde{p}_2|$.

Majme dve štúdie A a B s rovnakým kladným rozdielom rizík $\mathrm{RD} = 0.1$ (t.j. $p_1 < p_2$), ale pravdepodobnosti sú rozdielne, t.j. $\mathrm{RD}_A = \mathrm{RD}_B$, $p_{A1} > p_{B1}$ a $p_{A2} > p_{B2}$. Potom $\mathrm{RR}_A \neq \mathrm{RR}_B$ a štúdia s väčším p_2 bude mať RR bližšie k jednotke (RR bude menšie). Pomer šancí sa tiež bude odlišovať, t.j. $\mathrm{OR}_A \neq \mathrm{OR}_B$, ak obe p_2 sú menšie ako $\widetilde{p}_2$ alebo obe p_2 sú väčšie ako $\widetilde{p}_2$. Ak $p_{A2} < \widetilde{p}_2 < p_{B2}$, môže nastať situácia, že $\mathrm{OR}_A = \mathrm{OR}_B$. Napr. ak $\widetilde{p}_2 = 0.5(1 - \mathrm{RD}) = 0.45$, potom $p_{A2} = 0.35$ a $p_{B2} = 0.55$, $\mathrm{OR}_A = \frac{p_{A1}/(1-p_{A1})}{p_{A2}/(1-_{A2})} = 1.519481$ a $\mathrm{OR}_B = \frac{p_{B1}/(1-p_{B1})}{p_{B2}/(1-_{B2})} = 1.519481$, pretože $|\widetilde{p}_2 - p_{A2}| = |p_{B2} - \widetilde{p}_2|$.



Obr. 7.4: Nelineárny vzťah rozdielu rizík, pomeru šancí a pomeru rizík (vľavo $\mathrm{OR}_{\max} = 0.669$; vpravo $\mathrm{OR}_{\max} = 112.235$ a $\mathrm{OR}_{\min} = 1.494$)

Medzi p_1 a OR existuje nasledovný vzťah

$$p_1 = \frac{p_2 \times \mathrm{OR}}{1 - p_2 + p_2 \times \mathrm{OR}}$$

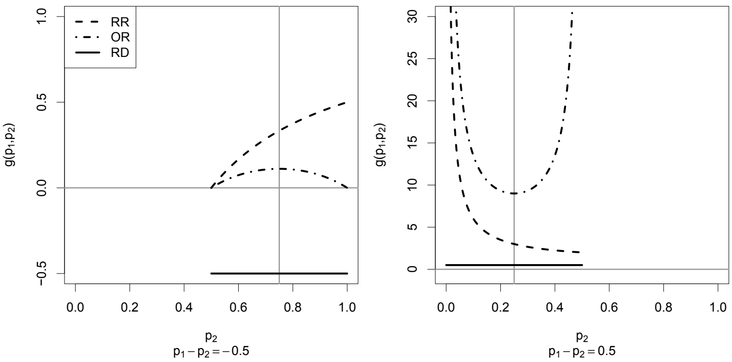
a medzi RR a OR platí

$$\mathrm{RR} = \mathrm{OR}\frac{1 + n_{21}/n_{22}}{1 + n_{11}/n_{12}}.$$

Rozdiel rizík p_1 **a** p_2. Majme dvojice hypotéz $H_{01} : p_1 - p_2 = p_0$ oproti $H_{11} : p_1 - p_2 \neq p_0$, $H_{02} : p_1 - p_2 \leq p_0$ oproti $H_{12} : p_1 - p_2 > p_0$, $H_{03} : p_1 - p_2 \geq p_0$ oproti $H_{13} : p_1 - p_2 < p_0$, ktoré chceme testovať. Hypotézy H_0 sa často nazývajú **hypotézy homogenity dvoch binomických rozdelení**, ak $p_0 = 0$.

Ak H_0 platí, potom

$$Z_W = \frac{X_1/N_1 - X_2/N_2 - p_0}{S_g} \overset{\mathcal{D}}{\sim} N(0, 1),$$

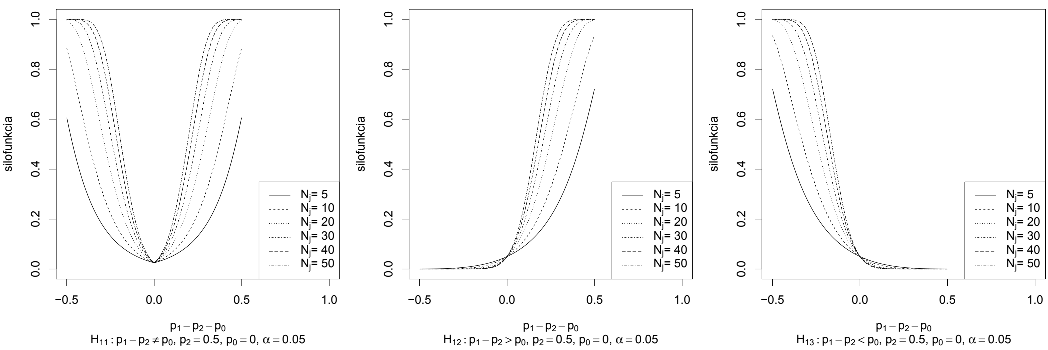


Obr. 7.5: Nelineárny vzťah rozdielu rizík, pomeru šancí a pomeru rizík (vľavo $\mathrm{OR}_{\max} = 0.111$; vpravo $\mathrm{OR}_{\max} = 1003.004$ a $\mathrm{OR}_{\min} = 9.000$)

kde $S_g^2 = \frac{X_1/N_1(1-X_1/N_1)}{N_1} + \frac{X_2/N_2(1-X_2/N_2)}{N_2} \neq 0$. Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **dvojvýberový** Z**-test o rozdiele pravdepodobností** $p_1 - p_2$.

Definícia 79 (Kritický obor a silofunkcia Z_W testu o $p_1 - p_2$) *Kritický obor a silofunkcia je definovaná nasledovne (ozn. $\sigma_g^2 = \frac{p_1(1-p_1)}{N_1} + \frac{p_2(1-p_2)}{N_2}$; pozri obrázok 7.6):*

H_0	H_1	$\mathcal{W}$	$1-\beta(p_1,p_2)$
$p_1 - p_2 = p_0$	$p_1 - p_2 \neq p_0$	$\mathcal{W}_1 = \{Z_W; \lvert Z_W\rvert \geq u(\alpha/2)\}$	$1 - \Phi\left(u_{\alpha/2} - \frac{\lvert p_0-(p_1-p_2))\rvert}{\sigma_g}\right)$
$p_1 - p_2 \leq p_0$	$p_1 - p_2 > p_0$	$\mathcal{W}_2 = \{Z_W; Z_W \geq u(\alpha)\}$	$1 - \Phi\left(u_{\alpha} + \frac{p_0-(p_1-p_2)}{\sigma_g}\right)$
$p_1 - p_2 \geq p_0$	$p_1 - p_2 < p_0$	$\mathcal{W}_3 = \{Z_W; Z_W \leq -u(\alpha)\}$	$1 - \Phi\left(u_{\alpha} - \frac{p_0-(p_1-p_2)}{\sigma_g}\right)$



Obr. 7.6: Silofunkcie Waldovho testu o $p_1 - p_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 80 (p-hodnota Z_W testu o $p_1 - p_2$) *Nech Z_W je* ***nejaká testovacia štatistika*** *a $z_W = \frac{\widehat{p}_1-\widehat{p}_2-p_0}{s_g}$, kde $\widehat{\sigma}_g^2 = s_g^2 = \frac{\widehat{p}_1(1-\widehat{p}_1)}{N_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{N_2}$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W \geq |z_W| | H_{01}), \ ak & H_{11}: p_1 - p_2 \neq p_0 \\ Pr(Z_W \geq z_W | H_{02}), \ ak & H_{12}: p_1 - p_2 > p_0 \\ Pr(Z_W \leq z_W | H_{03}), \ ak & H_{13}: p_1 - p_2 < p_0 \end{cases}$$

Definícia 81 ((Klasické) Waldove $100 \times (1-\alpha)\%$ empirické IS pre $p_1 - p_2$) *(Klasické) Waldove $100 \times (1-\alpha)\%$ empirické IS pre $p_1 - p_2$ pre všetky tri typy hypotéz majú nasledovný tvar*

H_0	H_1	*hranice (d,h) pre $100(1-\alpha)\%$ empirický IS*
$p_1 - p_2 = p_0$	$p_1 - p_2 \neq p_0$	$\mathcal{CS}_{1-\alpha} = \left\{p_0 : p_0 \in \left(\widehat{p}_1 - \widehat{p}_2 - u_{\alpha/2}s_g, \widehat{p}_1 - \widehat{p}_2 + u_{\alpha/2}s_g\right)\right\}$
$p_1 - p_2 \leq p_0$	$p_1 - p_2 > p_0$	$\mathcal{CS}_{1-\alpha} = \{p_0 : p_0 \in (\widehat{p}_1 - \widehat{p}_2 - u_\alpha s_g, 1)\}$
$p_1 - p_2 \geq p_0$	$p_1 - p_2 < p_0$	$\mathcal{CS}_{1-\alpha} = \{p_0 : p_0 \in (-1, \widehat{p}_1 - \widehat{p}_2 + u_\alpha s_g)\}$

Majme H_{11}. Potom **minimálny rozsah** definujeme ako

$$N_1 \geq \left(\frac{u_{\alpha/2} + u_\beta}{p_1 - p_2}\right)^2 \left(p_1\,(1-p_1) + \frac{p_2\,(1-p_2)}{k}\right), N_2 = kN_1.$$

Pre H_{12} a H_{13} zameníme $u_{\alpha/2}$ za u_α.

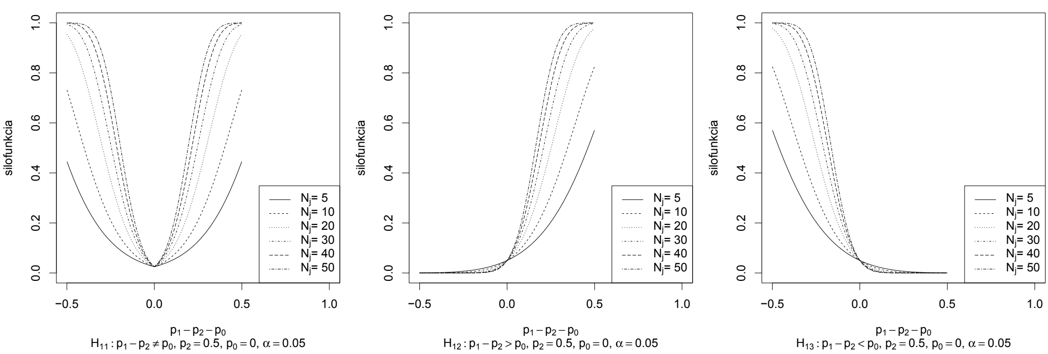
Ak za platnosti H_0 berieme pri výpočte do úvahy, že $p_1 = p_2$, potom dostaneme alternatívnu testovaciu štatistiku, ktorá bude mať za platnosti H_0 nasledovný tvar

$$Z_W^{(\mathrm{alt})} = \frac{X_1/N_1 - X_2/N_2 - p_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde $S_g^2 = X/N\,(1 - X/N)\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \neq 0$, $N = N_1 + N_2$ a $X = N_1 X_1/N_1 + N_2 X_2/N_2$. $Z_W^{(\mathrm{alt})}$ sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **alternatívny dvojvýberový** Z-**test o rozdiele pravdepodobností** $p_1 - p_2$.

Definícia 82 (Kritický obor a silofunkcia $Z_W^{(\mathrm{alt})}$ testu o $p_1 - p_2$) *Kritický obor a silofunkcia sú definované nasledovne (ozn. $\sigma_g^2 = p\,(1-p)\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$; pozri obrázok 7.7):*

H_0	H_1	$\mathcal{W}$	$\beta\,(p_1, p_2)$
$p_1 - p_2 = p_0$	$p_1 - p_2 \neq p_0$	$\mathcal{W}_1 = \left\{Z_W^{(alt)}; \lvert Z_W^{(alt)}\rvert \geq u\,(\alpha/2)\right\}$	$\Phi\left(u_{\alpha/2} - \frac{\lvert p_0 - (p_1 - p_2)\rvert}{\sigma_g}\right)$
$p_1 - p_2 \leq p_0$	$p_1 - p_2 > p_0$	$\mathcal{W}_2 = \left\{Z_W^{(alt)}; Z_W^{(alt)} \geq u\,(\alpha)\right\}$	$\Phi\left(u_\alpha + \frac{p_0 - (p_1 - p_2)}{\sigma_g}\right)$
$p_1 - p_2 \geq p_0$	$p_1 - p_2 < p_0$	$\mathcal{W}_3 = \left\{Z_W^{(alt)}; Z_W^{(alt)} \leq -u\,(\alpha)\right\}$	$\Phi\left(u_\alpha - \frac{p_0 - (p_1 - p_2)}{\sigma_g}\right)$



Obr. 7.7: Silofunkcie alternatívneho Waldovho testu o $p_1 - p_2$ pre H_{11} (vľavo), H_{12} (uprostred) a H_{13} (vpravo)

Definícia 83 (p-hodnota $Z_W^{(\mathbf{alt})}$ testu o $p_1 - p_2$) *Nech $Z_W^{(alt)}$ je* ***nejaká testovacia štatistika*** *a $z_W^{(alt)} = \frac{\widehat{p}_1 - \widehat{p}_2 - p_0}{s_g}$, kde $\widehat{\sigma}_g^2 = s_g^2 = \widehat{p}\,(1-\widehat{p})\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$, je jej realizácia (****pozorovaná, vypočítaná testovacia štatistika****), potom p-hodnotu počítame nasledovne:*

$$\text{p-hodnota} = \begin{cases} 2Pr(Z_W^{(alt)} \geq |z_W^{(alt)}||H_{01}),\ ak & H_{11} : p_1 - p_2 \neq p_0 \\ Pr(Z_W^{(alt)} \geq z_W^{(alt)}|H_{02}),\ ak & H_{12} : p_1 - p_2 > p_0 \\ Pr(Z_W^{(alt)} \leq z_W^{(alt)}|H_{03}),\ ak & H_{13} : p_1 - p_2 < p_0 \end{cases}$$

Definícia 84 (Alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $p_1 - p_2$) *Alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $p_1 - p_2$ pre všetky tri typy hypotéz majú nasledovný tvar:*

H_0	H_1	*hranice (d, h) pre $100(1-\alpha)\%$ empirický IS*
$p_1 - p_2 = p_0$	$p_1 - p_2 \neq p_0$	$\mathcal{CS}_{1-\alpha} = \left\{p_0 : p_0 \in \left(\widehat{p}_1 - \widehat{p}_2 - u_{\alpha/2}s_g, \widehat{p}_1 - \widehat{p}_2 + u_{\alpha/2}s_g\right)\right\}$
$p_1 - p_2 \leq p_0$	$p_1 - p_2 > p_0$	$\mathcal{CS}_{1-\alpha} = \{p_0 : p_0 \in (\widehat{p}_1 - \widehat{p}_2 - u_\alpha s_g, \infty)\}$
$p_1 - p_2 \geq p_0$	$p_1 - p_2 < p_0$	$\mathcal{CS}_{1-\alpha} = \{p_0 : p_0 \in (-\infty, \widehat{p}_1 - \widehat{p}_2 + u_\alpha s_g)\}$

Majme H_{11}. Potom **minimálny rozsah** definujeme ako

$$N_1 \geq \left(\frac{u_{\alpha/2} + u_\beta}{p_1 - p_2}\right)^2 p\,(1-p)\left(\frac{k+1}{k}\right), N_2 = kN_1.$$

Ak $N_1 = N_2 = N$, tak $p = \frac{Np_1 + Np_2}{2N} = \frac{p_1 + p_2}{2}$. Ak $N_1 \neq N_2$, tak $p = \frac{N_1 p_1 + kN_1 p_2}{N_1 + kN_1} = \frac{p_1 + kp_2}{1+k}$. Pre H_{12} a H_{13} zameníme $u_{\alpha/2}$ za u_α.

Príklad 237 (minimálny rozsah) *Sledujeme výskyt určitej choroby v ČR a v SR u žien. V minulosti sme zistili, že daná choroba sa nevyskytuje u žien s $p_1 = 0.6$ a $p_2 = 0.7$ a stanovili sme klinicky dôležitý rozdiel ako $p_2 - p_1 = 0.1$ na základe tejto minulej skúsenosti a vedomostí o tejto chorobe. Ďalej nech sila Z-testu pri dojstrannej alternatíve je $1 - \beta = 0.90$ (pravdepodobnosť odhalenia tohoto rozdielu) a hladina významnosti $\alpha = 0.05$. Minimálne koľko žien musíme sledovať v ČR a v SR pri danej α a $1-\beta$? Vypočítajte (a) pre $N_1 = N_2 = N$ ako aj pre (b) $N_1 = 2N_2$ (skutočnosť na základe počtu obyvateľov v ČR a v SR). Použite (1) Z_W a (2) $Z_W^{(alt)}$.*

Riešenie

```
"min.rozsah.n.p" <- function(p1,p2,alfa,sila,k){
N1 <- ((qnorm(1-alfa/2)+qnorm(sila))/(p1-p2))^2*(p1*(1-p1)+p2*(1-p2)/k)
N2 <- k*N1
N <- round(c(N1,N2))
p <- (p1+k*p2)/(1+k)
N1.alt <- ((qnorm(1-alfa/2)+qnorm(sila))/(p1-p2))^2*(p*(1-p)*(k+1)/k)
N2.alt <- k*N1.alt
N.alt <- ceiling(c(N1.alt,N2.alt))
rozsahy <- list(N=N,N.alt=N.alt)
return(rozsahy)
}
min.rozsah.n.p(0.6,0.7,0.05,0.9,2)
min.rozsah.n.p(0.6,0.7,0.05,0.9,1)
```

Pri použití Z_W sú minimálne rozsahy nasledovné – $N_1 = N_2 = 473$ pre homogénny dizajn; $N_1 = 363$ a $N_2 = 725$ pre nehomogénny dizajn. Pri použití $Z_W^{(\mathrm{alt})}$ sú minimálne rozsahy nasledovné – $N_1 = N_2 = 479$ pre homogénny dizajn; $N_1 = 351$ a $N_2 = 701$ pre nehomogénny dizajn.

Príklad 238 (porovnanie testovacích štatistík) *Ukážte, že ak $N_1 = N_2$, potom $|Z_W^{(alt)}| \leq |Z_W|$. Naviac ak $X_1 = X_2$, potom $|Z_W^{(alt)}| = |Z_W|$.*

Riešenie
Funkcia $f(x) = x(1-x)$ je konkávna na $\langle 0,1\rangle$. Pre ľubovoľné $\gamma \in \langle 0,1\rangle$ a pre ľubovoľné čísla $a, b \in \langle 0,1\rangle$ platí

$$f(\gamma a + (1-\gamma) b) \geq \gamma f(a) + (1-\gamma) f(b).$$

Ak zvolíme $\gamma = \frac{1}{2}, a = p_1, b = p_2$, potom dostaneme

$$\frac{1}{2}(p_1+p_2)\left[1 - \frac{1}{2}(p_1+p_2)\right] \geq \frac{1}{2}p_1(1-p_1) + \frac{1}{2}p_2(1-p_2).$$

Ak $N_1 = N_2$, máme $p = \frac{p_1+p_2}{2}$. Menovateľ Z_W má pod odmocninou

$$\tfrac{2}{N_1}\left[\tfrac{1}{2}p_1(1-p_1) + \tfrac{1}{2}p_2(1-p_2)\right],$$

menovateľ $Z_W^{(\text{alt})}$ má pod odmocninou

$$\tfrac{2}{N_1}\left[\tfrac{p_1+p_2}{2}\left(1 - \tfrac{p_1+p_2}{2}\right)\right].$$

Z nerovnosti uvedenej vyššie vyplýva, že $|Z_W^{(\text{alt})}| \leq |Z_W|$. Rovnosť v nerovnosti nastáva len vtedy, ak $p_1 = p_2$.

Pri $N_1 \neq N_2$, nie je medzi $Z_W^{(\text{alt})}$ a Z_W žiadny jednoduchý vzťah. Zdanlivo by pri $N_1 = N_2$ test založený na Z_W bol výhodnejší, lebo ma väčšiu silu $1-\beta$ pri danej alternatíve, ale pri malom rozsahu výberu nám prináša väčšiu pravdepodobnosť CHPD než požadovaná α. Ak napr. pri $\alpha = 0.05$, $N_1 = N_2 = 20$, Pr(CHPD) $= 0.081$ a pri $N_1 = 20, N_2 = 40$ je Pr(CHPD) $= 0.085$.

Z predchádzajúceho príkladu vyplýva, že ak $N_1 = N_2$, potom minimálny rozsah N_1 vypočítaný z testovacej štatistiky $Z_W^{(\text{alt})}$ bude vždy väčší ako rozsah N_1 vypočítaný zo Z_W.

Dvojvýberový χ^2-test o rozdiele pravdepodobností $p_1 - p_2$ v R (funkcia **prop.test()**).
Argumenty (vstupy) funkcie:

1. vektor početností úspechov (n_1 a n_2) x;
2. vektor rozsahov (početnosti úspechov spolu s neúspechmi N_1 a N_2) n;
3. vektor pravdepodobností p_{01} a p_{02} za platnosti H_0 p ($p_{0j} \in (0,1)$, kde $j = 1,2$);
4. formulácia alternatívy alternative="two.sided" je prednastavené, ďalšie voľby sú "greater", "less";
5. spoľahlivosť conf.level=p, prednastavené je conf.level=.95;
6. Yatesova korekcia na spojitosť (Newcombe, 1998), prednastavené correct = TRUE.

Výstupy funkcie:

1. testovacia štatistika statistic ($\chi^2 = (Z_W^{(\text{alt})})^2$);
2. stupne voľnosti df (parameter; $df = 1$);
3. p-hodnota p.value;
4. odhad pravdepodobností $\widehat{p}_1$ a $\widehat{p}_2$ estimate;
5. interval spoľahlivosti pre $p_1 - p_2$ conf.int (prislúchajúci Z_W a nie $Z_W^{(\text{alt})}$);
6. alternatívna hypotéza alternative hypothesis;
7. názov použitého testu method.

Príklad 239 (programovanie; Z-testy rozdielu $p_1 - p_2$) *Naprogramujte v R klasický a alternatívny (a) Z-test o rozdiele pravdepodobností p_1 a p_2 a (b) Waldov 95% empirický DIS rozdielu pravdepodobností p_1 a p_2.*

Riešenie (aj v R)

```
"test.rozdielu.prav" <- function(n1,n2,N1,N2){
p1.hat <- n1/N1
p2.hat <- n2/N2
# klasicky Z-test a IS
sg.sq <- p1.hat*(1-p1.hat)/N1+p2.hat*(1-p2.hat)/N2
zW <- (p1.hat-p2.hat)/sqrt(sg.sq)
p.hodn <- 1-pnorm(abs(zW))
IS <- p1.hat-p2.hat+c(-1,1)*qnorm(0.975)*sqrt(sg.sq)
Ztest <- c(IS,zW,p.hodn)
# alternativny Z-test a IS
p.hat <- (N1*p1.hat+N2*p2.hat)/(N1+N2)
sg.sq.alt <- p.hat*(1-p.hat)*(1/N1+1/N2)
zW.alt <- (p1.hat-p2.hat)/sqrt(sg.sq.alt)
p.hodn.alt <- 1-pnorm(abs(zW.alt))
ISalt <- p1.hat-p2.hat+c(-1,1)*qnorm(0.975)*sqrt(sg.sq.alt)
Ztest.alt <- c(ISalt,zW.alt,p.hodn.alt)
# vysledky
Z.test <- rbind(Ztest,Ztest.alt)
dimnames(Z.test)[[2]] <- c("DH","HH","Z-stat","p-hodnota")
dimnames(Z.test)[[1]] <- c("Ztest(p1-p2)","Ztest.alt(p1-p2)")
# odhady
odhady <- c(p.hat,p1.hat,p2.hat,p1.hat-p2.hat,sqrt(sg.sq),sqrt(sg.sq.alt))
names(odhady) <- c("p","p1","p2","rozdiel","sd","sd.alt")
VYSL <- list(odhady=round(odhady,4),Z.test=round(Z.test,4))
return(VYSL)
}
```

Príklad 240 (zuby a tepelný šok; rozdiel rizík) *Z 50 zubov vystavených tepelnému šoku sa 21 zlomilo. Z kontrolných 50 zubov sa zlomilo len 11. (a) Znižuje tepelný šok mechanickú odolnosť zubov? Porovnajte výsledky klasického a alternatívneho Z-testu rozdielu pravdepodobností. (b) Vypočítajte klasický a alternatívny Waldov 95% empirický DIS rozdielu pravdepodobností p_1 a p_2 zlomenia sa zuba.*

Riešenie (aj v R)
$N_1 = N_2 = 50, n_1 = 21, n_2 = 11, \widehat{p}_1 = 0.42, \widehat{p}_2 = 0.22, \widehat{p} = 0.32.$
$z_W \doteq 2.195$, p-hodnota $\doteq 0.014$ a H_0 zamietame na $\alpha = 0.05$.
Waldov 95% empirický DIS rozdielu pravdepodobností p_1 a p_2: $(d,h) \doteq (0.0214, 0.3786)$
$z_W^{(\mathrm{alt})} \doteq 2.144$, p-hodnota $\doteq 0.016$ a H_0 zamietame na $\alpha = 0.05$.
Alternatívny Waldov 95% empirický DIS rozdielu pravdepodobností p_1 a p_2: $(d,h) \doteq (0.017, 0.383)$

```
test.rozdielu.prav(21,11,50,50)
##$odhady
#      p     p1     p2  rozdiel     sd  sd.alt
# 0.3200 0.4200 0.2200  0.2000  0.0911  0.0933
##$Z.test
#                      DH     HH  Z-stat  p-hodnota
#Ztest(p1-p2)      0.0214 0.3786  2.1948     0.0141
#Ztest.alt(p1-p2)  0.0171 0.3829  2.1437     0.0160
```

Relatívne riziko p_1/p_2. Majme dvojice hypotéz $H_{01} : \ln \mathrm{RR} = \ln \mathrm{RR}_0$ oproti $H_{11} : \ln \mathrm{RR} \neq \ln \mathrm{RR}_0$, $H_{02} : \ln \mathrm{RR} \leq \ln \mathrm{RR}_0$ oproti $H_{12} : \ln \mathrm{RR} > \ln \mathrm{RR}_0$, $H_{03} : \ln \mathrm{RR} \geq \ln \mathrm{RR}_0$ oproti $H_{13} : \ln \mathrm{RR} < \ln \mathrm{RR}_0$, ktoré chceme testovať. Hypotézy H_0 sa často nazývajú **hypotézy homogenity dvoch binomických rozdelení**, ak $\ln \mathrm{RR}_0 = 0$.

Ak H_0 platí, potom

$$Z_W = \frac{\ln \frac{X_1/N_1}{X_2/N_2} - \ln \mathrm{RR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde $S_g^2 = Var[\ln(X_1/N_1)] + Var[\ln(X_2/N_2)] = \frac{1-X_1/N_1}{X_1} + \frac{1-X_2/N_2}{X_2} \neq 0$. Z_W sa nazýva **Waldova testovacia štatistika** (alebo **Z-štatistika**) a test **dvojvýberový Z-test o logaritme relatívneho rizika**. Rozptyl σ_g^2 vypočítame pomocou δ-metódy z rozptylov $Var[\ln p_j], j = 1,2$, kde

$$Var[\ln p_j] \approx \left(\frac{\partial \ln p_j}{\partial p_j}\right)^2 Var[p_j] = \left(\frac{1}{p_j}\right)^2 \frac{p_j(1-p_j)}{N_j} = \frac{1-p_j}{N_j p_j}.$$

Asymptoticky

$$Z_W = \frac{\ln(X_j/N_j) - \ln p_0}{\sqrt{Var[\ln(X_j/N_j)]}} \overset{\mathcal{D}}{\sim} N(0,1).$$

Keďže $\ln \mathrm{RR} = \ln p_1 - \ln p_2$ a p_1 a p_2 sú nezávislé, potom (Katz a kol., 1978)

$$\sigma_g^2 = Var[\ln \mathrm{RR}] = Var[\ln p_1] + Var[\ln p_2] = \frac{1-p_1}{N_1 p_1} + \frac{1-p_2}{N_2 p_2} = \frac{1-p_1}{n_1} + \frac{1-p_2}{n_2} = \frac{1}{n_1} - \frac{1}{N_1} + \frac{1}{n_2} - \frac{1}{N_2}.$$

Definícia 85 (p-hodnota Z_W testu o $\ln \mathbf{RR}$) *Nech Z_W je **nejaká testovacia štatistika** a $z_W = \frac{\ln \widehat{RR} - \ln RR_0}{s_g}$, kde $\ln \widehat{RR} = \ln \frac{\widehat{p}_1}{\widehat{p}_2}$ a $\widehat{\sigma}_g^2 = s_g^2 = \frac{1-\widehat{p}_1}{N_1 \widehat{p}_1} + \frac{1-\widehat{p}_2}{N_2 \widehat{p}_2}$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$\textit{p-hodnota} = \begin{cases} 2Pr(Z_W \geq |z_W| | H_{01}), \ ak & H_{11} : \ln RR \neq \ln RR_0 \\ Pr(Z_W \geq z_W | H_{02}), \ ak & H_{12} : \ln RR > \ln RR_0 \\ Pr(Z_W \leq z_W | H_{03}), \ ak & H_{13} : \ln RR < \ln RR_0 \end{cases}$$

Ak za platnosti H_0 berieme pri výpočte do úvahy, že $p_1 = p_2$, potom dostaneme alternatívny rozptyl, ktorý bude rovný

$$\sigma_g^2 = Var[\ln \mathrm{RR}] = \frac{1-p}{p}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) = \frac{1-p}{p}\frac{N}{N_1 N_2},$$

kde $N = N_1 + N_2$. Potom

$$Z_W^{(\mathrm{alt})} = \frac{\ln \frac{X_1/N_1}{X_2/N_2} - \ln \mathrm{RR}_0}{S_g} \overset{\mathcal{D}}{\approx} N(0,1),$$

kde $S_g^2 = \frac{1-X/N}{X/N}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \neq 0$, kde $X = N_1 X_1/N_1 + N_2 X_2/N_2$. $Z_W^{(\mathrm{alt})}$ sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **alternatívny dvojvýberový Z-test o logaritme relatívneho rizika**.

Definícia 86 (p-hodnota $Z_W^{(\mathrm{alt})}$ testu o $\ln \mathbf{RR}$) *Nech $Z_W^{(alt)}$ je **nejaká testovacia štatistika** a $z_W^{(alt)} = \frac{\ln \widehat{RR} - \ln RR_0}{s_g}$, kde $\widehat{\sigma}_g^2 = s_g^2 = \frac{1-\widehat{p}}{\widehat{p}}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$\textit{p-hodnota} = \begin{cases} 2Pr(Z_W^{(alt)} \geq |z_W^{(alt)}| | H_{01}), \ ak & H_{11} : \ln RR \neq \ln RR_0 \\ Pr(Z_W^{(alt)} \geq z_W^{(alt)} | H_{02}), \ ak & H_{12} : \ln RR > \ln RR_0 \\ Pr(Z_W^{(alt)} \leq z_W^{(alt)} | H_{03}), \ ak & H_{13} : \ln RR < \ln RR_0 \end{cases}$$

Definícia 87 (Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $\ln \mathbf{RR}$) *Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $\ln RR$ (dosadí sa smerodajná odchýlka s_g prislúchajúca jednému z testov) pre všetky tri typy hypotéz majú nasledovný tvar:*

H_0	H_1	*hranice (d,h) pre $100(1-\alpha)\%$ empirický IS*
$\ln RR = \ln RR_0$	$\ln RR \neq \ln RR_0$	$\mathcal{CS}_{1-\alpha} = \left\{\ln RR_0 : \ln RR_0 \in \left(\ln \widehat{RR} - u_{\alpha/2} s_g, \widehat{RR} + u_{\alpha/2} s_g\right)\right\}$
$\ln RR \leq \ln RR_0$	$\ln RR > \ln RR_0$	$\mathcal{CS}_{1-\alpha} = \left\{\ln RR_0 : \ln RR_0 \in \left(\ln \widehat{RR} - u_\alpha s_g, \infty\right)\right\}$
$\ln RR \geq \ln RR_0$	$\ln RR < \ln RR_0$	$\mathcal{CS}_{1-\alpha} = \left\{\ln RR_0 : \ln RR_0 \in \left(-\infty, \ln \widehat{RR} + u_\alpha s_g\right)\right\}$

Majme dvojice hypotéz $H_{01} : \mathrm{RR} = \mathrm{RR}_0$ oproti $H_{11} : \mathrm{RR} \neq \mathrm{RR}_0$, $H_{02} : \mathrm{RR} \leq \mathrm{RR}_0$ oproti $H_{12} : \mathrm{RR} > \mathrm{RR}_0$, $H_{03} : \mathrm{RR} \geq \mathrm{RR}_0$ oproti $H_{13} : \mathrm{RR} < \mathrm{RR}_0$, ktoré chceme testovať. Hypotézy H_0 sa často nazývajú **hypotézy homogenity dvoch binomických rozdelení**, ak $\mathrm{RR}_0 = 1$.

Ak H_0 platí, potom

$$Z_W = \frac{\frac{X_1/N_1}{X_2/N_2} - \mathrm{RR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde z δ-metódy pomocou rozptylu $\ln \mathrm{RR}$ (a rovnosti $Var[\mathrm{RR}] \approx \mathrm{RR}^2 Var[\ln \mathrm{RR}]$) dostaneme $S_g^2 \approx \left(\frac{X_1/N_1}{X_2/N_2}\right)^2 \left(\frac{1-X_1/N_1}{X_1} + \frac{1-X_2/N_2}{X_2}\right) \neq 0$. Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **dvojvýberový Z-test o relatívnom riziku**.

Definícia 88 (p-hodnota Z_W testu o RR) *Nech Z_W je **nejaká testovacia štatistika** a $z_W = \frac{\widehat{RR}-RR_0}{s_g}$, kde $\widehat{RR} = \frac{\widehat{p}_1}{\widehat{p}_2}$ a $\widehat{\sigma}_g^2 = s_g^2 = \widehat{RR}^2 \left(\frac{1-\widehat{p}_1}{N_1\widehat{p}_1} + \frac{1-\widehat{p}_2}{N_2\widehat{p}_2}\right)$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W \geq |z_W| | H_{01}), \ ak & H_{11} : RR \neq RR_0 \\ Pr(Z_W \geq z_W | H_{02}), \ ak & H_{12} : RR > RR_0 \\ Pr(Z_W \leq z_W | H_{03}), \ ak & H_{13} : RR < RR_0 \end{cases}$$

Ak za platnosti H_0 berieme pri výpočte do úvahy, že $p_1 = p_2$, potom dostaneme alternatívnu testovaciu štatistiku, ktorá bude mať nasledovný tvar

$$Z_W^{(\mathrm{alt})} = \frac{\frac{X_1/N_1}{X_2/N_2} - \mathrm{RR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde z δ-metódy pomocou rozptylu $\ln \mathrm{RR}$ dostaneme $S_g^2 \approx \left(\frac{X_1/N_1}{X_2/N_2}\right)^2 \times \frac{1-X/N}{X/N}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \neq 0$. $Z_W^{(\mathrm{alt})}$ sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **alternatívny dvojvýberový Z-test o relatívnom riziku**.

Definícia 89 (p-hodnota $Z_W^{(\mathrm{alt})}$ testu o RR) *Nech $Z_W^{(alt)}$ je **nejaká testovacia štatistika** a $z_W^{(alt)} = \frac{\widehat{RR}-RR_0}{s_g}$, kde $\widehat{\sigma}_g^2 = s_g^2 = \widehat{RR}^2 \frac{1-\widehat{p}}{\widehat{p}} \left(\frac{1}{N_1} + \frac{1}{N_2}\right)$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W^{(alt)} \geq |z_W^{(alt)}| | H_{01}), \ ak & H_{11} : RR \neq RR_0 \\ Pr(Z_W^{(alt)} \geq z_W^{(alt)} | H_{02}), \ ak & H_{12} : RR > RR_0 \\ Pr(Z_W^{(alt)} \leq z_W^{(alt)} | H_{03}), \ ak & H_{13} : RR < RR_0 \end{cases}$$

Definícia 90 (Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre RR) *Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre RR (dosadí sa smerodajná odchýlka s_g prislúchajúca jednému z testov) pre všetky tri typy hypotéz majú nasledovný tvar*

H_0	H_1	*hranice* (d,h) *pre* $100(1-\alpha)\%$ *empirický IS*
$RR = RR_0$	$RR \neq RR_0$	$\mathcal{CS}_{1-\alpha} = \left\{RR_0 : RR_0 \in \left(\widehat{RR} - u_{\alpha/2}s_g, \ln \widehat{RR} + u_{\alpha/2}s_g\right)\right\}$
$RR \leq RR_0$	$RR > RR_0$	$\mathcal{CS}_{1-\alpha} = \left\{RR_0 : RR_0 \in \left(\widehat{RR} - u_\alpha s_g, \infty\right)\right\}$
$RR \geq RR_0$	$RR < RR_0$	$\mathcal{CS}_{1-\alpha} = \left\{RR_0 : RR_0 \in \left(0, \widehat{RR} + u_\alpha s_g\right)\right\}$

Príklad 241 (programovanie; Z-testy pre $\ln$ RR a RR) *Naprogramujte v R (a) klasický a alternatívny Z-test o logaritme relatívneho rizika $\ln p_1/p_2$ a (b) klasický a alternatívny Z-test relatívneho rizika p_1/p_2, (c) klasický a alternatívny Waldov 95% empirický DIS relatívneho rizika (spätne transformovaný klasický a alternatívny Waldov 95% empirický DIS logaritmu relatívneho rizika) a (d) klasický a alternatívny Waldov 95% empirický DIS relatívneho rizika. Funkciu nazvite* `test.relat.rizika()`.

Príklad 242 (zuby a tepelný šok; podiel rizík) *Z 50 zubov vystavených tepelnému šoku sa 21 zlomilo. Z kontrolných 50 zubov sa zlomilo len 11. (a) Znižuje tepelný šok mechanickú odolnosť zubov? Porovnajte výsledky klasického a alternatívneho Z-testu (1) logaritmu podielu rizík a (2) podielu rizík. Vypočítajte (b) klasický a alternatívny Waldov 95% empirický DIS relatívneho rizika (spätne transformovaný klasický a alternatívny Waldov 95% empirický DIS logaritmu relatívneho rizika) zlomenia sa zuba a (c) klasický a alternatívny Waldov 95% empirický DIS relatívneho rizika zlomenia sa zuba.*

Riešenie (aj v R)

$N_1 = N_2 = 50, n_1 = 21, n_2 = 11, \widehat{p}_1 = 0.42, \widehat{p}_2 = 0.22, \widehat{p} = 0.32, \widehat{RR} = 1.91.$

Z-testy logaritmu relatívneho rizika a Waldove 95% empirické DIS pre relatívne riziko (spätne transformované DIS pre logaritmus relatívneho rizika):

$z_W \doteq 2.060$, p-hodnota $\doteq 0.020$ a H_0 zamietame na $\alpha = 0.05$.

Waldov 95% empirický DIS pre relatívne riziko RR: $(d, h) \doteq (1.032, 3.532)$

$z_W^{(\text{alt})} \doteq 2.220$, p-hodnota $\doteq 0.013$ a H_0 zamietame na $\alpha = 0.05$.

Alternatívny Waldov 95% empirický DIS pre relatívne riziko RR: $(d, h) \doteq (1.078, 3.381)$

Z-testy relatívneho rizika a Waldove 95% empirické DIS pre relatívne riziko:

$z_W \doteq 1.517$, p-hodnota $\doteq 0.065$ a H_0 nezamietame na $\alpha = 0.05$.

Waldov 95% empirický DIS pre relatívne riziko RR: $(d, h) \doteq (0.735, 3.084)$

$z_W^{(\text{alt})} \doteq 1.633$, p-hodnota $\doteq 0.051$ a H_0 nezamietame na $\alpha = 0.05$.

Alternatívny Waldov 95% empirický DIS pre relatívne riziko RR: $(d, h) \doteq (0.818, 3.000)$

```
test.relat.rizika(21,11,50,50)
##$odhady
#       p     p1     p2      RR  sd(lnRR)  sd.alt(lnRR)  sd(RR)  sd.alt(RR)
# 0.3200 0.4200 0.2200 1.9091    0.3139        0.2915  0.5992      0.5566
##$Ztest
#                      DH     HH  Z-stat  p-hodnota
#Ztest(lnRR)      1.0319 3.5319 2.0600     0.0197
#Ztest.alt(lnRR)  1.0781 3.3806 2.2179     0.0133
#Ztest(RR)        0.7346 3.0836 1.5171     0.0646
#Ztest.alt(RR)    0.8182 3.0000 1.6333     0.0512
```

Pomer šancí $\frac{p_1/(1-p_1)}{p_2/(1-p_2)}$. Majme dvojice hypotéz $H_{01} : \ln \mathrm{OR} = \ln \mathrm{OR}_0$ oproti $H_{11} : \ln \mathrm{OR} \neq \ln \mathrm{OR}_0$, $H_{02} : \ln \mathrm{OR} \leq \ln \mathrm{OR}_0$ oproti $H_{12} : \ln \mathrm{OR} > \ln \mathrm{OR}_0$, $H_{03} : \ln \mathrm{OR} \geq \ln \mathrm{OR}_0$ oproti $H_{13} : \ln \mathrm{OR} < \ln \mathrm{OR}_0$, ktoré chceme testovať. Hypotézy H_0 sa často nazývajú **hypotézy homogenity dvoch binomických rozdelení**, ak $\ln \mathrm{OR}_0 = 0$.

Ak H_0 platí, potom

$$Z_W = \frac{\ln \frac{(X_1/N_1)/(1-X_1/N_1)}{(X_2/N_2))/(1-X_2/N_2)} - \ln \mathrm{OR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde $S_g^2 = \frac{1}{N_1 X_1/N_1(1-X_1/N_1)} + \frac{1}{N_2 X_2/N_2(1-X_2/N_2)} = \frac{1}{N_1} + \frac{1}{N_1 - X_1} + \frac{1}{N_2} + \frac{1}{N_2 - X_2}$. Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **dvojvýberový Z-test o logaritme pomeru šancí**. $Var[\ln \mathrm{OR}]$ vypočítame na základe invariantnosti MLE $\boldsymbol{\theta} = (p_1, p_2)^T$ platí (Woolf, 1955)

$$\mathcal{I}^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{p_1(1-p_1)}{N_1} & 0 \\ 0 & \frac{p_2(1-p_2)}{N_2} \end{pmatrix}$$

a pre $g(\boldsymbol{\theta}) = \ln\frac{p_1}{1-p_1} - \ln\frac{p_2}{1-p_2}$ platí

$$\boldsymbol{\Delta} = \frac{\partial}{\partial\boldsymbol{\theta}} g(\boldsymbol{\theta}) = \left(\frac{\partial}{\partial p_1} g(\boldsymbol{\theta}), \frac{\partial}{\partial p_2} g(\boldsymbol{\theta})\right)^T = \left(\frac{1}{p_1(1-p_1)}, \frac{-1}{p_2(1-p_2)}\right)^T.$$

Potom

$$\sigma_g^2 = Var[\ln \mathrm{OR}] = \boldsymbol{\Delta}^T \mathcal{I}^{-1}(\boldsymbol{\theta})\boldsymbol{\Delta} = \frac{1}{N_1 p_1(1-p_1)} + \frac{1}{N_2 p_2(1-p_2)}.$$

Definícia 91 (p-hodnota Z_W testu o $\ln$ OR) *Nech Z_W je **nejaká testovacia štatistika** a $z_W = \frac{\ln\widehat{OR} - \ln OR_0}{s_g}$, kde $\ln\widehat{OR} = \ln\frac{\widehat{p}_1/(1-\widehat{p}_1)}{\widehat{p}_2/(1-\widehat{p}_2)}$ a $\widehat{\sigma}_g^2 = s_g^2 = \frac{1}{N_1\widehat{p}_1(1-\widehat{p}_1)} + \frac{1}{N_2\widehat{p}_2(1-\widehat{p}_2)}$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W \geq |z_W||H_{01}), \ ak & H_{11}: \ln OR \neq \ln OR_0 \\ Pr(Z_W \geq z_W|H_{02}), \ ak & H_{12}: \ln OR > \ln OR_0 \\ Pr(Z_W \leq z_W|H_{03}), \ ak & H_{13}: \ln OR < \ln OR_0 \end{cases}$$

Ak za platnosti H_0 berieme pri výpočte do úvahy, že $p_1 = p_2$, potom dostaneme alternatívny rozptyl, ktorý bude rovný

$$\sigma_g^2 = Var[\ln \mathrm{OR}] = \frac{1}{p(1-p)}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) = \frac{N}{p(1-p)N_1N_2},$$

Potom

$$Z_W^{(\mathrm{alt})} = \frac{\ln\frac{(X_1/N_1)/(1-X_1/N_1)}{(X_2/N_2))/(1-X_2/N_2)} - \ln \mathrm{OR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde $S_g^2 = \frac{1}{X/N(1-X/N)}\left(\frac{1}{N_1} + \frac{1}{N_2}\right) = \frac{N1}{X/N(1-X/N)N_1N_2} \neq 0$. $Z_W^{(\mathrm{alt})}$ sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **alternatívny dvojvýberový Z-test o logaritme pomeru šancí**.

Definícia 92 (p-hodnota $Z_W^{(\mathrm{alt})}$ testu o $\ln$ OR) *Nech $Z_W^{(alt)}$ je **nejaká testovacia štatistika** a $z_W^{(alt)} = \frac{\ln\widehat{OR} - \ln OR_0}{s_g}$, kde $\widehat{\sigma}_g^2 = s_g^2 = \frac{1}{\widehat{p}(1-\widehat{p})}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$, je jej realizácia (**pozorovaná, vypočítaná testovacia štatistika**), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W^{(alt)} \geq |z_W^{(alt)}||H_{01}), \ ak & H_{11}: \ln OR \neq \ln OR_0 \\ Pr(Z_W^{(alt)} \geq z_W^{(alt)}|H_{02}), \ ak & H_{12}: \ln OR > \ln OR_0 \\ Pr(Z_W^{(alt)} \leq z_W^{(alt)}|H_{03}), \ ak & H_{13}: \ln OR < \ln OR_0 \end{cases}$$

Definícia 93 (Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $\ln$ OR) *Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $\ln OR$ (dosadí sa smerodajná odchýlka s_g prislúchajúca jednému z testov) pre všetky tri typy hypotéz majú nasledovný tvar*

H_0	H_1	*hranice* (d,h) *pre* $100(1-\alpha)\%$ *empirický IS*
$\ln OR = \ln OR_0$	$\ln OR \neq \ln OR_0$	$\mathcal{CS}_{1-\alpha} = \left\{\ln OR_0 : \ln OR_0 \in (\ln\widehat{OR} - u_{\alpha/2}s_g, \ln\widehat{OR} + u_{\alpha/2}s_g)\right\}$
$\ln OR \leq \ln OR_0$	$\ln OR > \ln OR_0$	$\mathcal{CS}_{1-\alpha} = \left\{\ln OR_0 : \ln OR_0 \in \left(\ln\widehat{OR} - u_\alpha s_g, \infty\right)\right\}$
$\ln OR \geq \ln OR_0$	$\ln OR < \ln OR_0$	$\mathcal{CS}_{1-\alpha} = \left\{\ln OR_0 : \ln OR_0 \in \left(-\infty, \ln\widehat{OR} + u_\alpha s_g\right)\right\}$

Majme dvojice hypotéz H_{01} : OR = OR$_0$ oproti H_{11} : OR $\neq$ OR$_0$, H_{02} : OR $\leq$ OR$_0$ oproti H_{12} : OR $>$ OR$_0$, H_{03} : OR $\geq$ OR$_0$ oproti H_{13} : OR $<$ OR$_0$, ktoré chceme testovať. Hypotézy H_0 sa často nazývajú **hypotézy homogenity dvoch binomických rozdelení**, ak OR$_0 = 1$.

Ak H_0 platí, potom

$$Z_W = \frac{\frac{(X_1/N_1)/(1-X_1/N_1)}{(X_2/N_2))/(1-X_2/N_2)} - \mathrm{OR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde z δ-metódy ($Var[\mathrm{OR}] \approx \mathrm{OR}^2 Var[\ln \mathrm{OR}]$) pomocou rozptylu ln OR dostaneme

$$S_g^2 \approx \left(\frac{(X_1/N_1)/(1-X_1/N_1)}{(X_2/N_2))/(1-X_2/N_2)}\right)^2 \left(\frac{1}{N_1} + \frac{1}{N_1 - X_1} + \frac{1}{N_2} + \frac{1}{N_2 - X_2}\right) \neq 0.$$

Z_W sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **dvojvýberový Z-test o pomere šancí**.

Definícia 94 (p-hodnota Z_W testu o OR) *Nech Z_W je* ***nejaká testovacia štatistika*** *a $z_W = \frac{\widehat{OR} - OR_0}{s_g}$, kde $\widehat{OR} = \frac{\widehat{p}_1/(1-\widehat{p}_1)}{\widehat{p}_2/(1-\widehat{p}_2)}$) a $\widehat{\sigma}_g^2 = s_g^2 = \widehat{OR}^2 \left(\frac{1}{N_1\widehat{p}_1(1-\widehat{p}_1)} + \frac{1}{N_2\widehat{p}_2(1-\widehat{p}_2)}\right)$, je jej realizácia (****pozorovaná, vypočítaná testovacia štatistika****), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W \geq |z_W| | H_{01}), \ ak & H_{11} : OR \neq OR_0 \\ Pr(Z_W \geq z_W | H_{02}), \ ak & H_{12} : OR > OR_0 \\ Pr(Z_W \leq z_W | H_{03}), \ ak & H_{13} : OR < OR_0 \end{cases}$$

Ak za platnosti H_0 berieme pri výpočte do úvahy, že $p_1 = p_2$, potom dostaneme alternatívnu testovaciu štatistiku, ktorá bude mať nasledovný tvar

$$Z_W^{(\mathrm{alt})} = \frac{\frac{(X_1/N_1)/(1-X_1/N_1)}{(X_2/N_2))/(1-X_2/N_2)} - \mathrm{OR}_0}{S_g} \overset{\mathcal{D}}{\sim} N(0,1),$$

kde z δ-metódy pomocou rozptylu ln OR dostaneme

$$S_g^2 \approx \left(\frac{(X_1/N_1)/(1-X_1/N_1)}{(X_2/N_2))/(1-X_2/N_2)}\right)^2 \times \frac{N_1}{X/N(1-X/N)N_1N_2} \neq 0.$$

$Z_W^{(\mathrm{alt})}$ sa nazýva **Waldova testovacia štatistika** (alebo Z-**štatistika**) a test **alternatívny dvojvýberový Z-test o pomere šancí**.

Definícia 95 (p-hodnota $Z_W^{(\mathrm{alt})}$ testu o OR) *Nech $Z_W^{(alt)}$ je* ***nejaká testovacia štatistika*** *a $z_W^{(alt)} = \frac{\widehat{OR} - OR_0}{s_g}$, kde $\widehat{\sigma}_g^2 = s_g^2 = \widehat{OR}^2 \frac{1}{\widehat{p}(1-\widehat{p})}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$, je jej realizácia (****pozorovaná, vypočítaná testovacia štatistika****), potom p-hodnotu počítame nasledovne:*

$$p\text{-}hodnota = \begin{cases} 2Pr(Z_W^{(alt)} \geq |z_W^{(alt)}| | H_{01}), \ ak & H_{11} : OR \neq OR_0 \\ Pr(Z_W^{(alt)} \geq z_W^{(alt)} | H_{02}), \ ak & H_{12} : OR > OR_0 \\ Pr(Z_W^{(alt)} \leq z_W^{(alt)} | H_{03}), \ ak & H_{13} : OR < OR_0 \end{cases}$$

Definícia 96 (Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre OR) *Klasické a alternatívne Waldove $100 \times (1-\alpha)\%$ empirické IS pre $\ln OR$ (dosadí sa smerodajná odchýlka s_g prislúchajúca jednému z testov) pre všetky tri typy hypotéz majú nasledovný tvar*

H_0	H_1	*hranice (d, h) pre $100(1-\alpha)\%$ empirický IS*
$OR = OR_0$	$OR \neq OR_0$	$\mathcal{CS}_{1-\alpha} = \left\{OR_0 : OR_0 \in \left(\widehat{OR} - u_{\alpha/2}s_g, \widehat{OR} + u_{\alpha/2}s_g\right)\right\}$
$OR \leq OR_0$	$OR > OR_0$	$\mathcal{CS}_{1-\alpha} = \left\{OR_0 : OR_0 \in \left(\widehat{OR} - u_{\alpha}s_g, \infty\right)\right\}$
$OR \geq OR_0$	$OR < OR_0$	$\mathcal{CS}_{1-\alpha} = \left\{OR_0 : OR_0 \in \left(0, \widehat{OR} + u_{\alpha}s_g\right)\right\}$

Príklad 243 (programovanie; Z-testy pre OR a $\ln$ OR) *Naprogramujte v R (a) klasický a alternatívny Z-test o logaritme pomeru šancí $\ln \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$ a (b) klasický a alternatívny Z-test pomeru šancí $p_1/(1-p_1)p_2/(1-p_2)$, (c) klasický a alternatívny Waldov 95% empirický DIS pomeru šancí (spätne transformovaný klasický a alternatívny Waldov 95% empirický DIS logaritmu pomeru šancí) a (d) klasický a alternatívny Waldov 95% empirický DIS pomeru šancí. Funkciu nazvite* test.pomeru.sanci().

Príklad 244 (zuby a tepelný šok; pomer šancí) *Z 50 zubov vystavených tepelnému šoku sa 21 zlomilo. Z kontrolných 50 zubov sa zlomilo len 11. (a) Znižuje tepelný šok mechanickú odolnosť zubov? Porovnajte výsledky klasického a alternatívneho Z-testu (1) logaritmu pomeru šancí a (2) pomeru šancí. Vypočítajte (b) klasický a alternatívny Waldov 95% empirický DIS pomeru šancí (spätne transformovaný klasický a alternatívny Waldov 95% empirický DIS logaritmu pomeru šancí) zlomenia sa zuba a (c) klasický a alternatívny Waldov 95% empirický DIS pomeru šancí zlomenia sa zuba.*

Riešenie (aj v R)
$N_1 = N_2 = 50, n_1 = 21, n_2 = 11, \widehat{p}_1 = 0.42, \widehat{p}_2 = 0.22, \widehat{p} = 0.32, \widehat{OR} = 1.91.$
Z-testy logaritmu pomeru šancí a Waldove 95% empirické DIS pre pomer šancí (spätne transformované DIS pre logaritmus pomeru šancí):
$z_W \doteq 2.116$, p-hodnota $\doteq 0.017$ a H_0 zamietame na $\alpha = 0.05$.
Waldov 95% empirický DIS pre pomer šancí OR: $(d, h) \doteq (1.072, 6.150)$.
$z_W^{(\text{alt})} \doteq 2.120$, p-hodnota $\doteq 0.014$ a H_0 zamietame na $\alpha = 0.05$.
Alternatívny Waldov 95% empirický DIS pre pomer šancí OR: $(d, h) \doteq (1.108, 5.949)$.

Z-testy pomeru šancí a Waldove 95% empirické DIS pre pomer šancí:
$z_W \doteq 1.370$, p-hodnota $\doteq 0.085$ a H_0 nezamietame na $\alpha = 0.05$.
Waldov 95% empirický DIS pre pomer šancí OR: $(d, h) \doteq (0.325, 4.810)$.
$z_W^{(\text{alt})} \doteq 1.424$, p-hodnota $\doteq 0.077$ a H_0 nezamietame na $\alpha = 0.05$.
Alternatívny Waldov 95% empirický DIS pre pomer šancí OR: $(d, h) \doteq (0.410, 4.724)$.

```
test.pomeru.sanci(21,11,50,50)
##$odhady
#        p      p1      p2      OR   sd(lnOR)  sd.alt(lnOR)  sd(OR)   sd.alt(OR)
# 0.3200  0.4200  0.2200  2.5674    0.4457          0.4287  1.1443       1.1008
##$Ztest
#                     DH      HH  Z-stat  p-hodnota
#Ztest(lnOR)      1.0718  6.1500  2.1155     0.0172
#Ztest.alt(lnOR)  1.1080  5.9490  2.1992     0.0139
#Ztest(OR)        0.3246  4.8102  1.3697     0.0854
#Ztest.alt(OR)    0.4099  4.7249  1.4239     0.0772
```

Test pomerom vierohodnosti pre $p_1 - p_2$.

Majme H_{01} oproti H_{11}. Nech $X_j \sim Bin(N_j, p_j), j = 1, 2$, sú nezávislé náhodné premenné. Funkcia vierohodnosti binomického rozdelenia pre p_1 a p_2 bude mať tvar

$$L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = p_1^{n_1}(1-p_1)^{N_1-n_1} p_2^{n_2}(1-p_2)^{N_2-n_2},$$

kde $\boldsymbol{\theta} = (p_1, p_2)^T$. MLE $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{p}_1, \widehat{p}_2)^T$, kde $\widehat{p}_1 = n_1/N_1$ a $\widehat{p}_2 = n_2/N_2$, t.j. $\Theta_1 = \{\boldsymbol{\theta} : p_1 \neq p_2 \neq p\}$. Za platnosti H_{01} je $\boldsymbol{\theta}_0 = (\widehat{p}, \widehat{p})^T$, kde $\widehat{p} = \frac{N_1\widehat{p}_1 + N_2\widehat{p}_2}{N_1 + N_2}$ a $\Theta_0 = \{\boldsymbol{\theta} : p_1 = p_2 = p\}$. Potom -1

krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$\begin{aligned}
-\ln(\lambda(\mathbf{x}_1,\mathbf{x}_2)) &= l\left(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1,\mathbf{x}_2\right) - l\left(\boldsymbol{\theta}_0|\mathbf{x}_1,\mathbf{x}_2\right) \\
&= \sum_{j=1}^{2} n_j \ln \widehat{p}_j + \sum_{j=1}^{2}(N_j-n_j)\ln(1-\widehat{p}_j) - \sum_{j=1}^{2} n_j \ln \widehat{p} - \sum_{j=1}^{2}(N_j-n_j)\ln(1-\widehat{p}) \\
&= \sum_{j=1}^{2} n_j \ln \frac{\widehat{p}_j}{\widehat{p}} + \sum_{j=1}^{2}(N_j-n_j)\ln\frac{1-\widehat{p}_j}{1-\widehat{p}}.
\end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{X}_1,\mathbf{X}_2\right)\right) \overset{\mathcal{D}}{\sim} \chi_1^2$. Preznačme p_1 na p_{11}, $1-p_1$ na p_{12}, p_2 na p_{21}, $1-p_2$ na p_{22}. Ekvivalentne preznačíme n_1 na n_{11}, N_1-n_1 na n_{12}, n_2 na n_{21}, N_2-n_2 na n_{22}. Potom $\boldsymbol{\theta} = \mathbf{p} = (p_{11},p_{12},p_{21},p_{22})^T$, $\widehat{\boldsymbol{\theta}} = \widehat{\mathbf{p}} = (\widehat{p}_{11},\widehat{p}_{12},\widehat{p}_{21},\widehat{p}_{22})^T$, kde $\widehat{p}_{11} = \frac{n_{11}}{N_1}$, $\widehat{p}_{12} = \frac{n_{12}}{N_1}$, $\widehat{p}_{21} = \frac{n_{21}}{N_2}$, $\widehat{p}_{22} = \frac{n_{22}}{N_2}$ a $\boldsymbol{\theta}_0 = \mathbf{p}_0 = (p_{0,11},p_{0,12},p_{0,21},p_{0,22})^T = (\widehat{p},1-\widehat{p},\widehat{p},1-\widehat{p})^T$, kde $p_{j1}+p_{j2}=1$, $n_{j1}+n_{j2}=N_j$ a $j=1,2$. Potom dostaneme

$$u_{\mathrm{LR}} = 2\sum_{j=1}^{2}\sum_{i=1}^{2} n_{ji}\ln\frac{\widehat{p}_{ji}}{p_{0,ji}} = 2\sum_{j,i}\text{pozorované}_{ji}\times\ln\frac{\text{pozorované}_{ji}}{\text{očakávané}_{ji}}.$$

Vierohodnostný $100\times(1-\alpha)\%$ **empirický DIS pre** p_1-p_2 bude mať tvar

$$\mathcal{CS}_{1-\alpha} = \left\{p_0 : U_{\mathrm{LR}} < \chi_1^2(\alpha)\right\}.$$

Na výpočet tohoto DIS by sme museli reparametrizovať funkciu vierohodnosti na parameter záujmu (p_1-p_2) a rušivý parameter. Alternatívne by bolo možné vypočítať u_{LR} pre rôzne kombinácie p_1 a p_2 (na vhodne zvolených intervaloch) a pre tieto p_j, $j=1,2$, vypočítať rozdiely p_1-p_2. Potom $\mathcal{CS}_{1-\alpha} = \left\{p_0 : U_{\mathrm{LR}}(p_1-p_2) < \chi_1^2(\alpha)\right\}$.

Parametrom záujmu $g(\boldsymbol{\theta}) = \theta$ môže byť namiesto samotných pravdepodobností p_1 a p_2 ich rozdiel $\theta = p_1-p_2$ (často nazývaný aj rozdiel rizík), relatívne riziko $\theta = p_1/p_2$, pomer šancí $\theta = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$ alebo logaritmus pomeru šancí $\theta = \ln\frac{p_1/(1-p_1)}{p_2/(1-p_2)}$. Každý parameter záujmu má svoje výhody a nevýhody vo vzťahu k interpretácii alebo štatistickým vlastnostiam. Pri malých rozsahoch výberov je funkcia vierohodnosti **logaritmu pomeru šancí** regulárnejšia než funkcia vierohodnosti ostatných parametrov záujmu, preto vyjadríme funkciu vierohodnosti pre logaritmus pomeru šancí θ s rušivým parametrom $\eta = \ln\frac{p_2}{1-p_2}$. Potom $p_2 = \frac{e^\eta}{1+e^\eta}$ a $p_1 = \frac{e^{\theta+\eta}}{1+e^{\theta+\eta}}$ a funkcia vierohodnosti bude rovná

$$\begin{aligned}
L(\boldsymbol{\theta}|\mathbf{x}_1,\mathbf{x}_2) &= \left(\frac{p_1}{1-p_1}\right)^{n_1}(1-p_1)^{N_1}\left(\frac{p_2}{1-p_2}\right)^{n_2}(1-p_2)^{N_2} \\
&= \left(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\right)^{n_1}\left(\frac{p_2}{1-p_2}\right)^{n_1+n_2}(1-p_1)^{N_1}(1-p_2)^{N_2} \\
&= e^{\theta n_1}e^{\eta(n_1+n_2)}(1+e^{\theta+\eta})^{-N_1}(1+e^{\eta})^{-N_2}.
\end{aligned}$$

kde $\boldsymbol{\theta} = (\theta,\eta)^T$. Maximálne vierohodný odhad logaritmu pomeru šancí $\widehat{\theta} = \ln\frac{n_1/(N_1-n_1)}{n_2/(N_2-n_2)}$. Aby sme dostali profilovú funkciu vierohodnosti pre θ, mohli by sme vypočítať maximálne vierohodný odhad η v každej fixovanej hodnote θ avšak to nie je možné. Preto musíme profilovú funkciu vierohodnosti počítať numericky ako

$$L(\theta|\mathbf{x}_1,\mathbf{x}_2) = \max_{\forall\eta} L(\boldsymbol{\theta}|\mathbf{x}_1,\mathbf{x}_2).$$

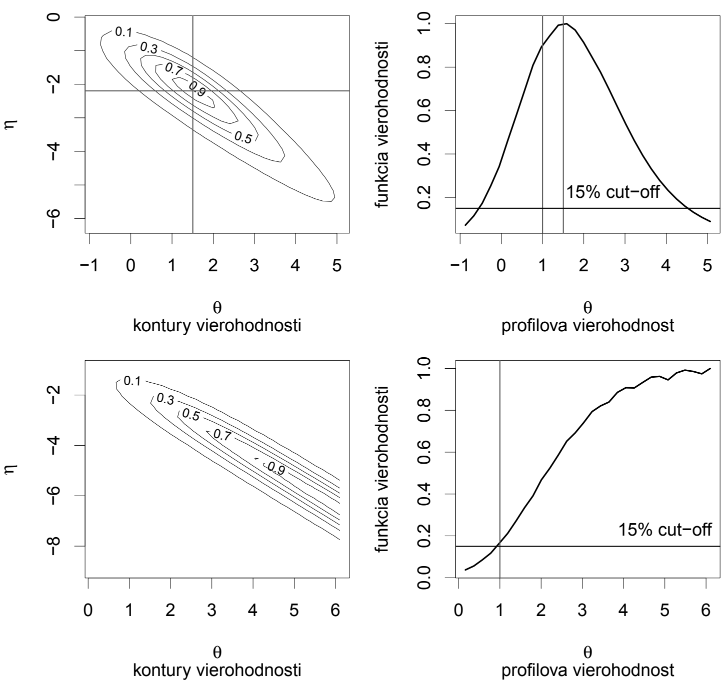
Príklad 245 (maximálne vierohodné odhady; nádor prsníka) *Majme početnosti subjektov X_1, ktoré majú rozšírené metastázy nádoru prsníka, kde $X_1 \sim Bin(N_1,p_1)$ a početnosti subjektov X_2, ktoré* *majú lokalizované metastázy nádoru prsníka, kde $X_2 \sim Bin(N_2, p_2)$; pozri Pawitan (2001). (a) Aplikujte funkciu vierohodnosti $L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2)$, kde $\boldsymbol{\theta} = (\theta, \eta)^T$, logaritmus pomeru šancí $\theta = \ln \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$ a rušivý parameter $\eta = \ln \frac{p_2}{1-p_2}$ na dáta v tabuľke 7.2 a vypočítajte $\widehat{\theta}$. (b) Nakreslite funkciu vierohodnosti ako aj profilovú funkciu vierohodnosti a DIS. Zopakujte pre $n_1 = 6$ a $n_2 = 0$. (c) Vypočítajte vierohodnostný 95% DIS pre θ pomocou metodiky 15% cut-off štandardizovanej profilovej funkcie vierohodnosti. DIS dokreslite do jedného obrázka k profilovej funkcii vierohosnoti v jej 15% cut-off.*

Tabuľka 7.2: Početnosti subjektov s rozšírenými a lokalizovanými metastázami

metastázy	rozšírené	lokalizované	spolu
áno	5	1	6
nie	10	9	19
spolu	15	10	25

Riešenie (pozri obrázok 7.8)

Keď $n_1 = 5$ a $n_2 = 1$, logaritmus pomeru šancí $\widehat{\theta} = \ln \frac{5/10}{1/9} = 1.5$.
Vierohodnostný 95% DIS pre θ je rovný $(-0.544, 4.524)$ a vierohodnostný 95% DIS pre pomer šancí je rovný $(0.580, 92.226)$.
Keď $n_1 = 6$ a $n_2 = 0$, $\widehat{\theta} = \infty$ a vierohodnosť nie je regulárna.
Vierohodnostný 95% JIS pre θ je rovný $(0.921, \infty)$.



Obr. 7.8: Funkcia vierohodnosti a profilová funkcia vierohodnosti binomického rozdelenia pre θ; $n_1 = 5, n_2 = 1$ (prvý riadok), $n_1 = 6, n_2 = 0$ (druhý riadok)

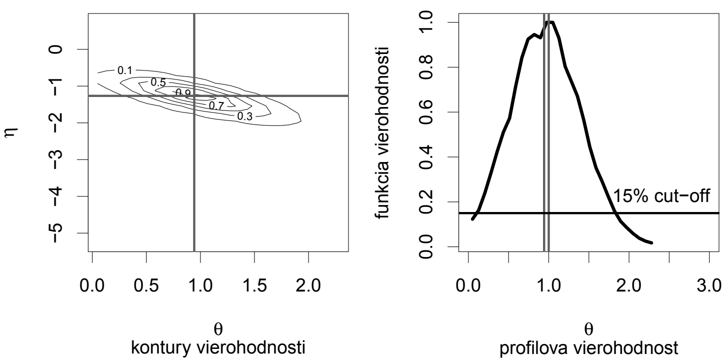
Príklad 246 (zuby a tepelný šok; pomer šancí) *Z 50 zubov vystavených tepelnému šoku sa 21 zlomilo. Z kontrolných 50 zubov sa zlomilo len 11. (a) Znižuje tepelný šok mechanickú odolnosť zubov?*

Na testovanie použite test pomerom vierohodnosti. Vypočítajte (b) vierohodnostný 95% empirický DIS pre pomer šancí.

Riešenie (aj v R)

```
n1 <- 21
n2 <- 11
N1 <- 50
N2 <- 50
N <- N1+N2
p1.hat <- n1/N1
p2.hat <- n2/N2
p.hat <- c(p1.hat,1-p1.hat,p2.hat,1-p2.hat)
n <- c(n1,N1-n1,n2,N2-n2)
p.exp <- c((n1+n2)/N,(N1-n1+N2-n2)/N,(n1+n2)/N,(N1-n1+N2-n2)/N)
uLR <- 2*(sum(n*log(p.hat/p.exp))) # 4.653895
1-pchisq(uLR,df=1) # 0.03098316
```

Testovacia štatistika $u_{\mathrm{LR}} = 4.654$ a prislúchajúca p-hodnota $= 0.031$, t.j. $H_0 : p_1 - p_2 = 0$ zamietame na $\alpha = 0.05$. Na výpočet DIS použijeme funkciu vierohodnosti pre logaritmus pomeru šancí (pozri obrázok 7.9) a spätne ho transformujeme na DIS pre pomer šancí. DIS pre logaritmus pomeru šancí je rovný $(d, h) = (0.103, 1.834)$. DIS pre pomer šancí je rovný $(d, h) = (1.107, 6.262)$.



Obr. 7.9: Funkcia vierohodnosti a profilová funkcia vierohodnosti binomického rozdelenia pre logaritmus pomeru šancí zlomenia sa zuba

Príklad 247 (test o pomere šancí) *Majme dáta* two-samples-probabilities-sexratio.txt, *premennú počet starších súrodencov* o.sib.N *a pohlavie* sex. *Predpokladáme, že početnosť chlapcov, ak nemajú staršieho súrodenca, $X_{m,0} \sim Bin(N_{m,0}, p_{m,0})$; početnosť chlapcov, ak ho majú, $X_{m,1} \sim Bin(N_{m,1}, p_{m,1})$; početnosť dievčat, ak nemajú staršieho súrodenca, $X_{f,0} \sim Bin(N_{f,0}, p_{f,0})$; početnosť dievčat, ak ho majú, $X_{f,1} \sim Bin(N_{f,1}, p_{f,1})$. (a) Otestujte hypotézu o pomere šancí narodenia chlapca (ak nemá staršieho súrodenca voči situácii, že ho má) na hladine významnosti $\alpha = 0.05$. (b) Vypočítajte $100 \times (1-\alpha)$% empirický DIS pre tento pomer šancí, kde koeficient spoľahlivosti $1 - \alpha = 0.95$. Použite Waldovu testovaciu štatistiku Z_W pre logaritmus pomeru šancí a jej zodpovedajúci DIS.*

Funkcia vierohodnosti **Poissonovho rozdelenia** má nasledovný tvar

$$L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = e^{-(\lambda_1+\lambda_2)}\lambda_1^{n_1}\lambda_2^{n_2},$$

kde $\boldsymbol{\theta} = (\lambda_1, \lambda_2)^T$. Zaujíma nás len **relatívne riziko** $\theta = \lambda_2/\lambda_1$ a nie λ_1 a λ_2 samostatne, pričom λ_1 budeme považovať za rušivý parameter. Model reparametrizujeme z $\boldsymbol{\theta} = (\lambda_1, \lambda_2)^T$ na $\boldsymbol{\theta} = (\theta, \lambda_1)^T$. Potom

$$L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = e^{-\lambda_1(1+\theta)}\lambda_1^{n_1+n_2}\theta^{n_2},$$

Tabuľka 7.3: Odhad relatívneho rizika úmrtia a 95% IS relatívneho rizika úmrtia celkovo a pre každú vekovú skupinu

kód vekovej skupiny	veková skupina	minulý rok	tento rok	$\widehat{\theta}_j$	(d,h) pre 95% DIS
vek 1	< 20	20	35	1.75	$(1.024, 3.073)$
vek 2	$\langle 20, 40)$	40	54	1.35	$(0.901, 2.039)$
vek 3	$\langle 40, 50)$	75	65	0.87	$(0.620, 1.207)$
vek 4	≥ 50	10	2	0.20	$(0.028, 0.755)$

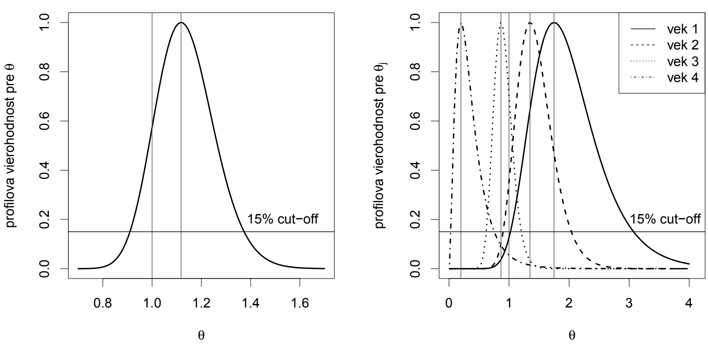
kde $\boldsymbol{\theta} = (\theta, \lambda_1)^T$. Pre každé fixované θ je maximálne vierohodný odhad λ_1 je rovný $\widehat{\lambda}_1(\theta) = (n_1+n_2)/(1+\theta)$. Potom profilová funkcia vierohodnosti

$$L(\theta|\mathbf{x}_1, \mathbf{x}_2) = e^{-\widehat{\lambda}_1(\theta)(1+\theta)} \left(\widehat{\lambda}_1(\theta)\right)^{n_1+n_2} \theta^{n_2} = c\left(\frac{1}{1+\theta}\right)^{n_1}\left(1-\frac{1}{1+\theta}\right)^{n_2},$$

kde maximalizácia $L(\theta|\mathbf{x}_1, \mathbf{x}_2)$ nezávisí na c. Potom $\widehat{\theta} = n_2/n_1$.

Príklad 248 (maximálne vierohodné odhady; úmrtia) *Nech početnosti úmrtí na cestách v minulom a tomto roku, ozn.* X_1 *a* X_2, *majú nezávislé Poissonove rozdelenia s parametrami* λ_1 *a* λ_2,*.t.j.* $X_1 \sim Poiss(\lambda_1)$ *a* $X_2 \sim Poiss(\lambda_2)$. *(a) Aplikujte funkciu vierohodnosti* $L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2)$, *kde* $\boldsymbol{\theta} = (\theta, \lambda_1)^T$, *relatívne riziko* $\theta = \lambda_2/\lambda_1$ *a rušivý parameter* λ_1 *na dáta v tabuľke 7.3 (*$n_1 = 170$ *a* $n_2 = 190$*; pozri Pawitan (2001)) a vypočítajte* $\widehat{\theta}$. *(b) Nakreslite funkciu vierohodnosti ako aj profilové funkcie vierohodnosti pre každú vekovú skupinu. (c) Vypočítajte vierohodnostný* 95% *DIS pre* θ *a* θ_j *pomocou metodiky* 15% *cut-off štandardizovanej funkcie vierohodnosti.*

Riešenie (pozri obrázok 7.10)
Relatívne riziko $\widehat{\theta} = 1.12$, $\widehat{\theta}_j, j = 1,2,3,4$ pozri v tabuľke.
Vierohodnostný 95% DIS pre θ je rovný $(0.910, 1.374)$.



Obr. 7.10: Profilová funkcia vierohodnosti Poissonovho rozdelenia pre θ a $\theta_j, j = 1,2,3,4$

Aproximácia binomického rozdelenia Poissonovým rozdelením. Nech $X_j \sim Bin(N_j, p_j), j = 1,2$, kde p_j sú veľmi malé čísla a N_j sú veľké čísla. Parametrom záujmu je **relatívne riziko** $\theta = p_1/p_2$. Preto môžeme predpokladať, že $X_1 \sim Poiss(\lambda_1), \lambda_1 = N_1p_1$ a $X_2 \sim Poiss(\lambda_2), \lambda_2 = N_2p_2$. Tiež môžeme situáciu zjednodušiť použitím $N_1 \approx N_2$. Potom funkcia vierohodnosti bude mať tvar

$$L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2) = e^{-(N_1p_1+N_2p_2)}(N_1p_1)^{n_1}(N_2p_2)^{n_2} = ce^{-p_2(N_1\theta+N_2)}\theta^{n_1}p_2^{n_1+n_2}$$

kde $\boldsymbol{\theta} = (\theta, p_2)^T$, $c = N_1^{n_1}N_2^{n_2}$ (maximalizácia $L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2)$ nezávisí na c) a profilova funkcia vierohodnosti

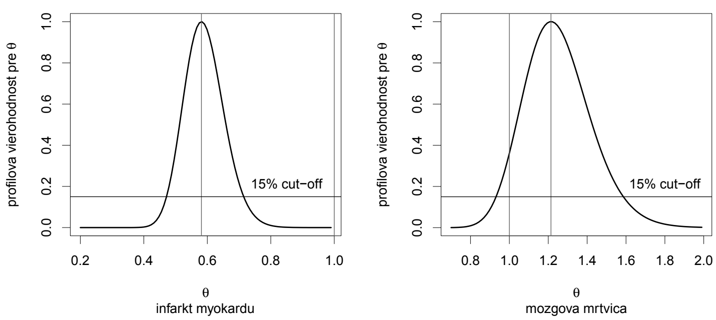
$$L(\theta|\mathbf{x}_1, \mathbf{x}_2) = \left(\frac{N_1\theta}{N_1\theta+N_2}\right)^{n_1}\left(1-\frac{N_1\theta}{N_1\theta+N_2}\right)^{n_2}.$$

Príklad 249 (maximálne vierohodné odhady; nové liek vs. placebo) *Predpokladajme, že početnosť subjektov s infarktom myokardu (IM)* X_1 *v skupine A (nový liek) má binomické rozdelenie s parametrami* N_1 *a* p_1*, t.j.* $X_1 \sim Bin(N_1, p_1)$*, a počet subjektov s IM v skupine B (placebo) má tiež binomické rozdelenie s parametrami* N_2 *a* p_2*, t.j.* $X_2 \sim Bin(N_2, p_2)$*. Pozorovali sme* $n_1 = 139$ *z celkového počtu* $N_1 = 11037$ *subjektov a* $n_2 = 239$ *z celkového počtu* $N_2 = 11034$ *subjektov. Kedže* N_1 *a* N_2 *sú vysoké čísla a pravdepodobnosti IM malé čísla, môžeme predpokladať, že* $X_1 \sim Poiss(\lambda_1), \lambda_1 = N_1 p_1$ *a* $X_2 \sim Poiss(\lambda_2), \lambda_2 = N_2 p_2$*. Zjednodušme situáciu použitím* $N_1 \approx N_2$*. (a) Aplikujte funkciu vierohodnosti* $L(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{x}_2)$*, kde* $\boldsymbol{\theta} = (\theta, p_2)^T$ *a relatívne riziko* $\theta = p_1/p_2$ *na dáta v tabuľke 7.4 (pozri Pawitan (2001)) – pre (1) IM ako aj (2) pre mozgovú mŕtvicu (MM). (b) Nakreslite profilové funkcie vierohodnosti pre každú skupinu. (c) Vypočítajte vierohodnostný* 95% *DIS pre (1)* θ_{IM} *a (2)* θ_{MM} *pomocou metodiky* 15% *cut-off štandardizovanej funkcie vierohodnosti.*

Tabuľka 7.4: Početnosti subjektov s infarktom myokardu a mozgovou mŕtvicou v skupine A a B

skupina	IM	MM	spolu
skupina A	139	119	11037
skupina B	239	98	11034

Riešenie (pozri tabuľku 7.4 a obrázok 7.11)
Relatívne riziko $\widehat{\theta}_{IM} = \frac{139/11037}{239/11034} \doteq 0.581$. Vierohodnostný 95% DIS pre θ_{IM} je rovný $(0.471, 0.715)$.
Relatívne riziko $\widehat{\theta}_{MM} = \frac{119/11037}{98/11034} \doteq 1.214$. Vierohodnostný 95% DIS pre θ_{MM} je rovný $(0.931, 1.586)$.



Obr. 7.11: Profilová funkcia vierohodnosti pre θ

8 Testovanie hypotéz o viacerých parametroch

Aj keď nám v praxi v mnohých situáciách stačí testovanie hypotéz o jednom parametri alebo dvoch parametroch (čo je náplňou predchádzajúcich dvoch kapitol), bežne sa v antropológii stretávame aj s požiadavkou otestovať rozdiely J parametrov súčasne (kde $J > 2$). Tieto parametre charakterizujú rozdelenie nejakej premennej v J populáciách, ktoré sú identifikovateľné pomocou nejakej ordinálnej alebo nominálnej premennej, ktorú v tomto prípade nazývame **faktor**. **Kategórie** (**úrovne**) daného faktora

- existujú buď prirodzene – napr. faktor pohlavie, mesto, štát, archeologická kultúra (lokalita), zamestnanie, faktor počet predchádzajúcich detí biologickej matky hodnoteného jedinca, faktor populácia odrážajúca podtypy jednej choroby odlišnej etiológie, faktor poradie narodenia (prvorodený, druhorodený, narodený vo vyššom poradí), faktor počet narodených detí (*nullipara*, *primipara*, *secundipara* atď.), vekové intervaly, faktor vzdelanie alebo

- sú dané metodicky na základe použitej štandardnej kategorizačnej metódy – napr. faktor sexuálna orientácia (sedembodová Likertova škála, Kinseyho škála – od sexuálnej orientácie výlučne na opačné pohlavie (0) až po sexuálnu orientáciu výlučne na rovnaké pohlavie (6)), faktor zdravotný stav novorodenca (kategórie orientačného hodnotenia zdravotného stavu novorodencov na základe Apgar skóre).

V takýchto situáciách máme dve hlavné úlohy

1. zistiť, či vôbec existuje nejaký rozdiel aspoň medzi dvoma populáciami, t.j. či sa stredné hodnoty aspoň dvoch populácií líšia (ide o **viacvýberový test rovnosti stredných hodnôt**) a

2. zistiť, medzi ktorými populáciami existuje štatisticky významný rozdiel a ktoré sa navzájom nelíšia (ide o tzv. **mnohonásobné porovnávania** alebo **post-hoc testy**).

Ako zdanlivo správna alternatíva k testovaniu viacerých parametrov súčasne (viac kategórií, viac populácií, viac premenných a pod.) sa ponúka viacnásobné použitie príslušného počtu testov dvoch parametrov (napr. tri dvojvýberové Studentove t-testy v prípade troch úrovní faktora a testovania rozdielov stredných hodnôt). Pri viacnásobnom porovnávaní dvoch parametrov je však pravdepodobnosť vzniku chyby prvého druhu vyššia (často výrazne vyššia v závislosti od počtu testovaných nulových hypotéz) a sila testu nižšia ako pri porovnávaní dvoch parametrov (pozri kapitolu 4 Testovanie hypotéz). V prípade viacerých výberov alebo viacerých premenných by sme preto mali vždy použiť test pre viac parametrov s nadväzujúcimi post-hoc testami a nie sekvenciu dvojvýberových testov.

Pri testovaní hypotéz o viacerých parametroch však treba pamätať na splnenie predpokladov, pri ktorých možno tieto testy použiť. Vo všeobecnosti jednotlivé náhodné výbery nemusia mať rovnaké rozsahy a môžu sa líšiť ich rozptyly alebo ich rozdelenie (tieto podmienky však možno ošetriť pri plánovaní štúdie/experimentu). Jednotlivé úrovne môžu a nemusia tvoriť stupne v rámci nejakého gradientu. Okrem toho môžu byť rozdiely medzi niektorými úrovňami (populáciami) výrazne väčšie než medzi inými. K riešeniu musíme pristupovať rôznym spôsobom, na základe toho, čo o vzorkách a populáciách, z ktorých vzorky pochádzajú, vopred vieme. Napr. pri porovnaní šiestich miest, z ktorých tri sú z jedného a tri z druhého kraja, môžeme vopred predpokladať existenciu nejakých skupín, v ktorých si môžu byť výbery bližšie, zatiaľ čo pri iných porovnaniach to vopred predpokladať nie je možné a skupiny sa objavia až ako výsledok štatistickej analýzy. Rozdiely takýchto skupín možno testovať použitím vhodne zvolených **kontrastov** pred analýzou ale i v priebehu analýzy.

V antropológii a biomedicínsky zameraných disciplínach sa analýza rozptylu (ANOVA) bežne používa pri testovaní rozdielov stredných hodnôt u viacerých výberov alebo pri testovaní vplyvu iných faktorov na tieto rozdiely. Menej používané však už sú viacvýberové testy rozdielov v rozptyloch, korelačných koeficientoch a pravdepodobnostiach, hoci sú rovnako potrebné a užitočné. Dôvodom môže byť snáď aj to, že väčšinou nie sú implementované v najčastejšie používaných „klikacích“ štatistických balíkoch. Aj z tohto dôvodu sa domnievame, že ich zaradenie do tejto kapitoly spolu s riešenými a neriešenými príkladmi môže prispieť k ich väčšiemu rozšíreniu v relevantných oblastiach. Nejde pritom len o ich prínos pri plánovaní štúdií a experimentov, ale aj o prínos pri štatistických analýzach samotných, kedy znalosť týchto testov môže pomôcť bližšie pochopiť komplexné vzťahy v analyzovaných dátach a tým zjednodušiť interpretáciu štatistických výsledkov.

Z praktických dôvodov (použitie testov založených na predpokladoch zhody v rozptyloch) nás najčastejšie zaujíma zhoda rozptylov nejakej premennej v dvoch populáciách. Ako už bolo uvedené v predchádzajúcich kapitolách (kapitola 6 a 7), štúdium rozdielov v rozptyloch medzi porovnávanými populáciami môže byť podstatné i meritórne z hľadiska štúdia *evolučných procesov* (stabilizujúca selekcia) alebo *ontogenetických procesov* (vplyv modifikujúcich faktorov prostredia). Predstavuje preto dôležitý krok na ceste k rozlíšeniu rozdielov medzi porovnávanými populáciami. Variabilita v dátach predstavuje súhrn *variability pôvodnej* (napr. biologickej povahy), vznikajúcej v dôsledku procesov, ktoré nás zaujímajú, a *variability vnesenej druhotne*, t.j. buď prirodzenou selekciou časti populácie, z ktorej vzorku vyberáme (napr. časť pohrebiska), náhodou (šum, nepresnosť merania) alebo úplne umelo metódou vzorkovania či technikou merania (napr. rozdiely v meradlách). Pokiaľ sa vzorky v rozptyloch nelíšia, neznamená to, že ide o identické výsledky spolupôsobenia rovnakých zdrojov variability pôvodných a druhotných. Pokiaľ sa ale vzorky v rozptyloch významne líšia, je potrebné hľadať príčinu tejto odlišnosti. Na rozdiel od porovnania rozptylov dvoch populácií môže byť našou úlohou pri porovnávaní viacerých populácií (podobne ako pri porovnávaní viacerých stredných hodnôt) 1. zistiť, či sa aspoň jedna z nich líši v rozptyle od ostatných a 2. testovať, medzi ktorými populáciami existujú štatisticky významné rozdiely v rozptyle.

Použitie viacvýberového testu zhody korelačných koeficientov je obdobou hodnotenia zhody viacerých stredných hodnôt alebo rozptylov pomocou viacvýberových testov. K dispozícii máme dáta z niekoľkých populácií popisujúce vzťah dvoch spojitých premenných a pýtame sa, či 1. je súvislosť (asociácia) oboch premenných v sledovaných populáciách rovnaká alebo nie. Pokiaľ sa aspoň jedna populácia v tejto asociácii líši, zaujíma nás 2., ktorá to je, resp. ktoré populácie sa od seba navzájom líšia. Typickým príkladom môžu byť medzipopulačné rozdiely vo vzťahu medzi šírkou a výškou nosa. V proporciách nosa existujú rozdiely medzi veľkými, geograficky vymedzenými populačnými skupinami (tradične a nesprávne hľadané rozdiely medzi „rasami“), ktoré však vychádzajú najmä z prílišného zjednodušenia situácie, sprevádzaného zanedbávaním funkčných adaptácií ako zdroja variability vo veľkosti a proporciách ľudského tela. Charakter nosa je v skutočnosti výsledkom evolučných adaptácií na teplotu a vlhkosť klímy – dlhé a relatívne úzke nosy sú výhodnejšie v aridných (suchých) a chladných podmienkach kvôli lepšiemu zvlhčeniu a ohriatiu vdychovaného vzduchu (tzv. Thomsonovo biogeografické pravidlo). Táto adaptácia pritom môže zanechať určitý obraz aj na tvare podkladových lebečných štruktúr nosa. V rôznych ľudských populáciách tak spolu môže šírka a výška nosa súvisieť rôzne. Pritom si treba uvedomiť, že našou úlohou *nie je* zistenie rozdielu medzi populáciami v strednej hodnote pomeru šírky a výšky nosa (t.j. či sa populácie líšia alebo nelíšia v pomere oboch rozmerov, alebo či rozdiely medzi populáciami v tomto pomere závisia na rozdieloch v stredných hodnotách niektorého z oboch rozmerov), *ale zistenie rozdielu v asociácii* jedného a druhého rozmeru v sledovaných populáciách. Na pamäti treba mať, že nejde o zisťovanie rozdielov v miere veľkostnej izometrie/alometrie tvaru nosa, t.j. či sa oba korelované rozmery zväčšujú proporčne rovnako (izometria) alebo jeden z nich napr. v polovičnej či dvojnásobnej miere. Test zhody viacerých korelačných koeficientov zisťuje, či je v nulovej hypotéze vo všetkých populáciách súvislosť oboch rozmerov rovnaká. Nulovú hypotézu zamietame, ak existuje aspoň jedna dvojica populácií, kde je rozdiel korelačných koeficientov nenulový. Zamietnutie nulovej hypotézy nezávisí od tesnosti vzťahu dvoch veličín v dvoch populáciách, ale od samotného rozdielu korelačných koeficientov. V kombinácii s testem shody stredných hodnôt týchto rozmerov a ich rozptylov to dovoľuje uvažovať o funkčnom význame morfológie nosa (veľkosť, tvar, variabilita a vzájomný vzťah rozmerov) v širšom kontexte.

Napriek tomu, že na celý rad riešených úloh biologickej antropológie stačí jednovýberové alebo dvojvýberové porovnanie pravdepodobností, mnohokrát nie je možné situáciu takto zjednodušiť bez straty podstatných informácií. V takom prípade je potrebné testovať rozdiely medzi populáciami pomocou testov porovnávajúcich viac než dve pravdepodobnosti. V antropológii môže byť typickým prípadom porovnanie frekvencií mužov a žien, klasifikovaných v rámci nejakého faktoru s tromi alebo viacerými úrovňami, napr. päť vekových intervalov, štyri stupne vzdelania, osem krajských miest, rôzne druhy zamestnania, ap. Pri testovaní nulových hypotéz s viacerými pravdepodobnosťami je nevyhnutné presne špecifikovať, aké prípadné rozdiely nás konkrétne zaujímajú, napr. sú podstatné rozdiely medzi pohlaviami v každom krajskom meste alebo rozdiely medzi rôznymi krajskými mestami v rámci každého pohlavia? Testovanie všetkých možných rozdielov totiž zvyšuje počet stupňov voľnosti a zmenšuje schopnosť testu zamietnuť nulovú hypotézu pri rovnakých frekvenciách. V prípade kostrových nálezov je typickou úlohou porovnanie frekvencií viacúrovňového epigenetického znaku medzi dvoma alebo viacerými populáciami s cieľom štúdia ich genetickej príbuznosti. Môže sa však stať, že aj v relatívne veľkých kostrových súboroch (čo už samo osebe nie je úplne bežné) budú frekvencie v jednotlivých úrovniach (kategóriách) daného znaku veľmi malé, t.j. rádovo v jednotkách prípadov (napr. Kaur a kol., 2012). V malých kostrových súboroch či v populáciách s extrémne nízkym podielom daného epigenetického znaku sa dokonca znak nemusí v niektorej z jeho úrovní vyskytovať vôbec. V prípade neprítomnosti niektorej úrovne znaku, t.j. nulovej frekvencie, je potom vhodné a niekedy dokonca nevyhnutné použitie špeciálne navrhnutých testov, ktorých problematika však už presahuje rámec tejto knihy (bližšie Rücker a kol., 2009).

Príkladom viacvýberového testu zhody stredných hodnôt môže byť test rozdielov stredných hodnôt výšky hornej časti tváre v piatich rôznych populáciách (dáta: anova-means-skull.txt). Príkladom viacvýberového testu rozptylov môže byť test rozdielov v rozptyloch dĺžky kľúčnej kosti v štyroch populáciách kľúčnych kostí (dáta: more-samplex-variances-clavicle.txt). Testom zhody viacerých korelačných koeficientov môžeme napr. zistiť, či sa päť populácií líši v asociácii výšky a šírky nosa (dáta: more-samples-correlations-skull.txt). Zhoda viacerých pravdepodobností nás môže zaujímať napr. pri testovaní rozdielov vo výskyte troch stupňov útvarov negatívneho reliéfu na vnútornej strane lonovej spony u žien z troch populácií (dáta: more-samples-probabilities-pubis.txt).

V tejto kapitole sa teda budeme venovať testovaniu hypotéz o stredných hodnách $\mu_1, \mu_2, \dots, \mu_J$, za predpokladov normality, t.j. $X_{ji} \sim N(\mu_j, \sigma_j^2)$, kde $j = 1, 2, \dots, J$ a $i = 1, 2, \dots, n_j$, o rozptyloch $\sigma_1^2, \sigma_2^2, \dots, \sigma_J^2$, za predpokladov normality, t.j. $X_{ji} \sim N(\mu_j, \sigma_j^2)$, o korelačných koeficientoch $\rho_1, \rho_2, \dots, \rho_J$, za predpokladov normality, t.j. $(X_{ji}, Y_{ji})^T \sim N_2\left(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j\right)$, a o pravdepodobnostiach $p_1, p_2, \dots, p_J$, za prepokladov $X \sim Mult_J\left(N, \mathbf{p}\right)$, kde $\mathbf{p} = (p_1, p_2, \dots, p_J)^T$. Keďže X_{ji} v tomto prípade zapisujeme do štatistického modelu, kde sa na jeho ľavej strane nachádzajú závislé premenné (ozn. Y), preznačíme ich na Y_{ji}. Ozn. X_{ij} rezervujeme pre nezávislé premenné, napr. fixné efekty (zaradenie do skupín a pod.).

8.1 Asymptotické testy o stredných hodnotách

Nech $Y_{ji} \sim N(\mu_j, \sigma_j^2)$, kde $j = 1, 2, \dots, J$ a $i = 1, 2, \dots, n_j$, sú nezávislé náhodné premenné. Budeme rozlišovať dve situácie

1. $Y_{ji} \sim N(\mu_j, \sigma_j^2)$, kde $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_J^2 = \sigma_e^2$ (**homogenita rozptylov**), σ_j^2 sú neznáme a
2. $Y_{ji} \sim N(\mu_j, \sigma_j^2)$, kde existuje aspoň jedna dvojica $i \neq j$ taká, že $\sigma_i^2 \neq \sigma_j^2$ (**nehomogenita rozptylov**), σ_j^2 sú neznáme.

8.1.1 Jednofaktorový ANOVA model s fixnými efektami pri rovnakých rozptyloch

Nech $Y_{ji} \sim N(\mu_j, \sigma_j^2)$, kde $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_J^2 = \sigma_e^2$ a zároveň σ_j^2 sú neznáme. Majme **jednofaktorový model analýzy rozptylu (ANOVA) s fixnými (pevnými) efektami**, ozn. $\mathcal{F}_{H_1}$, definovaný ako

$$Y_{ji} = \mu_j + \varepsilon_{ji} = \mu + \alpha_j + \varepsilon_{ji},$$

kde $\mu = \sum_{j=1}^{J} \mu_j / J$, $\mu_j = \mu + \alpha_j$, $\sum_{j=1}^{J} \alpha_j = 0$, μ je **celková (spoločná) úroveň spoločná všetkým populáciám** (alebo **celková stredná hodnota**), α_j je j**-ta úroveň faktora** A (j**-ty efekt faktora** A) a znamená odchýlku strednej hodnoty j-tej populácie od μ. Pre chyby ε_{ji} platí $\varepsilon_{ji} \sim N\left(0, \sigma_e^2\right)$. Model $Y_{ji} \sim N(\mu + \alpha_j, \sigma_e^2)$ sa nazýva aj **model podmieneného normálneho rozdelenia**.

Majme dvojicu hypotéz H_0: $\mu_1 = \mu_2 = ... = \mu_J = \mu$ oproti H_1: existuje aspoň jedno $i < j$ také, že $\mu_i \neq \mu_j$. Ak H_0 platí, potom $\alpha_1 = \alpha_2 = \ldots = \alpha_J = 0$, dostaneme submodel, ozn. $\mathcal{F}_{H_0}$, definovaný ako

$$Y_{ji} = \mu + \varepsilon_{ji}.$$

Ak H_0 platí, potom

$$F_W = \frac{\frac{SS_A}{J-1}}{\frac{SS_e}{n-J}} \overset{\mathcal{D}}{\sim} F_{J-1, n-J},$$

kde $df_A = J-1$, $df_e = (n-1)-(J-1) = n-J$ sú stupne voľnosti, $n = \sum_{j=1}^{J} n_j$ je celkový rozsah, n_j sú rozsahy jednotlivých výberov, SS_A je **výberový súčet štvorcov rozdielov medzi súbormi** a je definovaný ako

$$SS_A = \sum_{j=1}^{J} n_j \left(\overline{Y}_{j.} - \overline{Y}_{..}\right)^2 = \sum_{j=1}^{J} \frac{Y_{j.}^2}{n_j} - \frac{1}{n} Y_{..}^2,$$

$\widehat{\mu} = \overline{Y}_{..} = Y_{..}/n$ je maximálne vierohodný odhad μ, $Y_{..} = \sum_{j=1}^{J} \sum_{i=1}^{n_j} Y_{ji}$, $\widehat{\mu}_j = \overline{Y}_{j.} = Y_{j.}/n$ je maximálne vierohodný odhad μ_j, $Y_{j.} = \sum_{i=1}^{n_j} Y_{ji}$, SS_e je **výberový súčet štvorcov rozdielov vnútri súborov** a je definovaný ako

$$SS_e = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(Y_{ji} - \overline{Y}_{j.}\right)^2 = \sum_{j=1}^{J} \sum_{i=1}^{n_j} Y_{ji}^2 - \sum_{j=1}^{J} \frac{Y_{j.}^2}{n_j}.$$

Súčet SS_A a SS_e sa rovná SS_T, čo je **celkový výberový súčet štvorcov rozdielov** a je definovaný ako

$$SS_T = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(Y_{ji} - \overline{Y}_{..}\right)^2 = \sum_{j=1}^{J} \sum_{i=1}^{n_j} Y_{ji}^2 - \frac{1}{n} Y_{..}^2,$$

Rovnosti $SS_T = SS_A + SS_e$ hovoríme aj *rozklad celkovej sumy štvorcov*. Pre stupne voľnosti potom platí $df_T = df_A + df_e$, kde $df_T = n-1$. Sumy štvorcov sa najčastejšie zapisujú do **ANOVA tabuľky**:

zdroj variability	suma štvorcov	df	priemerné štvorce
medzi súbormi	$SS_{A,\text{obs}}$	$J-1$	$MS_{A,\text{obs}} = SS_{A,\text{obs}} / (J-1)$
vnútri súborov	$SS_{e,\text{obs}}$	$n-J$	$MS_{e,\text{obs}} = SS_{e,\text{obs}} / (n-J) = \widehat{\sigma}_e^2 = s^2$
celkovo	$SS_{T,\text{obs}}$	$n-1$	

F_W sa nazýva **Fisherova testovacia štatistika** (alebo **ANOVA F-štatistika**) a test **viacvýberový F-test o rovnosti stredných hodnôt** $\mu_1, \mu_2, \ldots, \mu_J$ (alebo **ANOVA F-test**). Realizáciou F_W je F_{obs} a p-hodnota = $\Pr(F_W \geq F_{\text{obs}} | H_0)$.

Úlohu môžeme interpretovať tak, že stredná hodnota μ_j náhodnej veličiny Y_{ji} závisí na **faktore** A, čo je premenná v nominálnej škále. Jednotlivým **úrovniam** (hladinám) tejto premennej zodpovedajú **fixné efekty** $\alpha_j = \mu_j - \mu$. Úrovne premennej volí experimentátor, sú teda nenáhodné, dopredu dané (fixné). Potom chápeme α_j ako neznáme parametre, ktorých maximálne vierohodné odhady definujeme ako $\widehat{\alpha}_j = \overline{y}_{j.} - \overline{y}_{..}$. Samotné rozhodovanie o H_0 bude založené na porovnaní priemerných súm štvorcov $SS_{A,\text{obs}}/df_A$ a $SS_{e,\text{obs}}/df_e$. Väčšie rozdiely $\overline{y}_{j.}$ a $\overline{y}_{..}$ (v absolútnej hodnote) sa prejavia vo väčšej hodnote štatistiky $SS_{A,\text{obs}}$. Štatistika $SS_{e,\text{obs}}$ zasa umožňuje odhadnúť rozptyl σ_e^2 a súčasne dáva mieru pre hodnotenie veľkosti variability medzi súbormi.

Rozdelenia používané v ANOVA modeli môžeme zhrnúť do nasledovných bodov:

1. Nech $Z_{ji} \sim N(0,1)$, kde $j = 1,2,\ldots,J; i = 1,2,\ldots,n_j$, potom $\sum_{j=1}^{J} \sum_{i=1}^{n_j} Z_{ji}^2 \sim \chi^2_{df}$, kde $df = n = \sum_{j=1}^{J} n_j$;

2. Nech $Z_{ji} \sim N\left(0, \sigma_e^2\right)$, potom $\sum_{j=1}^{J} \sum_{i=1}^{n_j} Z_{ji}^2 \sim \sigma_e^2 \chi_n^2$ a $\sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(Z_{ji} - \overline{Z}_{..}\right)^2 \sim \sigma_e^2 \chi_{n-1}^2$;

3. Nech $Z_{ji} \sim N\left(\alpha_j, \sigma_e^2\right)$, potom $\sum_{j=1}^{J} \sum_{i=1}^{n_j} Z_{ji}^2 \sim \sigma_e^2 \chi^2_{df,\lambda^2}$, $\sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(Z_{ji} - \overline{Z}_{..}\right)^2 \sim \sigma_e^2 \chi^2_{df-1,\lambda^2}$, kde $\lambda^2 = \sum_{j=1}^{J} \alpha_j^2 / \sigma_e^2$ a zároveň $\frac{SS_T}{n-1} \sim \frac{\sigma_e^2}{n-1} \chi^2_{n-1}$, $SS_e \sim \sigma_e^2 \chi^2_{df_e}$;

4. Nech $\overline{Z}_{j.} \sim N\left(\alpha_j, \sigma_e^2/n_j\right)$, potom $\frac{SS_A}{J-1} \sim \frac{\sigma_e^2}{J-1} \chi^2_{J-1,\lambda^2}$, kde $\lambda^2 = \sum_{j=1}^{J} \frac{\alpha_j^2}{\sigma_e^2/n_j}$.

Majme **vyvážený model** (vyvážené triedenie), t.j. nech $n_1 = n_2 = \ldots = n_J = K$, kde K je počet opakovaní v podtriede. Ak v jednofaktorovom ANOVA modeli platí nulová hypotéza H_0, potom má testovacia štatistika F_W centrálne F rozdelenie so stupňami voľnosti $df_A = J-1$ a $df_e = n-J = KJ - J = J(K-1)$. Teda

$$F_W = \frac{MS_A}{MS_e} = \frac{SS_A/df_A}{SS_e/df_e} \approx \frac{\chi^2_{df_A}/df_A}{\chi^2_{df_e}/df_e} \overset{\mathcal{D}}{\approx} F_{df_A, df_e}$$

a ide o podiel dvoch centrovaných χ^2 (čítame „chí-kvadrátov") podelených príslušnými stupňami voľnosti. Nech $Z_{ji} = Y_{ji} - \mu$. Potom $Z_{ji} \sim N\left(0, \sigma_e^2\right)$, $\overline{Z}_{j\cdot} \sim N\left(0, \sigma_e^2/K\right)$ a $SS_A \sim \sigma_e^2 \chi^2_{df_A}$.

Ale ak H_0 neplatí, potom $Z_{ji} \sim N\left(\alpha_j, \sigma_e^2\right)$ a $\overline{Z}_j \sim N\left(\alpha_j, \sigma_e^2/K\right)$, kde aspoň jedno $\alpha_j \neq 0$ a testovacia štatistika má necentrálne F rozdelenie

$$F_{W,\lambda} \approx \frac{\chi^2_{df_A,\lambda^2}/df_A}{\chi^2_{df_e}/df_e} \overset{\mathcal{D}}{\sim} F_{df_A,df_e,\lambda^2},$$

kde $\chi^2_{df_A,\lambda^2}$ je necentrálne chi-kvadrát rozdelenie s parametrom necentrality

$$\lambda^2 = \frac{\sum_{j=1}^{J} \alpha_j^2}{\sigma_e^2/K} = \frac{\sum_{j=1}^{J} \left(\mu_j - \mu\right)^2}{\sigma_e^2/K} = K\left(J-1\right)\frac{\sigma_A^2}{\sigma_e^2},$$

kde $\sigma_A^2 = \frac{1}{J-1}\sum_{j=1}^{J}\left(\mu_j - \mu\right)^2$.

Sila testu. Sila ANOVA F-testu $1-\beta$ je definovaná ako

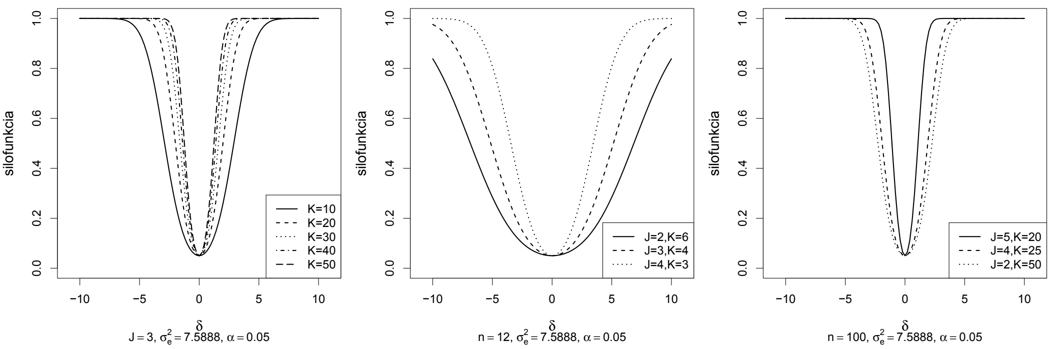
$$1 - \beta = \Pr\left(F_{df_A,df_e,\lambda^2} \geq F_{df_A,df_e}\left(\alpha\right)\right),$$

kde $F_{df_A,df_e}\left(\alpha\right)$ je kritická hodnota centrálneho F-rozdelenia a F_{df_A,df_e,λ^2} je náhodná premenná z necentrálneho F-rozdelenia (pozri obrázok 8.1). Potom **minimálny rozsah** K bude iteračným riešením rovnice

$$1 - \beta = \Pr\left(F_{J-1,J(K-1),\lambda^2} \geq F_{J-1,J(K-1)}\left(\alpha\right)\right),$$

ktorú môžeme napísať aj nasledovne (Rasch a kol., 2011)

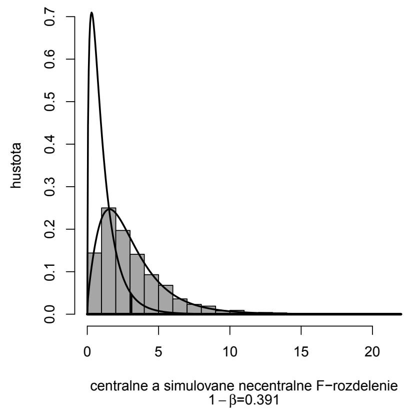
$$F_{J-1,J(K-1),\lambda^2}(1-\beta) = F_{J-1,J(K-1)}\left(\alpha\right).$$



Obr. 8.1: Silofunkcie ANOVA F-testu pri rôznych J a K – vľavo $J = 3$, uprostred $JK = 12$ a vpravo $JK = 100$

Príklad 250 (sila ANOVA F-testu a minimálny rozsah) *Majme štyri populácie, ktorých stredné hodnoty sú $\mu_1 = 390$, $\mu_2 = 405$, $\mu_3 = 415$ a $\mu_4 = 410$. Predpokladajme, že $Y_{ji} \sim N(\mu_j, \sigma_e^2)$.*
(a) Vypočítajte silu $1-\beta$ ANOVA F-testu rovnosti stredných hodnôt za predpokladu, že $K = 6$, $\sigma_e^2 = 20^2$ a $\alpha = 0.05$.
(b) Vypočítajte silu $1-\beta$ ANOVA F-testu rovnosti stredných hodnôt za predpokladu, že $K = 6$, $\sigma_e^2 = 10^2$ a $\alpha = 0.05$.
(c) Použite R na simuláciu hustoty necentrálneho F_{df_A,df_e,λ^2} testovacích štatistík $F_{W,\lambda}^{(m)}$ (necentrálne F rozdelenie s df_A a df_e stupňami voľnosti a parametrom necentrality λ^2), kde $\alpha = 0.05$, $K = 6$, $\sigma_e^2 = 20^2$, $\lambda^2 = \sum_{j=1}^{J}\left(\mu_j - \widehat{\mu}\right)^2/(\sigma_e^2/K)$, $m = 1, 2, \ldots, M$, pri $M = 1000$ opakovaniach.
(d) Vypočítajte minimálny rozsah K, ak $\sigma_e^2 = 20^2$ a $\alpha = 0.05$. K bude iteračným riešením rovnice $F_{J-1,J(K-1),\lambda^2}(\beta) = F_{J-1,J(K-1)}\left(1-\alpha\right)$.

Riešenie
(a) $\lambda^2 = \frac{6(390-405)^2+6(405-405)^2+6(415-405)^2+6(410-405)^2}{20^2} = 5.25,$
$1-\beta(\lambda^2) = \Pr\left(F_{df_A,df_e,\lambda^2} \geq F_{df_A,df_e}(\alpha)\right) = \Pr\left(F_{3,20,5.25} \geq F_{3,20}(0.05)\right) = 0.386.$
(b) $\lambda^2 = \frac{6(405-405)^2+6(405-405)^2+6(415-405)^2+6(410-405)^2}{10^2} = 21,$
$1-\beta(\lambda^2) = \Pr\left(F_{df_A,df_e,\lambda^2} \geq F_{df_A,df_e}(\alpha)\right) = \Pr\left(F_{3,20,21} \geq F_{3,20}(0.05)\right) = 0.95.$
(c) simulácia silofunkcie (pozri obrázok 8.2)



Obr. 8.2: Hustota centrálneho a necentrálneho F-rozdelenia superponovaná histogramom simulácií pre $Y_j \sim N(\mu_j, \sigma_e^2)$

(d) minimálny rozsah $K = 14$ a $n = KJ = 56$.

Alternatívnym prístupom je zistenie sily špecifikáciou **najmenšieho detegovateľného rozdielu**, ktorý chceme mať medzi dvoma najvzdialenejšími strednými hodnotami. Môžeme písať, že najmenšia vzdialenosť medzi najvzdialenejšími strednými hodnotami bude $\delta = c\sigma$, kde $c > 0$. Dá sa ukázať, že $\sum_{j=1}^J \alpha_j^2$ je minimálna, ak budú ľubovoľné dva efekty $\alpha_i = -c\sigma/2$ a $\alpha_j = c\sigma/2$ a ostatných $J-2$ efektov bude rovných nule. Preto

$$\sum_{j=1}^J \alpha_j^2 = \left(-\frac{c\sigma}{2}\right)^2 + \left(\frac{c\sigma}{2}\right)^2 + 0 + \underbrace{...}_{J-2} + 0 = 2\frac{c^2\sigma^2}{4}.$$

Potom

$$\lambda^2 = \frac{\sum_{j=1}^J \alpha_j^2}{\sigma_e^2/K} = \frac{\delta^2}{2\sigma_e^2/K}.$$

Maticový zápis modelu $\mathcal{F}_{H_1}$ a $\mathcal{F}_{H_0}$.
Modely $\mathcal{F}_{H_1}$ a $\mathcal{F}_{H_0}$ sú lineárnymi regresnými modelmi a môžeme ich všeobecne zapísať v tvare $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, kde $\mathbf{Y}$ n-rozmerný náhodný vektor, $\mathbf{X}$ je **matica plánu** s rozmermi $n \times (J+1)$ a $\boldsymbol{\varepsilon}$ je n-rozmerný **vektor chýb**. Potom model $\mathcal{F}_{H_1}$ bude mať tvar

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_J \end{pmatrix} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \begin{pmatrix} \mathbf{1}_1 & \mathbf{1}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{1}_2 & \mathbf{0} & \mathbf{1}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{1}_J & \mathbf{0} & \mathbf{0} & \dots & \mathbf{1}_J \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_J \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_J \end{pmatrix},$$

kde $\mathbf{Y}_j = (Y_{j1}, Y_{j2}, \dots, Y_{jn_j})^T$ je n_j-rozmerný vektor, $\mathbf{1}_j$ je n_j-rozmerný vektor jednotiek a $\boldsymbol{\varepsilon}_j$ je n_j-rozmerný vektor chýb. Potom $\mathbf{Y}_j \sim N_{n_j}(\mathbf{X}\boldsymbol{\beta}, \sigma_e^2\mathbf{I}_{n_j \times n_j})$, vektor chýb $\boldsymbol{\varepsilon}_j \sim N_{n_j}(\mathbf{0}, \sigma_e^2\mathbf{I}_{n_j \times n_j})$, vektor parametrov $\widehat{\boldsymbol{\beta}} \sim N_{J+1}(\boldsymbol{\beta}, \sigma_e^2(\mathbf{X}^T\mathbf{X})^{-1})$, kde maximálne vierohodný odhad $\widehat{\boldsymbol{\beta}}$ vypočítame pomocou **metódy najmenších štvorcov**, t.j. $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$.

Model $\mathcal{F}_{H_0}$ bude mať tvar

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_J \end{pmatrix} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \begin{pmatrix} \mathbf{1}_1 \\ \mathbf{1}_2 \\ \vdots \\ \mathbf{1}_J \end{pmatrix} \mu + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_J \end{pmatrix}.$$

Test pomerom vierohodnosti o rovnosti stredných hodnôt.

Nech $Y_j \sim N(\mu_j, \sigma^2)$, kde $j = 1, 2, \ldots, J$ a σ^2 je neznáma. Logaritmus funkcie vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_J) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \mu_j)^2,$$

kde $n = \sum_{j=1}^{J} n_j$. MLE $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_J, \widetilde{\sigma}^2)^T$, kde

$$\widehat{\mu}_j = \overline{y}_j, \widetilde{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j)^2,$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \mu_i \neq \mu_j; i < j, i = 1, 2, \ldots, J-1; j = 2, 3, \ldots, J\}$ pre aspoň jedno i a j. Za platnosti H_0 je $\widehat{\boldsymbol{\theta}}_0 = (\widehat{\mu}, \widehat{\mu}, \ldots, \widehat{\mu}, \widetilde{\sigma}_0^2)^T$ platí

$$\overline{x} = \widehat{\mu} = \frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j} y_{ji}, \widetilde{\sigma}_0^2 = \frac{1}{n}\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \widehat{\mu})^2,$$

Tiež platí $\Theta_0 = \{\boldsymbol{\theta} : \mu_1 = \mu_2 = \ldots = \mu_J = \mu\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$-\ln(\lambda(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_J)) = \frac{n}{2}\ln\left(\frac{\widetilde{\sigma}_0^2}{\widetilde{\sigma}^2}\right)$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_J\right)\right) \overset{\mathcal{D}}{\sim} \chi^2_{J-1}$, kde H_0 bude zamietnutá pre veľké hodnoty podielu $\frac{\widetilde{\sigma}_0^2}{\widetilde{\sigma}^2}$.

Dá sa ukázať, že $-\ln(\lambda(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_J))$ je rastúcou funkciou F_{obs}. Úpravou podielu $\widetilde{\sigma}_0^2/\widetilde{\sigma}^2$ dostaneme

$$\begin{aligned}
-\ln(\lambda(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_J)) &= \frac{n}{2}\ln\left(\frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j + \overline{y}_j - \widehat{\mu})^2}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j)^2}\right) \\
&= \frac{n}{2}\ln\left(\frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j)^2 + \sum_{j=1}^{J} n_j(\overline{y}_j - \widehat{\mu})^2}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j)^2}\right) \\
&= \frac{n}{2}\ln\left(1 + \frac{J-1}{n-J}\frac{\sum_{j=1}^{J} n_j(\overline{y}_j - \widehat{\mu})^2/(J-1)}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}(y_{ji} - \overline{y}_j)^2/(n-J)}\right) \\
&= \frac{n}{2}\ln\left(1 + \frac{J-1}{n-J}F_{\mathrm{obs}}\right).
\end{aligned}$$

Potom môžeme U_{LR} prepísať

$$u_{\mathrm{LR}} = n\ln\left(1 + \frac{J-1}{n-J}F_{\mathrm{obs}}\right).$$

ANOVA model v R (funkcia **aov()**)
Argumenty (vstupy) funkcie aov():

1. ANOVA model formula v podobe y∼x;
2. dátová tabuľka data;
3. nastavenie výstupu v podobe tabuľky s rozmermi $n \times 3$ obsahujúcej odhady $\widehat{\mu}$, $\widehat{\alpha}_j$ a reziduály (chyby) ε_{ji}, projections=FALSE (prednastavené);

Výstupy funkcie aov():

1. tabuľky s rozmermi $n \times 3$ obsahujúca odhady $\widehat{\mu}$, $\widehat{\alpha}_j$ a reziduály ε_{ji}, projections;
2. odhady $\widehat{y}_{ji}$ fitted.values;
3. reziduály ε_{ji} residuals;

Výstupy funkcie summary(aov()):

1. ANOVA tabuľka, kde
   - stupne voľnosti df_A a df_e summary(MODEL)[[1]][,1],
   - sumy štvorcov $SS_{A,\text{obs}}$ a $SS_{e,\text{obs}}$ summary(MODEL)[[1]][,2],
   - priemerné štvorce $MS_{A,\text{obs}}$ a $MS_{e,\text{obs}}$ summary(MODEL)[[1]][,3];
2. realizáciu testovacej štatistiky F_{obs} summary(MODEL)[[1]][1,4];
3. p-hodnota summary(aov())[[1]][1,5].

Funkcia aov() používa na výpočty funkciu lineárny regresný model lm(). Pri priamom použití fukcie lm() dostaneme ANOVA tabuľku ako anova(lm()). Odmocninu z rozptylu $\widehat{\sigma}_e^2$ dostaneme pomocou summary(lm())$sig. Alternatívne je možné použiť funkciu oneway.test(), ktorej vstupom je ANOVA model formula v podobe y∼x, dátová tabuľka data a nastavenie rovnosti rozptylov var.equal=TRUE. Výstupom sú realizácia testovacej štatistiky F_{obs}, stupne voľnosti df_A a df_e a p-hodnota.

Príklad 251 (test pomerom vierohodnosti o rovnosti stredných hodnôt) *Naprogramujte v R test pomerom vierohodnosti o rovnosti stredných hodnôt za predpokladu homogenity rozptylov. Aplikujte na dáta anova-head.txt a premennú bizyg.W u mužov a porovnajte výsledok testu pomerom vierohodnosti s výsledkom ANOVA F-testu. Odôvodnite rozdiely vo výsledkoch.*

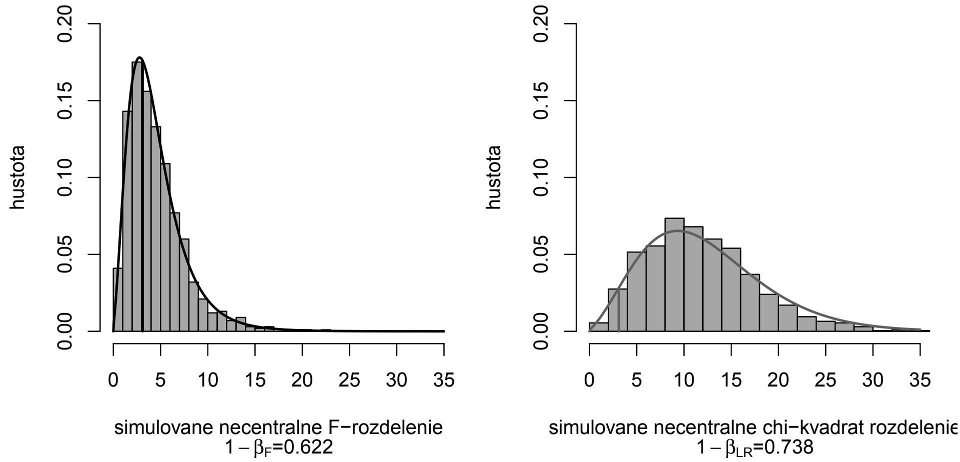
Riešenie v R

```
"testy.o.stred.hodnotach" <- function(x,y){
MODEL <- aov(y~x)
F.stat <- summary(MODEL)[[1]][1,4]
p.hodnota.F <- summary(MODEL)[[1]][1,5]
VYSL.F.test <- c(F.stat,p.hodnota.F)
names(VYSL.F.test) <- c("test-stat","p-hodnota")
J <- nlevels(x)
n <- length(x)
u.LT <- n*log(1+(J-1)/(n-J)*F.stat)
p.hodnota.LR <- 1-pchisq(u.LT,df=J-1)
VYSL.LR.test <- c(u.LT,p.hodnota.LR)
names(VYSL.LR.test) <- c("test-stat","p-hodnota")
VYSL <- list(VYSL.F.test,VYSL.LR.test)
return(VYSL)
}
testy.o.stred.hodnotach(sex.sexor[sex=="m"],bizyg.W[sex=="m"])
#[[1]]
# test-stat  p-hodnota
#3.32743009 0.07303222
#[[2]]
#test-stat p-hodnota
#9.8476790 0.0199063
```

Príklad 252 (testy o rovnosti stredných hodnôt; simulačná štúdia) *Majme 4 populácie, ktorých stredné hodnoty sú $\mu_1 = 390$, $\mu_2 = 405$, $\mu_3 = 415$ a $\mu_4 = 410$. Predpokladajme, že $Y_{ji} \sim N(\mu_j, \sigma_e^2)$, kde $\sigma_e^2 = 15^2$. Použite R na simuláciu hustoty (1) necentrálneho F_{df_A,df_e,λ^2} testovacích štatistík $F_{W,\lambda}^{(m)}$ (necentrálne F rozdelenie s df_A a df_e stupňami voľnosti a parametrom necentrality*

λ^2) a (2) necentrálneho $\chi^2_{df_A,\lambda^2}$ testovacích štatistík $u^{(m)}_{LR,\lambda}$ (necentrálne χ^2 rozdelenie s df_A stupňami voľnosti a parametrom necentrality λ^2), kde $\alpha = 0.05$, $K = 6$, $\lambda^2 = \sum_{j=1}^{J} (\mu_j - \widehat{\mu})^2 / (\sigma_e^2/K)$, $m = 1, 2, \ldots, M$, pri $M = 1000$ opakovaniach. Vypočítajte silu v prípade (1) a (2). Okomentujte výsledok.

Riešenie v R (riešenie pozri na obrázku 8.3)
Na základe výsledkov príkladu 252 je možné konštatovať, že sila testu pomerom vierohodnosti pri danej alternatíve je väčšia ako sila ANOVA F-testu.



Obr. 8.3: Histogramy vygenerovaných testovacích štatistík za platnosti alternatívnej hypotézy v relatívnej škále superponované s teoretickými krivkami hustoty F_W (vľavo) a U_{LR} (vpravo)

Príklad 253 (testy o rovnosti stredných hodnôt; simulačná štúdia) *Na základe simulačnej štúdie preverte, že ak (a) $Y_j \sim N(\mu_j, \sigma^2)$, kde $\mu_j = 1, \sigma^2 = 1/4$, $j = 1, 2, \ldots, J$ a (b) $Y_j \sim [(1-p)N(\mu_j, \sigma^2) + pN(\mu_j, \sigma_1^2)]$, kde $p = 0.05$ a $\sigma_1^2 = 1$, potom testovacia štatistika (1) F_W má $F_{J-1,n-J}$ rozdelenie s $J-1$ a $n-J$ a (2) U_{LR} má χ^2_{J-1} rozdelenie s $J-1$ stupňami voľnosti. Použite rozsahy náhodných výberov $n_j = 10$ a $n_j = 50$ ($n = \sum_{j=1}^{J} n_j$). Pre každú simuláciu vypočítajte $F_{obs,m}$ a $u_{LR,m}$, $m = 1, 2, \ldots, M$, kde $M = 10000$. Superponujte histogram vygenerovaných testovacích štatistík v relatívnej škále s teoretickou krivkou hustoty (1) F_W a (2) U_{LR}.*

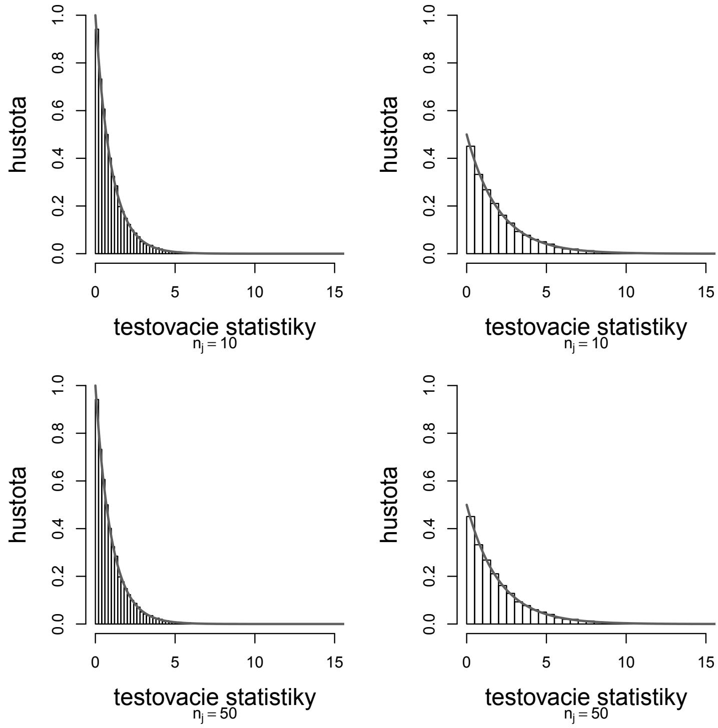
Riešenie (riešenie (a) pozri na obrázku 8.4)
Na základe simulačnej štúdie z príkladu 253 je lepšie používať ANOVA F-test ako test pomerom vierohodnosti pre $n \leq 10$, avšak rozdiely medzi výsledkami sú veľmi malé, pretože obe testovacie štatistiky rýchlo konvergujú k svojím teoretickým rozdeleniam.

8.1.2 Metódy mnohonásobného porovnávania

Ak ANOVA F-test zamietne H_0, potom je potrebné zistiť, ktoré rozdiely dvojíc stredných hodnôt sú štatisticky signifikantné na nominálnej hladine významnosti α. Môžeme tak urobiť pomocou **post-hoc testov**. Základným predpokladom ich použitia je, rovnako ako pre ANOVA model, splnenie podmienky homogenity rozptylov a normality Y_{ji} a chýb ε_{ji}. Ekvivalentnou H_0 je nasledovná hypotéza $H_0 : \mu_i = \mu_j$ pre $\forall i, j; i < j$. Prepíšme H_0 do všeobecnejšieho tvaru

$$H_0 : \sum_{j=1}^{J} a_j \mu_j = \sum_{j=1}^{J} a_j \mu_{0j} \text{ oproti } \sum_{j=1}^{J} a_j \mu_j \neq \sum_{j=1}^{J} a_j \mu_{0j} \text{ pre nejaké } \mathbf{a} = (a_1, a_2, \ldots, a_J)^T \in \mathcal{A},$$

kde $\mathcal{A} = \{\mathbf{a} : \sum_{j=1}^{J} a_j = 0\}$ a $\mathbf{a}$ je **vektor kontrastov**. Kontrasty nie sú nezávislé, pozri napr. Casella a Berger (2002). Vo všeobecnosti však môžeme prepokladať, že H_0 generuje podpriestor s hodnosťou h. Potom definujme $H_0 = \cap_{k=1}^{h} H_{0k}$, kde $h = \binom{J}{2} = J(J-1)/2$, ak ide o všetky párové porovnania.



Obr. 8.4: Histogramy vygenerovaných testovacích štatistík (za platnosti nulovej hypotézy) v relatívnej škále superponované s teoretickými krivkami hustoty F_W (vľavo) a U_{LR} (vpravo)

V prípade, že J-ta z porovnavaných populácií je **kontrolná** (charakterizovaná μ_J) a ostatné majú byť porovnávané len s touto kontrolnou populáciou a nie medzi sebou, potom volíme $h = J-1$ a zaujímame sa len napr. o rozdiely tvaru $\left|\overline{y}_{j\cdot} - \overline{y}_{J\cdot}\right|$, kde $j = 1, 2, \ldots, J-1$. Najprv testujeme H_0 viacvýberovým ANOVA F-testom na hladine významnosti α použitím ANOVA F-štatistiky. Ak H_0 nezamietame, nepokračujeme ďalej. Ak H_0 zamietame, chceme identifikovať, ktorú z hypotéz $\mathbf{a}^T\boldsymbol{\mu} = \mathbf{a}^T\boldsymbol{\mu}_0 = 0$, kde $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_J)^T$, zamietame (pre fixné $\mathbf{a}$). Počet hypotéz h poznáme vopred, ale množiny $\mathcal{H}_0 = \{k : H_{0k} = 0\}$ a $\mathcal{H}_1 = \{k : H_{0k} = 1\}$, t.j. množiny nezamietnutých a zamietnutých nulových hypotéz z množiny všetkých nulových hypotéz $\mathcal{H} = \mathcal{H}_0 \cup \mathcal{H}_1 = \{1, 2, \ldots, h\}$, kde $h = h_0 + h_1$, $h_0 = \text{card}\{\mathcal{H}_0\}$ a $h_1 = \text{card}\{\mathcal{H}_1\}$, dopredu nepoznáme.

Pre $H_{0,ij} : \mu_i = \mu_j, i < j$, bude vektor kontrastov $\mathbf{a}_k$ mať na i-tom mieste -1, na j-tom mieste 1, ostatné sú nuly, napr.

$$\mathbf{a}_1 = (1, -1, 0, \ldots, 0)^T, \mathbf{a}_2 = (0, 1, -1 \ldots, 0)^T, \ldots, \mathbf{a}_{J-1} = (0, 0, \ldots, 1, -1)^T,$$

z čoho vyplýva, že

$$\mathbf{a}_1 \Rightarrow \mu_1 = \mu_2, \mathbf{a}_2 \Rightarrow \mu_2 = \mu_3, \ldots, \mathbf{a}_{J-1} \Rightarrow \mu_{J-1} = \mu_J,$$

čo implikuje $\mu_1 = \mu_2 = \ldots = \mu_J = \mu$.

Pre nejaký vektor $\mathbf{a}$ je stredná hodnota $E[\sum_{j=1}^J a_j \overline{Y}_{j\cdot}] = \sum_{j=1}^J a_j\mu_j$ a rozptyl $Var[\sum_{j=1}^J a_j \overline{Y}_{j\cdot}] = \sigma_e^2 \sum_{j=1}^J \frac{a_j^2}{n_j}$. Potom

$$Z_W = \frac{\sum_{j=1}^J a_j \overline{Y}_{j\cdot} - \sum_{j=1}^J a_j \mu_{j0}}{\sqrt{\sigma_e^2 \sum_{j=1}^J \frac{a_j^2}{n_j}}} \overset{\mathcal{D}}{\sim} N(0, 1).$$

Rozptyl σ_e^2 nepoznáme a musíme ho odhadnúť. Výberový rozptyl v j-tej populácii je rovný $S_j^2 = \frac{1}{n_j-1}\sum_{j=1}^{n_j}(Y_{ji}-\overline{Y}_{j\cdot})^2$, kde $j=1,2,\ldots,J$, sú nezávislé. Potom platí $(n_j-1)S_j^2/\sigma_e^2 \sim \chi^2_{n_j-1}$. Keďže v modeloch $\mathcal{F}_{H_0}$ a $\mathcal{F}_{H_1}$ predpokladáme rovnosť rozptylov, potom môžeme písať $\widehat{\sigma}_e^2$ ako $s^2 = \frac{1}{n-J}\sum_{j=1}^{J}(n_j-1)s_j^2 = \frac{1}{n-J}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left(y_{ji}-\overline{y}_{j\cdot}\right)^2 = \frac{1}{n-J}SS_{e,\text{obs}}$, kde $n-J=\sum_{j=1}^{J}(n_j-1)$. Potom $(n-J)S^2/\sigma_e^2 \sim \chi^2_{n-J}$. Navyše S^2 je nezávislé na $\overline{Y}_{j\cdot}$, a teda môžeme písať

$$T_{\mathbf{a}} = \frac{\mathbf{a}^T\widehat{\boldsymbol{\mu}} - \mathbf{a}^T\boldsymbol{\mu}_0}{\sqrt{S^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}} = \frac{\sum_{j=1}^{J}a_j\overline{Y}_{j\cdot} - \sum_{j=1}^{J}a_j\mu_{j0}}{\sqrt{S^2\sum_{j=1}^{J}\frac{a_j^2}{n_j}}} \overset{\mathcal{D}}{\sim} t_{n-J},$$

kde je matica plánu $\mathbf{X}$ použitá bez prvého stĺpca (charakterizujúceho celkovú strednú hodnotu μ) a má preto rozmery $n\times J$ (podobne aj ďalej v celej kapitole 8). Realizáciou $T_{\mathbf{a}}$ je $t_{\mathbf{a}}$, p-hodnota $=\Pr(T_{\mathbf{a}} \geq |t_{\mathbf{a}}||H_0)$ a H_0 zamietame, ak $|t_{\mathbf{a}}| \geq t_{n-J}(\alpha/2)$; $t_{n-J}(\alpha/2)$ je kritická hodnota t rozdelenia s $n-J$ stupňami voľnosti. $T_{\mathbf{a}} = T_{LSD}$ je **Waldova testovacia štatistika**, často nazývaná aj **Fisherova LSD štatistika** (z angl. *least significant difference*, t.j. najmenší signifikantný rozdiel) a test **viacvýberový Fisherov LSD test o lineárnom kontraste** $\sum_{j=1}^{J}a_j\mu_j$. Potom môžeme definovať **Waldov** $100\times(1-\alpha)\%$ **empirický IS pre nejakú lineárnu kombináciu** $\sum_{j=1}^{J}a_j\mu_j$ (nazývaný aj **empirický IS Fisherovho typu**) ako

$$\left(\sum_{j=1}^{J}a_j\overline{y}_{j\cdot} - t_{n-J}(\alpha/2)\sqrt{s^2\sum_{j=1}^{J}\frac{a_j^2}{n_j}}, \sum_{j=1}^{J}a_j\overline{y}_{j\cdot} + t_{n-J}(\alpha/2)\sqrt{s^2\sum_{j=1}^{J}\frac{a_j^2}{n_j}}\right).$$

Označme

$$T_{\mathbf{a}}^2 = \frac{\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - \mathbf{a}^T\boldsymbol{\mu}_0\right)^2}{S^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}} = \frac{\left(\sum_{j=1}^{J}a_j\overline{Y}_{j\cdot} - \sum_{j=1}^{J}a_j\mu_{j0}\right)^2}{S^2\sum_{j=1}^{J}\frac{a_j^2}{n_j}} = \frac{\left(\sum_{j=1}^{J}a_j(\overline{Y}_{j\cdot} - \mu_{j0})\right)^2}{S^2\sum_{j=1}^{J}\frac{a_j^2}{n_j}}.$$

Potom

$$\sup_{\mathbf{a}:\sum_{j=1}^{J}a_j=0} T_{\mathbf{a}}^2 = \frac{\sum_{j=1}^{J}n_j\left((\overline{Y}_{j\cdot}-\overline{Y}_{\cdot\cdot})-(\mu_{j0}-\mu)\right)^2}{S^2} = (J-1)F_W,$$

kde $\overline{Y}_{\cdot\cdot} = \sum_{j=1}^{J}n_j\overline{Y}_{j\cdot}/\sum_{j=1}^{J}n_j$ a $\mu = \sum_{j=1}^{J}n_j\mu_J/\sum_{j=1}^{J}n_j$. Navyše

$$\sup_{\mathbf{a}:\sum_{j=1}^{J}a_j=0} T_{\mathbf{a}}^2 \overset{\mathcal{D}}{\sim} (J-1)F_{J-1,n-J}.$$

Čitateľ a menovateľ $(J-1)F_W$ sú nezávislé. Tiež platí $S^2 \sim \sigma_e^2\chi^2_{n-J}/(n-J)$ a

$$1/\sigma_e^2\left(\sum_{j=1}^{J}n_j\left((\overline{Y}_{j\cdot}-\overline{Y}_{\cdot\cdot})-(\mu_{j0}-\mu)\right)^2\right) \sim \chi^2_{J-1}.$$

Scheffe ukázal, že

$$F_{\mathbf{a}} = \frac{\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - \mathbf{a}^T\boldsymbol{\mu}_0\right)^2}{S^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}} = \frac{(\sum_{j=1}^{J}a_j\overline{Y}_{j\cdot} - \sum_{j=1}^{J}a_j\mu_{j0})^2}{S^2\sum_{j=1}^{J}a_j^2/n_j} \overset{\mathcal{D}}{\sim} (J-1)F_{J-1,n-J},$$

kde H_0 zamietame, ak $F_{\mathbf{a}} \geq (J-1)F_{J-1,n-J}(\alpha)$, kde $F_{J-1,n-J}(\alpha)$ je kritická hodnota F rozdelenia s $J-1$ a $n-J$ stupňami voľnosti. Je potrebné zdôrazniť, že H_0 musí platiť pre všetky kontrasty $\mathbf{a}$ simultánne a H_0 zamietame, ak zamietame hypotézu o supréme $T_{\mathbf{a}}^2$, t.j. zamietame H_0 v ANOVA F-teste. $F_{\mathbf{a}}$ sa nazýva **Waldova testovacia štatistika**, často nazývaná aj **Scheffeho štatistika** a test **viacvýberový Scheffeho test nulovosti všetkých kontrastov**. Realizáciou $F_{\mathbf{a}}$ je $F_{\mathbf{a},\text{obs}}$

a (adjustovaná) p-hodnota $= \Pr(F_{\mathbf{a}} \geq F_{\mathbf{a},\mathrm{obs}}|H_0)$. **Waldove simultánne** $100 \times (1-\alpha)\%$ **empirické intervaly spoľahlivosti Scheffeho typu** definujeme ako

$$\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - \sqrt{(J-1)F_{J-1,n-J}(\alpha)}\sqrt{\widehat{\sigma}_e^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}, \mathbf{a}^T\widehat{\boldsymbol{\mu}} + \sqrt{(J-1)F_{J-1,n-J}(\alpha)}\sqrt{\widehat{\sigma}_e^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}\right),$$

kde pravdepodobnosť pokrytia všetkých IS (simultánne) je rovná $1-\alpha$. Za simultánnu inferenciu (t.j. testovanie H_{0k}) platíme dĺžkou simultánnych IS Scheffeho typu oproti IS Fisherovho typu, t.j. keďže garantujeme simultánny koeficient spoľahlivosti $1-\alpha$, simultánne IS Scheffeho typu môžu byť dosť široké (platí $t_{n-J}(\alpha/2) \leq \sqrt{(J-1)F_{J-1,n-J}(\alpha)}$). Z čoho vyplýva, že Sheffeho testy majú menšiu silu ako t-testy (Casella a Berger, 2002).

Tukey ukázal, že

$$\sup_{\mathbf{a}:\sum_{j=1}^J a_j=0} T_{\mathbf{a}} = \frac{\overline{Y}_{\max\cdot} - \overline{Y}_{\min\cdot}}{S\sqrt{\frac{1}{2}\left(\frac{1}{n_{\max}\overline{Y}_{\max\cdot}} + \frac{1}{n_{\min}\overline{Y}_{\min\cdot}}\right)}} \overset{\mathcal{D}}{\sim} q_{J,n-J},$$

kde $\overline{Y}_{\max\cdot} = \max_{\forall j}\overline{Y}_{j\cdot}$ a jemu prislúchajúci rozsah $n_{\max}$, $\overline{Y}_{\min\cdot} = \min_{\forall j}\overline{Y}_{j\cdot}$ a jemu prislúchajúci rozsah $n_{\min}$. Potom

$$F_{\mathbf{a}} = \frac{\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - \mathbf{a}^T\boldsymbol{\mu}_0\right)^2}{S^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}} \overset{\mathcal{D}}{\sim} \frac{1}{2}q^2_{J,n-J},$$

kde H_0 zamietame, ak $F_{\mathbf{a}} \geq \frac{1}{2}q^2_{J,n-J}(\alpha)$, kde $q_{J,n-J}(\alpha)$ je kritická hodnota **studentizovaného rozpätia** s J a $n-J$ stupňami voľnosti. Je potrebné opäť zdôrazniť, že H_0 musí platiť pre všetky kontrasty $\mathbf{a}$ simultánne a H_0 zamietame, ak zamietame hypotézu o supréme $T_{\mathbf{a}}$. $F_{\mathbf{a}}$ sa nazýva **Waldova testovacia štatistika**, často nazývaná aj **Tukeyho HSD štatistika** (alebo **Tukey-Kramerova štatistika**; HSD z angl. *honest significant difference*, t.j. „skutočný“ signifikantný rozdiel) a test **viacvýberový Tukeyho HSD test nulovosti všetkých kontrastov**. Realizáciou $F_{\mathbf{a}}$ je $F_{\mathbf{a},\mathrm{obs}}$ a (adjustovaná) p-hodnota $= \Pr(F_{\mathbf{a}} \geq F_{\mathbf{a},\mathrm{obs}}|H_0)$. **Waldove simultánne** $100 \times (1-\alpha)\%$ **empirické intervaly spoľahlivosti Tukeyho typu** definujeme ako

$$\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - q_{J,n-J}(\alpha)\sqrt{\frac{1}{2}\widehat{\sigma}_e^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}, \mathbf{a}^T\widehat{\boldsymbol{\mu}} + q_{J,n-J}(\alpha)\sqrt{\frac{1}{2}\widehat{\sigma}_e^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}\right).$$

Príklad 254 (test trendu v jednofakorovom ANOVA modeli) *Majme jednofaktorový ANOVA model, v ktorom chceme testovať H_0 o **lineárnom trende**, kde $\alpha_1 < \alpha_2 < \ldots < \alpha_J$ a naviac $\alpha_j = aj$. Ak napr. $J = 3$, potom $H_0 : \alpha_2 - \alpha_1 = \alpha_3 - \alpha_2$ alebo $H_0 : \alpha_1 - 2\alpha_2 + \alpha_3 = 0$, kde $\mathbf{a} = (1, -2, 1)^T$.*

Definícia 97 (chyba porovnávania α_c) ***Chyba porovnávania*** *(comparison-wise error, CWER) α_c je pravdepodobnosť zamietnutia práve jednej H_{0k}, keď táto H_{0k} je pravdivá, t.j. ide o pravdepodobnosť, že nastane práve jedna CHPD v jednom párovom porovnaní.*

Definícia 98 (experimentálna chyba α_e) ***Experimentálna chyba*** *(experiment-wise error, EWER) α_e je pravdepodobnosť zamietnutia aspoň jednej H_{0k}, keď všetky H_{0k} sú pravdivé, t.j. ide o pravdepodobnosť, že nastane aspoň jedna CHPD medzi všetkými h nezávislými párovými porovnávaniami. Táto chyba je kontrolovaná na nominálnej hladine významnosti α.*

Z definícií 97 a 98 vyplýva, že Pr(CHPD) jedného testu je rovná α_c a pravdepodobnosť správneho rozhodnutia je $1-\alpha_c$. Za predpokladu, že máme h nezávislých párových porovnávaní, bude mať náhodná premenná V (počet CHPD) binomické rozdelenie, t.j. $V \sim Bin(h, \alpha_c)$. Keďže α_e je pravdepodobnosť, že nastane aspoň jedna CHPD medzi všetkými h nezávislými párovými porovnávaniami, môžeme ju definovať nasledovne

$$\alpha = \alpha_e = \Pr(V \geq 1) = 1 - \Pr(V = 0) = 1 - \binom{J}{2}\alpha_c^0(1-\alpha_c)^h = 1 - (1-\alpha_c)^h.$$

Z tejto rovnosti vyplýva, že ak sa počet párových porovnávaní zväčšuje, α_e sa blíži k jednotke (pozri tabuľku 8.1). Ak $h = 1$ (dvojvýberový prípad), potom $\alpha = \alpha_e = \alpha_c$.

Tabuľka 8.1: Experimentálna chyba α_e ako funkcia α_c a h

α_c/h	2	5	10	20	50
0.01	0.0199	0.0490	0.0956	0.1821	0.3950
0.05	0.0975	0.2262	0.4013	0.6415	0.9231
0.10	0.1900	0.4095	0.6513	0.8784	0.9948

Zamerajme sa na hodnotenie **zovšeobecnenej pravdepodobnosti CHPD** v podobe

1. **pravdepodobnosti najmenej jednej CHPD**, kde V je počet zamietnutých pravdivých H_{0k} (*family-wise error rate*, FWER: metódy napr. Fisher (1935) LSD, Tukey (1949, 1953, 1991) HSD, Scheffe (1953), Bonferroni (1936), Šidák (1967), Holm (1979), Hochberg (1988)); FWER $= \Pr(V \geq 1)$; **FWER adjustované (upravené) p-hodnoty** sú definované nasledovne
$$\widetilde{p}_k = \inf\left\{\alpha : H_{0k} \text{ zamietame na FWER } = \alpha\right\};$$

2. **očakávanej hodnoty podielu CHPD medzi zamietnutými hypotézami**, FDR $= E[V/R]$, ak $R > 0$ alebo 0, ak $R = 0$, kde R je počet zamietnutých pravdivých a nepravdivých H_0, FDP $= V/R$ (*false dicovery rate*, FDR, *false discovery proportion*, FDP: metódy napr. Benjamini a Hochberg (1995), Benjamini a Yekutieli (2001)); **FDR adjustované p-hodnoty** sú definované nasledovne
$$\widetilde{p}_k = \inf\left\{\alpha : H_{0k} \text{ zamietame na FDR } = \alpha\right\};$$

Kontrola FWER a FDR znamená nasledovné: FWER $\leq \alpha$ a $\Pr(\text{FDP} > \gamma) \leq \alpha$, kde $\gamma, \alpha \in (0,1)$. Rovnako je možné zovšeobecniť aj silu testu (pozri napr. Dudoit a van der Laan, 2008).

Aby bolo možné robiť simultánnu inferenciu, je potrebné modifikovať kritickú hodnotu $t_{n-J}(\alpha/2)$ rozdelenia Fisherovej LSD štatistiky pomocou substitúcie $\alpha/2$ použitím jedno- a viackrokových metód. **(Jednokroková) Bonferroniho**, resp. **Šidákova metóda** sú založené na princípe zmenšenia argumentu $\alpha/2$ kritickej hodnoty t-rozdelenia s $n-J$ stupňami voľnosti (obojstranný test) na základe *Bonferroniho*, resp. *Šidákovej nerovnosti*, $\Pr\left(\cup_{k=1}^h A_k\right) \leq \sum_{k=1}^h \Pr(A_k)$, $\Pr\left(\cup_{k=1}^h A_k\right) \leq 1-\alpha$, resp. $\Pr\left(\cap_{k=1}^h A_k\right) \geq \prod_{k=1}^h \Pr(A_k), \Pr\left(\cap_{k=1}^h A_k\right) \geq 1-\alpha$, na $\alpha/(2h)$, resp. $1-(1-\alpha/2)^{1/h}$, kde A_k je najaká udalosť. Bonferroniho metóda je *konzervatívnejšia* ako Šidákova (vedie ku menšiemu počtu zamietnutí, t.j. kritické hodnoty sú väčšie), lebo platí $(1-\alpha)^{1/h} < 1-\alpha/h$ pre všetky $\alpha > 0, h > 1$, teda $t_{n-J}(\alpha/h) > t_{n-J}(1-(1-\alpha)^{1/h})$. Rozdiel je ale zanedbateľný. V súvislosti s kontrolou FWER a adjustovanými p-hodnotami platí pre Bonferroniho nerovnosť
$$\text{FWER} = \Pr(V>0) = \Pr\left(\cup_{k=1}^h\left(\widetilde{P}_k \leq \alpha\right)\right) \leq \sum_{k=1}^{h_0} \Pr\left(\widetilde{P}_k \leq \alpha\right) \leq \sum_{k=1}^{h_0} \frac{\alpha}{h} = h_0\frac{\alpha}{h} \leq \alpha.$$

a pre Šidákovu nerovnosť
$$\begin{aligned}\Pr(V=0) &= \Pr\left(\cap_{k=1}^h\left(\widetilde{P}_k \geq \alpha\right)\right) = \prod_{k=1}^{h_0} \Pr\left(\widetilde{P}_k \geq \alpha\right) \\ &= \prod_{k=1}^{h_0} \Pr\left(P_k \geq 1-(1-\alpha)^{1/h}\right) = 1-(1-\alpha)^{h_0/h},\end{aligned}$$

z ktorej vyplýva, že FWER $= \Pr(V>0) = 1-\Pr(V=0) = (1-\alpha)^{h_0/h} \leq \alpha$.

Ak použijeme vyššie uvedené postupy na h párových porovnaní, potom pravdepodobnosť, že aspoň raz chybne zamietneme jednu z rovností $\mu_i = \mu_j$, ktorá platí, nie je väčšia ako α. Ak ide o vyvážené triedenie, kde $n_1 = n_2 = \ldots = n_J$, potom je táto pravdepodobnosť presne rovná α. T.j. ak sú všetky hypotézy pravdivé, pravdepodobnosť identifikácie, že niektorá z nich je nepravdivá, nie je viac ako α, pretože α je pravdepodobnosť zamietnutia ANOVA F-testu. Taktiež ANOVA F-test je test všetkých $\mathbf{a}_k^T\boldsymbol{\mu} = 0$, $k = 1, 2, \ldots, h$, a ak je tento test zamietnutý, ešte nemusí nastať situácia, že niektorá z vyššie spomenutých metód nezamietne nejakú hypotézu. Práve pre túto vlastnosť je experimentálna chyba menšia ako α. Ale ak ANOVA F-test zamieta nulovú hypotézu, potom Scheffeho metóda bude zamietať

H_0 aspoň pre jeden kontrast.

Nech $SS\left(\mathbf{a}_k^T\boldsymbol{\mu}\right) = \left(\sum_{j=1}^J a_{kj}\overline{y}_{j.} - \sum_{j=1}^J a_{kj}\mu_j\right)^2 / \left(\sum_{j=1}^J a_{kj}^2/n_j\right)$. LSD metóda je silnejšia ako Scheffeho metóda. Scheffeho metóda zamieta H_0, ak $SS\left(\mathbf{a}_k^T\boldsymbol{\mu}\right)/MS_e > hF_{h,df_e}(\alpha)$, LSD metóda zamieta H_0, ak $SS\left(\mathbf{a}_k^T\boldsymbol{\mu}\right)/MS_e > F_{1,df_e}(\alpha)$. Navyše platí, že $hF_{h,df_e}(\alpha) > F_{1,df_e}(\alpha)$.

Ak H_0 zamietneme ANOVA F-testom, potom je LSD metoda silnejšia ako Šidákova a Bonferroniho metóda, pretože $F_{1,df_e}(\alpha/h) > F_{1,df_e}(1-(1-\alpha)^{1/h}) > F_{1,df_e}(\alpha)$. Bonferroniho metóda je dizajnovaná na konečné množstvo hypotéz, a teda nie je prekvapujúce, že je silnejšia ako Scheffeho metóda pre testovanie h hypotéz, ak h nie je príliž veľké.

Príklad 255 (Studentov t-test vs. Tukeyho LSD metóda) *Pomocou simulačnej štúdie, vygenerujte $M = 10000$ náhodných výberov z $N(0,1)$ pre $n = 24$, $J = 6$, vypočítajte t-štatistiky (1) Fisherove $T_{LSD,m,j}$, $m = 1,2,\ldots,M$, $j = 1,2,\ldots,4$, a (2) maximum Fisherových štatistík $\max_{\forall j}(T_{LSD,m,j})$, čo je Tukeyho štatistika $T_{HSD,m}$. Nakreslite hustoty (1) a (2), odhadnuté pomocou funkcie* density()*, do jedného obrázka spolu s ich korešpondujúcimi kvantilmi.*

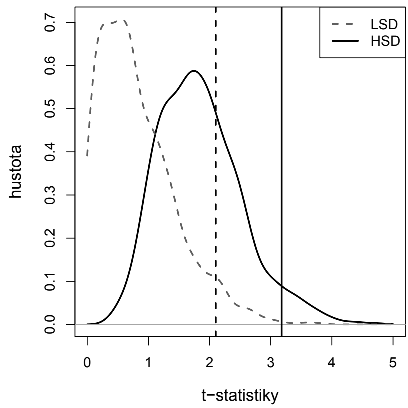
Riešenie v R (pozri obrázok 8.5)

```
x <- factor(rep(LETTERS[1:6], rep(4,6)))
n <- 24
J <- 6
alfa <- 0.05
model.sim <- summary(lm(rnorm(24)~x))
model.sim$coef[,1]  # regresne koeficienty
#(Intercept)          xB            xC            xD            xE            xF
# -0.1517165    0.4501536    0.1832767    0.1540238    0.2710701   -0.1914448
model.sim$sig # 1.063496 odmocnina z MSe
model.sim$coef[2,1]*sqrt(2)/model.sim$sig # 0.9493587; t-statistika pre A a B
kvantil.t <- qt(1-alfa/2,n-J) # 2.100922; kvantil t-rozdelenia
# rozsah t-statistik
rozsah.t.stat <- range(c(0,model.sim$coef[-1,1])) # -0.1914448  0.4501536
# maximum t-statistik
(rozsah.t.stat[2]-rozsah.t.stat[1])*sqrt(2)/model.sim$sig # 0.8531835
# simulacia pre M=1000
M <- 1000
TESTSTAT <- matrix(0,M,2)
for(i in 1:M){
model.sim <- summary(lm(rnorm(24)~x))
TESTSTAT[i,1] <- abs(model.sim$coef[2,1]*sqrt(2)/model.sim$sig)
rozsah.t.stat <- range(c(0,model.sim$coef[-1,1]))
TESTSTAT[i,2] <- (rozsah.t.stat[2]-rozsah.t.stat[1])*sqrt(2)/model.sim$sig
}
# podiel signifikantnych t-statistik, porovnanie A a B
sum(TESTSTAT[,1]>kvantil.t)/M # 0.05
# podiel signifikantnych max t-statistik
sum(TESTSTAT[,2]>kvantil.t)/M # 0.315
# kriticka hodnota nad ktorou uz zamietame
quantile(TESTSTAT[,2],1-alfa) # 3.158423
# kvantil studentizovaneho rozpatia
kvantil.qt <- qtukey(1-alfa,J,n-J)/sqrt(2) # 3.178035
# obrazok
windows(5,5)
par(mar=c(4.5,4.5,1,1))
hust.max <- density(TESTSTAT[,2],from=0,to=5)
hust.AB <- density(TESTSTAT[,1],from=0,to=5)
matplot(hust.AB$x,cbind(hust.max$y,hust.AB$y),type="l",lwd=2,
        xlab="t-statistiky",ylab="hustota",cex.lab=1.2)
abline(h=0,col="grey")
abline(v=kvantil.t,lty=2,lwd=2)
abline(v=kvantil.qt,lwd=2)
legend("topright",c("LSD","HSD"),col=c("red","black"),lty=c(2,1),lwd=c(2,2))
```

Porovnanie LSD a HSD metódy pomocou simulácie 1000 náhodných výberov z $N(0,1)$, kde $n = 24$, $J = 6$, ukazuje, že použitie LSD vedie ku zvýšeniu Pr(CHPD) až na $\alpha_{\mathrm{LSD}} = 0.32 \gg \alpha$ bez použitia ANOVA F-testu (pozri obrázok 8.5), t.j. použitie kritických hodnôt $t_{n-J}(\alpha/2)$ namiesto $q_{J,n-J}(\alpha)$ pri teste maximálneho rozdielu stredných hodnôt vedie ku 32% riziku zamietnutia H_0, keď H_0 platí.

Všeobecne LSD je metóda silnejšia ako ostatné metódy na detekciu pravdivých $\mathbf{a}_k^T\boldsymbol{\mu} \neq \mathbf{a}_k^T\boldsymbol{\mu}_0$, ale ak nastane $\mathbf{a}_k^T\boldsymbol{\mu} = \mathbf{a}_k^T\boldsymbol{\mu}_0$, potom je vyššia Pr(CHPD), ako pri iných metódach, t.j. pravdepodobnosť, že takúto hypotézu identifikujeme ako rozdielnu od nuly.

Čím je menšia hodnosť (dimenzia) h podpriestoru konečného počtu hypotéz, tým je väčšia sila Scheffeho testu. Majme dva podpriestory, jeden s dimenziou h_A a druhý s dimenziou h_B, kde $h_A > h_B$. Je garantované, že experimentálna chyba nepresiahne α. Silnejšia procedúra je tá, ktorá pravdivo zamieta častejšie. Na základe Scheffeho metódy zamietame v prvom prípade, ak $SS\left(\mathbf{a}_k^T\boldsymbol{\mu}\right)/MS_e > h_AF_{h_A,df_e}(\alpha)$ a v druhom prípade, ak $SS\left(\mathbf{a}_k^T\boldsymbol{\mu}\right)/MS_e > h_BF_{h_B,df_e}(\alpha)$. Keďže $h_A > h_B$, potom



Obr. 8.5: Odhadnuté hustoty vybraných t-štatistík T_{LSD} (pre rozdiel populácií A a B) a maximálnych t-štatistík (T_{HSD}) spolu s korešpondujúcimi teoretickými kritickými hodnotami $t_{n-J}(0.025)$, resp. a $q_{J,n-J}(0.05)$

$h_A F_{h_A,df_e}(\alpha) \geq h_B F_{h_B,df_e}(\alpha)$, teda očakávame viac zamietnutí v podpriestore s dimenziou h_B a navyše táto procedúra je silnejšia (Christensen, 2002). Jednou z výhod Scheffeho metódy je, že ide o možný test akejkoľvek hypotézy v podpriestore, môžeme sa pozrieť do dát a vybrať H_0 a test zostáva validný. Táto metóda je najvhodnejšia na testovanie vybraných kontrastov, ale Pr(CHPD) je za platnosti H_0 výrazne pod nominálnou α, sila testu je tiež dosť nízka v porovnaní s inými metódami.

Všeobecne môžeme písať nasledovné

$$hF_{h,df_e}(\alpha) > F_{1,df_e}(\alpha/h) > F_{1,df_e}(1-(1-\alpha)^{1/h}) > \frac{q^2_{J,df_e}(\alpha)}{2} > F_{1,df_e}(\alpha)\,, h > 1,$$

z čoho vyplýva pre silu nasledovné

$$(1-\beta)_{\text{LSD}} > (1-\beta)_{\text{HSD}} > (1-\beta)_{\text{Šidák}} > (1-\beta)_{\text{Bonferroni}} > (1-\beta)_{\text{Scheffe}}$$

a pre Pr(CHPD)

$$\Pr(\text{CHPD}_{\text{LSD}}) > \Pr(\text{CHPD})_{\text{HSD}} > \Pr(\text{CHPD}_{\text{Šidák}}) > \Pr(\text{CHPD}_{\text{Bonferroni}}) > \Pr(\text{CHPD}_{\text{Scheffe}}),$$

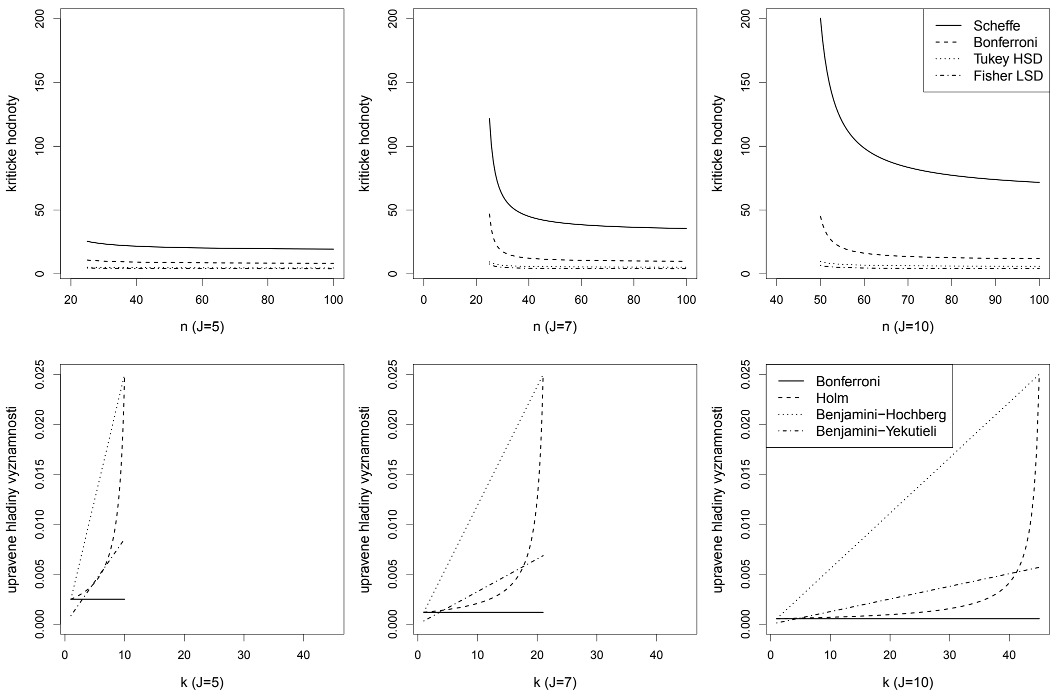
pričom sa rozdiely zväčšujú so zväčšujúcim sa J (pozri obrázok 8.6).

Tukeyho HSD nie je natoľko ovplyvnená malým rozsahom n pri veľkých J ako Bonferroniho a Šidákova metóda. Použitie Scheffeho metódy nie je pri takejto situácii vhodné (kritické hodnoty sú príliž vysoké, pozri obrázok 8.6, prvý riadok). Pri fixnom počte pozorovaní n je sila všetkých metód pri danej alternatíve vyššia pri nižšom J (pozri obrázok 8.1).

Ďalšou skupinou metód sú **sekvenčne zamietajúce (viackrokové) metódy** založené na Bonferroniho a Šidákovom princípe. Hlavným princípom je zoradenie p-hodnôt do neklesajúcej postupnosti vypočítaných na základe napr. T_{LSD}, kde $p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(h)}$. Ďalej nech $\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_h$ sú konštanty a $k = 1, 2, \ldots, h$.

Ak zamietanie prebieha nasledovne – ak $p_{(1)} > \alpha_1$, potom nezamietame žiadnu H_0, inak $p_{(1)} \leq \alpha_1, \ldots, p_{(r)} \leq \alpha_r$ a $H_{0(1)}, \ldots, H_{0(r)}$ sú zamietnuté, kde $r \leq h$ je najväčší index spĺňajúci nerovnosť – hovoríme o ***step-down* procedúre** (začíname s najmenšou p-hodnotou a pokračujeme zamietaním, pokiaľ sú korešpondujúce p-hodnoty malé). Holm (1979) navrhol na základe Bonferroniho princípu tzv. **Holm-Bonferroniho** $\alpha_k = \alpha/(h-k+1)$ a Dudoit a van der Laan (2008) navrhli na základe Šidákovho princípu tzv. **Holm-Šidákove** $\alpha_k = 1-(1-\alpha)^{1/(h-k+1)}$.

Ak zamietanie prebieha nasledovne – ak $p_{(h)} \leq \alpha_h$, potom zamietame všetky H_0, inak $p_{(1)} > \alpha_1, \ldots, p_{(r+1)} > \alpha_{r+1}$ a $H_{0(1)}, \ldots, H_{0(r+1)}$ nie sú zamietnuté, kde r je najmenší index spĺňajúci nerovnosť – hovoríme o ***step-up* procedúre** (začíname s najväčšou p-hodnotou a pokračujeme neza-



Obr. 8.6: Kritické hodnoty – Scheffeho $hF_{h,df_e}(\alpha)$, Bonferroniho $F_{1,df_e}(\alpha/h)$, Tukeyho HSD $q^2_{J,n-J}(\alpha)$ a Fisherove LSD $F_{1,df_e}(\alpha)$, $J = 5, 7, 10$, pre rôzne $n < 100$, kde $n - J > 0$ (prvý riadok); upravené hladiny významnosti α_k – Bonferroniho $\alpha_k = \alpha/h$, Holmove $\alpha_k = \alpha/(h - k + 1)$, Benjamini-Hochbergove $\alpha_k = k\alpha/h$, Benjamini-Yekutieliho $\alpha_k = k\alpha/(h \sum_{l=1}^{h} 1/l)$, $J = 5, 7, 10$, pre vzrastajúce $k = 1, 2, \ldots, h$ (druhý riadok)

mietaním pokiaľ sú korešpondujúce p-hodnoty veľké). Hochberg (1988) navrhol na základe Bonferroniho metódy tzv. **Hochberg-Bonferroniho** $\alpha_k = \alpha/(h - k + 1)$, na základe Šidákovho princípu tzv. **Hochberg-Šidákove** $\alpha_k = 1 - (1 - \alpha)^{1/(h-k+1)}$, Benjamini a Hochberg (1995) navrhli tzv. **Benjamini-Hochbergove** $\alpha_k = k\alpha/h$, Benjamini a Yekutieli (2001) navrhli tzv. **Benjamini-Yekutieliho** $\alpha_k = k\alpha/(h \sum_{l=1}^{h} 1/l)$. Konštanty α_k pre $k = 1$ a $k = h$ sú pre prvé dve metódy rovnaké, ale pre $k \neq 1, h > 0$, Benjamini-Hochbergove α_k sú väčšie, čo indikuje častejšie zamietanie. So zväščujúcim sa počtom porovnávaných skupín sa sklon priamky, prechádzajúcej bodmi (k, α_k) Benjamini-Hochbergových α_k, zmenšuje. Hochberg-Bonferroniho (Holm-Bonferroniho) α_k sa so zvyšujúcim sa J približujú hodnote α_k jednokrokovej Bonferroniho metódy s výnimkou vysokých indexov k, čo hovorí o *liberálnosti* (vedie ku väčšiemu počtu zamietnutí, t.j. kritické hodnoty sú menšie), podobne ako aj jednokroková Bonferroniho metóda (pozri obrázok 8.6, druhý riadok). Benjamini-Yekutieliho α_k boli navrhnuté pre prípady závislosti testovacích štatistík, čo je častý prípad pri aplikácii mnohonásobného porovnávania pri mnohorozmerných (h-rozmerných) problémoch napr. v genetike (závislosť porovnávaných génov) a antropológii (asymetria dlhých kostí), kde pri dvojvýberovom alebo párovom prípade aplikujeme najprv Hotellingov T^2-test a potom h post-hoc dvojvýberových Studentových t-testov pre rozdiel stredných hodnôt, pri viacvýberovom prípade najprv MANOVA F-test a potom h post-hoc ANOVA F-testov s príslušnou korekciou α_k podľa výberu metódy.

Adjustované hladiny významnosti) α_k sú definované nasledovne

- Bonferroniho $\alpha_k = \alpha/h$,
- Šidákove $\alpha_k = 1 - (1 - \alpha)^{1/h}$,
- Holm-Bonferroniho *step-down* $\alpha_k = \alpha/(h - k + 1)$,

- Holm-Šidákove *step-down* $\alpha_k = 1 - (1-\alpha)^{1/(h-k+1)}$,
- Hochberg-Bonferroniho *step-up* $\alpha_k = \alpha/(h-k+1)$,
- Hochberg-Šidákove *step-up* $\alpha_k = 1 - (1-\alpha)^{1/(h-k+1)}$,
- Benjamini-Hochbergove *step-up* $\alpha_k = k\alpha/h$,
- Benjamini-Yekutieliho *step-up* $\alpha_k = k\alpha/(h\sum_{l=1}^{h} 1/l)$.

Argument $\alpha/2$ kritickej hodnoty $t_{n-J}(\alpha/2)$ sa substituuje α_k. Potom budú **Waldove simultánne** $100 \times (1-\alpha)\%$ **empirické intervaly spoľahlivosti Fisherovho typu** definované nasledovne

$$\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - t_{n-J}(\alpha_k)\sqrt{\widehat{\sigma}_e^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}, \mathbf{a}^T\widehat{\boldsymbol{\mu}} + t_{n-J}(\alpha_k)\sqrt{\widehat{\sigma}_e^2\mathbf{a}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{a}}\right).$$

Adjustované p-hodnoty p_k sú definované nasledovne (Yekutieli a Benjamini, 1999)

- Bonferroniho $\widetilde{p}_k = \min\{hp_k, 1\}$,
- Šidákove $\widetilde{p}_k = 1 - (1-p_k)^h$,
- Holm-Bonferroniho *step-down* $\widetilde{p}_{(j)} = \max_{k=1,\ldots,j}\left\{\min\left\{(h-k+1)p_{(k)}, 1\right\}\right\}$,
- Holm-Šidákove *step-down* $\widetilde{p}_{(j)} = \max_{k=1,\ldots,j}\left\{(1-\left(1-p_{(k)}\right)^{h-k+1}\right\}$,
- Hochberg-Bonferroniho *step-up* $\widetilde{p}_{(j)} = \min_{k=j,\ldots,h}\left\{\min\left\{(h-k+1)p_{(k)}, 1\right\}\right\}$,
- Hochberg-Šidákove *step-up* $\widetilde{p}_{(j)} = \min_{k=j,\ldots,h}\left\{(1-\left(1-p_{(k)}\right)^{h-k+1}\right\}$,
- Benjamini-Hochbergove *step-up* $\widetilde{p}_{(j)} = \min_{k=j,\ldots,h}\left\{\min\left\{\frac{h}{k}p_{(k)}, 1\right\}\right\}$,
- Benjamini-Yekutieliho *step-up* $\widetilde{p}_{(j)} = \min_{k=j,\ldots,h}\left\{\min\left\{\frac{h\sum_{l=1}^{h}1/l}{k}p_{(k)}, 1\right\}\right\}$.

Metódy kontrolujúce FWER sú konzervatívnejšie, metódy kontrolujúce FDR sú liberálnejšie (Shaffer, 1995). Viackrokové metódy sú menej konzervatívne než jednokrokové. Teda pri fixovanej nominálnej α použitie FDR metód vedie ku vyššej sile post-hoc testov pri danej alternatíve. FDR je sľubnou alternatívou ku FWER, aj napriek vyššiemu počtu CHPD, pokiaľ ich počet je malý v porovnaní so všetkými zamietnutými hypotézami (pozri napr. Dudoit a van der Laan, 2008). Pri malom počte porovnávaných stredných hodnôt (a teda malom počte post-hoc testov) je vhodné použit napr. Bonferroniho metódu alebo Tukey HSD test. Pri veľkom počte porovnávaných stredných hodnôt je vhodnejšie použit sekvenčné metódy. Pri mnohorozmerných problémoch je vhodné použiť metódu Benjamini a Yekutieli (2001).

Mnohonásobné porovnávania v R

Na Tukeyho HSD metódu použijeme funkciu TukeyHSD(aov(),ordered=FALSE), kde argument ordered ponechá pôvodné poradie hypotéz H_{0k}. Výstupom je tabuľka obsahujúca odhady rozdielov stredných hodnôt $\overline{y}_{i.} - \overline{y}_{j.}$, dolné a horné hranice Waldových simultánnych $100 \times (1-\alpha)\%$ empirických IS Tukeyho typu a adjustované p-hodnoty $\widetilde{p}_k$. Na jednokrokové a viackrokové metódy (výpočet adjustovaných p-hodnôt) použijeme funkciu pairwise.t.test(y,x,p.adjust="metoda",pool.sd=TRUE), kde argument pool.sd=TRUE predstavuje použitie $\widehat{\sigma}_e^2$ a argument p.adjust="metoda" špecifikuje metódu nasledovne

1. Bonferroniho p.adjust="bonferroni",
2. Holm-Bonferroniho p.adjust="holm",
3. Hochberg-Bonferroniho p.adjust="hochberg",

4. Benjamini-Hochbergove p.adjust="BH" a

5. Benjamini-Yekutieliho p.adjust="BY".

Funkcia lm() má prednastavené kontrasty párových porovnaní s prvou populáciou, $\alpha_j - \alpha_1$, kde $J > 1$, kedy použijeme vstupné argumenty y a x. Pokiaľ by sme chceli testovať nulovosť jednotlivých α_j, použijeme vstupné argumenty y-mean(y) a x-1 (argument x-1 znamená model bez interceptu μ).

Minimálny rozsah K môžeme počítať aj na základe post-hoc testov, kde použijeme koeficient necentrality definovaný na základe (1) $\lambda = \sqrt{(J-1)\sigma_A^2/(\sigma_e^2/K)}$ alebo (2) $\lambda = |\delta|/\sqrt{2\sigma_e^2/K}$. Potom

1. Fisherova LSD metóda, kde minimálny rozsah bude riešením iteračných rovníc
$$F_{1,J(K-1),\lambda^2}(\beta) = F_{1,J(K-1)}(1-\alpha),$$
2. Tukeyho HSD metóda, kde minimálny rozsah bude riešením iteračných rovníc
$$F_{1,J(K-1),\lambda^2}(\beta) = \frac{q^2_{J,J(K-1)}(1-\alpha)}{2},$$
3. Šidákova metóda, kde minimálny rozsah bude riešením iteračných rovníc
$$F_{1,J(K-1),\lambda^2}(\beta) = F_{1,J(K-1)}((1-\alpha)^{1/h}),$$
4. Bonferroniho metóda, kde minimálny rozsah bude riešením iteračných rovníc
$$F_{1,J(K-1),\lambda^2}(\beta) = F_{1,J(K-1)}(1-\alpha/h),$$
5. Scheffeho metóda, kde minimálny rozsah bude riešením iteračných rovníc
$$F_{1,J(K-1),\lambda^2}(\beta) = (J-1)F_{J-1,J(K-1)}(1-\alpha),$$

Príklad 256 (minimálny rozsah; post-hoc testy) *Majme štyri populácie, ktorých stredné hodnoty sú $\mu_1 = 390$, $\mu_2 = 405$, $\mu_3 = 415$ a $\mu_4 = 410$. Predpokladajme, že $X \sim N(\mu_j, \sigma_e^2)$. Vypočítajte minimálny rozsah K, ak $\sigma_e^2 = 20^2$ a $\alpha = 0.05$, použitím (1) Fisherovej LSD metódy, (2) Tukeyho HSD metódy, (3) Šidákovej metódy, (4) Bonferroniho metódy a (5) Scheffeho metódy. Výsledky porovnajte s K vypočítaným v príklade 250. Funkcie nazvite nasledovne (1) minimalny.rozsah.Fisher(), (2) minimalny.rozsah.Tukey(), (3) minimalny.rozsah.Sidak(), (4) minimalny.rozsah.Bonferroni() a (5) minimalny.rozsah.Scheffe().*

Riešenie
(1) Fisherova LSD metóda – minimálny rozsah $K = 10$ a $n = KJ = 40$.
(2) Tukeyho HSD metóda – minimálny rozsah $K = 15$ a $n = KJ = 60$.
(3) Šidákova metóda – minimálny rozsah $K = 15$ a $n = KJ = 60$.
(4) Bonferroniho metóda – minimálny rozsah $K = 15$ a $n = KJ = 60$.
(5) Scheffeho metóda – minimálny rozsah $K = 16$ a $n = KJ = 64$.

Porovnaním všetkých použitých metód dostaneme nasledovné poradie minimálnych rozsahov
$$n_{\text{LSD}} < n_{\text{HSD}} = n_{\text{Šidák}} = n_{\text{Bonferroni}} < n_{\text{Scheffe}}.$$
Vzťahy medzi n korešpondujú poradiu kritických hodnôt. Výpočet je ovplyvnený malými rozsahmi K, zaokrúhlovaním n na nula desatinných miest a voľbou h pri Bonferroniho a Šidákovej metóde. Toho dôsledkom sú aj odlišné hodnoty sily, ktorá pre dané n pri rôznych metódach prvýkrát prekročí hranicu 0.8.

Príklad 257 (jednorozmerný prípad) *Majme koncentráciu stroncia Sr (mg/ml) v piatich vodných celkoch (pozri tabuľku 8.2). Otestujte rovnosť stredných hodnôt ANOVA F-testom pomocou funkcií (1)* aov() *(2)* oneway.test() *a (3)* lm()*. Ak je H_0 zamietnutá na $\alpha = 0.05$, potom použite (1) Tukeyho HSD metódu (T_{HSD} štatistiku), vypočítajte adjustované p-hodnoty, simultánne 95% empirické IS Tukeyho typu pre všetky párové porovnania rozdielov stredných hodnôt a zobrazte ich graficky. (2) Vypočítajte adjustované p-hodnoty a Waldove simultánne 95% empirické IS Fisherovho typu pre všetky párové porovnania rozdielov stredných hodnôt (a) Bonferroniho metódou, (b) Holmovou step-down metódou, (c) Hochbergovou step-up metódou, (d) Benjamini-Hochbergovou step-up metódou a (e) Benjamini-Yekutieliho step-up metódou. (3) Po náhľade na dáta vhodne zadefinujte kontrasty a aplikujte na ne Scheffeho metódu.*

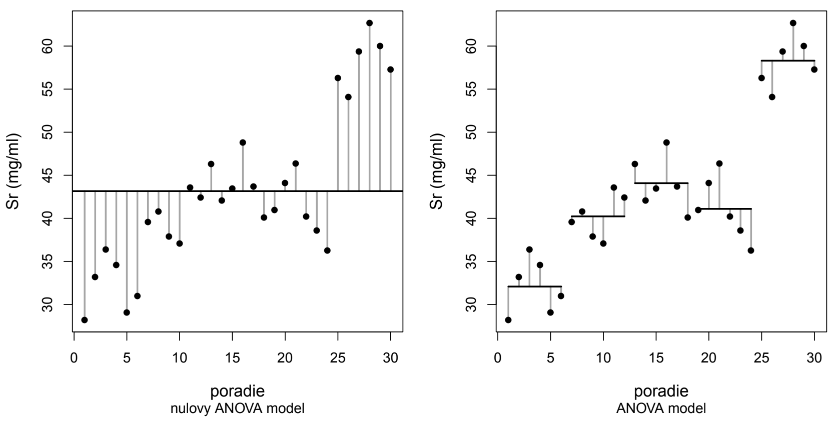
Tabuľka 8.2: Koncentrácia stroncia Sr (mg/ml) v piatich vodných celkoch

A (1)	B (2)	C (3)	D (4)	E (5)
28.2	39.6	46.3	41.0	56.3
33.2	40.8	42.1	44.1	54.1
36.4	37.9	43.5	46.4	59.4
34.6	37.1	48.8	40.2	62.7
29.1	43.6	43.7	38.6	60.0
31.0	42.4	40.1	36.3	57.3

Riešenie (**aj v** R, pozri obrázok 8.7)

```
K <- 6
J <- 5
VodCelk <- factor(rep(LETTERS[1:J],rep(K,J)))
ConcStr.1 <- c(28.2,33.2,36.4,34.6,29.1,31.0)
ConcStr.2 <- c(39.6,40.8,37.9,37.1,43.6,42.4)
ConcStr.3 <- c(46.3,42.1,43.5,48.8,43.7,40.1)
ConcStr.4 <- c(41.0,44.1,46.4,40.2,38.6,36.3)
ConcStr.5 <- c(56.3,54.1,59.4,62.7,60.0,57.3)
ConcStr <- c(ConcStr.1,ConcStr.2,ConcStr.3,ConcStr.4,ConcStr.5)
mean(ConcStr) # 43.16
PRIEM.ConcStr <- tapply(ConcStr,VodCelk,mean)
round(PRIEM.ConcStr,2)
#    A     B     C     D     E
#32.08 40.23 44.08 41.10 58.30
PRIEM.str <- tapply(ConcStr-mean(ConcStr),VodCelk,mean)
round(PRIEM.str,2)
#      A      B      C      D      E
# -11.08  -2.93   0.92  -2.06  15.14
# obrazok -- model FH0 a FH1
windows(10,5)
par(mfcol=c(1,2),mar=c(5,4.5,1,1))
plot(1:30,ConcStr,type="n",ylab="Sr (mg/ml)",xlab="poradie",
     sub="nulovy ANOVA model",cex.lab=1.2)
for(i in 1:30) lines(c(i,i),c(mean(ConcStr),ConcStr[i]),col="grey",lwd=2)
abline(mean(ConcStr),0,lwd=2)
points(1:30,ConcStr,pch=16)
plot(1:30,ConcStr,type="n",ylab="Sr (mg/ml)",xlab="poradie",
     sub="ANOVA model",cex.lab=1.2)
for(i in 1:6) lines(c(i,i),c(PRIEM.ConcStr[1],ConcStr[i]),col="grey",lwd=2)
for(i in 7:12) lines(c(i,i),c(PRIEM.ConcStr[2],ConcStr[i]),col="grey",lwd=2)
for(i in 13:18) lines(c(i,i),c(PRIEM.ConcStr[3],ConcStr[i]),col="grey",lwd=2)
for(i in 19:24) lines(c(i,i),c(PRIEM.ConcStr[4],ConcStr[i]),col="grey",lwd=2)
for(i in 25:30) lines(c(i,i),c(PRIEM.ConcStr[5],ConcStr[i]),col="grey",lwd=2)
lines(c(1,6),c(PRIEM.ConcStr[1],PRIEM.ConcStr[1]),lwd=2)
lines(c(7,12),c(PRIEM.ConcStr[2],PRIEM.ConcStr[2]),lwd=2)
lines(c(13,18),c(PRIEM.ConcStr[3],PRIEM.ConcStr[3]),lwd=2)
lines(c(19,24),c(PRIEM.ConcStr[4],PRIEM.ConcStr[4]),lwd=2)
lines(c(25,30),c(PRIEM.ConcStr[5],PRIEM.ConcStr[5]),lwd=2)
points(1:30,ConcStr,pch=16)

StrMOD01 <- aov(ConcStr~VodCelk)
summary(StrMOD01) # ANOVA tabulka
oneway.test(ConcStr~VodCelk,var.equal=TRUE) # vysledky ANOVA F-testu
StrMOD02 <- lm(ConcStr~VodCelk)
summary(StrMOD02) # vysledky ANOVA F-testu
anova(StrMOD02) # ANOVA tabulka
StrMOD03 <- lm(ConcStr-mean(ConcStr)~VodCelk-1)
summary(StrMOD03)$coef # efekty faktora VodCelk
sqrt((summary(StrMOD03)$sig)^2/K) # 1.275748; odmocnina z (MSe/K)
anova(StrMOD03) # ANOVA tabulka

# Tukeyho HSD metoda pre vybrany kontrast B-A
lambda.AB <- c(-1,1,0,0,0)
```

```
citatel.AB <- sum(lambda.AB*PRIEM.ConcStr) # 8.15
sigmasq.e.hat <- (summary(StrMOD03)$sig)^2 # 9.7652
menovatel.AB <- sqrt(sigmasq.e.hat/2*sum(lambda.AB^2/K)) # 1.275748
t.LSD.AB <- citatel.AB/menovatel.AB # 6.388408
qtukey(0.95,J,K*J-J) # 4.153363
p.hodn <- 1-ptukey(t.LSD.AB,J,K*J-J) # 0.001129311
IS.AB <- citatel.AB+c(-1,1)*qtukey(0.95,J,K*J-J)*menovatel.AB
# 2.851355 13.448645
mp.Tukey <- TukeyHSD(aov(ConcStr~VodCelk),ordered=FALSE) # tab. 41
# graficke znazornenie simultannuch IS
IS.Tukey <- mp.Tukey$VodCelk[,2:3]
mu.a.hat <- mp.Tukey$VodCelk[,1]
range.x <- range(IS.Tukey)
h <- J*(J-1)/2
windows(5,5)
par(mar=c(5,4.5,1,1))
plot(mu.a.hat,1:h,type="n",xlab="Sr (mg/ml)",ylab="",sub="Waldove simultanne 95% empiricke IS Tukeyho typu",
     xlim=range.x,bty="n",axes=FALSE,cex.lab=1.2)
abline(v=0,lty=2)
for (j in 1:h) abline(h=j,col="gray")
for (j in 1:h) arrows(mu.a.hat[h-j+1],j,IS.Tukey[h-j+1,1],j,length=0.05,angle=90,lwd=2)
for (j in 1:h) arrows(mu.a.hat[h-j+1],j,IS.Tukey[h-j+1,2],j,length=0.05,angle=90,lwd=2)
points(rev(mu.a.hat),1:h,pch=16)
axis(1)
axis(2,labels=rownames(mp.Tukey$VodCelk),at=h:1,las=2)
box()

# parove porovnavania
mp.Bonf <- pairwise.t.test(ConcStr,VodCelk,p.adjust="bonferroni",pool.sd=TRUE)
mp.Holm <- pairwise.t.test(ConcStr,VodCelk,p.adjust="holm",pool.sd=TRUE)
mp.Hoch <- pairwise.t.test(ConcStr,VodCelk,p.adjust="hochberg",pool.sd=TRUE)
mp.BH <- pairwise.t.test(ConcStr,VodCelk,p.adjust="BH",pool.sd=TRUE)
mp.BY <- pairwise.t.test(ConcStr,VodCelk,p.adjust="BY",pool.sd=TRUE)
mp.all <- cbind(mp.Tukey$VodCelk[,4],
                na.omit(c(mp.Bonf$p.value)),na.omit(c(mp.Holm$p.value)),
                na.omit(c(mp.Hoch$p.value)),na.omit(c(mp.BH$p.value)),
                na.omit(c(mp.BY$p.value)))
rownames(mp.all) <- rownames(mp.Tukey$VodCelk)
colnames(mp.all) <- c("Tukey HSD","Bonferroni","Holm","Hochberg","Benjamini-Hochberg","Benjamini-Yekutieli")
round(mp.all,10) # tab. 42

# Scheffeho metoda
lambda.1 <- c(0,1/3,1/3,1/3,-1)
citatel.1 <- sum(lambda.1*PRIEM.ConcStr) # -16.49444
sigmasq.e.hat <- (summary(StrMOD03)$sig)^2 # 9.7652
menovatel.1 <- sqrt(sigmasq.e.hat*sum(lambda.1^2/K)) # 1.473107
S1 <- abs(citatel.1)/menovatel.1 # 11.19704
qf(0.95,J-1,K*J-J) # 2.75871
krit.hodn <- sqrt((J-1)*qf(0.95,J-1,K*J-J)) # 3.321873
F.stat <- citatel.1^2/((J-1)*menovatel.1^2) # 31.34345
p.hodn <- 1-pf(S1,J-1,K*J-J) # 2.411503e-05
```

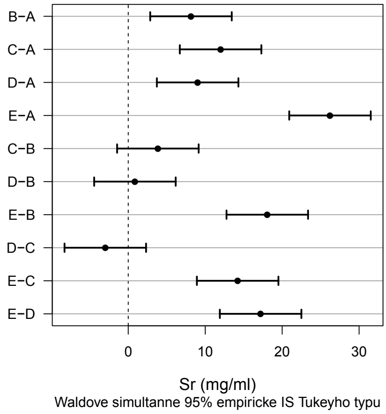


Obr. 8.7: Rozptylové grafy ANOVA modelov – $\mathcal{F}_{H_0}$ (vľavo) a $\mathcal{F}_{H_1}$ (vpravo)

Celkový aritmetický priemer je rovný $\overline{y} = 43.16$. Aritmetické priemery koncentrácií Sr v jednotlivých vodných celkoch sú nasledovné: $\overline{y}_{1.} = 32.08$, $\overline{y}_{2.} = 40.23$, $\overline{y}_{3.} = 44.08$, $\overline{y}_{4.} = 41.10$ a $\overline{y}_{5.} = 58.30$, pre $n_j = 6$, $j = 1, 2, \ldots, 6$. Centrované aritmetické priemery sú rovné – $\overline{y}_{1.} - \overline{y} = -11.08$, $\overline{y}_{2.} - \overline{y} = -2.93$, $\overline{y}_{3.} - \overline{y} = 0.92$, $\overline{y}_{4.} - \overline{y} = -2.06$ a $\overline{y}_{5.} - \overline{y} = 15.14$. Pre aritmetické priemery platí $\overline{y}_{1.} < \overline{y}_{2.} < \overline{y}_{4.} < \overline{y}_{3.} < \overline{y}_{5.}$. ANOVA tabuľka je nasledovná

zdroj variability	suma štvorcov	*df*	priemerné štvorce
medzi súbormi	$SS_{A,\text{obs}} \doteq 2193.442$	$J-1=4$	$MS_{A,\text{obs}} \doteq 548.361$
vnútri súborov	$SS_{e,\text{obs}} \doteq 244.130$	$n-J=25$	$MS_{e,\text{obs}} \doteq 9.765$
celkovo	$SS_{T,\text{obs}} \doteq 2437.572$	$n-1=29$	

Z ANOVA tabuľky vypočítame $F_W \doteq 56.2$, čo je väčšie ako $F_{J-1,n-J}(\alpha) = F_{4,25}(0.05) \doteq 2.76$ (p-hodnota $\ll 0.05$), t.j. H_0 zamietame na $\alpha = 0.05$. Ďalej sme zistili (Tukeyho HSD metódou), že $\mu_1 \neq \mu_2 = \mu_4 = \mu_3 \neq \mu_5$ (pozri tabuľku 8.3 a 8.4 a graf 8.8).



Obr. 8.8: Waldove simultánne 95% empirické IS Tukeyho typu pre rozdiely stredných hodnôt

Tabuľka 8.3: Výsledky Tukey HSD metódy – rozdiely aritmetických priemerov $\overline{y}_{i\cdot} - \overline{y}_{j\cdot}$, dolná a horná hranica Waldových simultánnych 95% empirických IS Tukeyho typu pre $\mu_i - \mu_j$ (DH a HH), adjustované p-hodnoty $\widetilde{p}_k$

	$\overline{y}_{i\cdot} - \overline{y}_{j\cdot}$	DH	HH	$\widetilde{p}_k$
B-A	8.15	2.85	13.45	0.00112931
C-A	12.00	6.70	17.30	0.00000534
D-A	9.02	3.72	14.32	0.00033392
E-A	26.22	20.92	31.52	<0.00000001
C-B	3.85	–1.45	9.15	0.23762175
D-B	0.87	–4.43	6.17	0.98848032
E-B	18.07	12.77	23.37	<0.00000001
D-C	–2.98	–8.28	2.32	0.47910996
E-C	14.22	8.92	19.52	0.00000029
E-D	17.20	11.90	22.50	0.00000001

Po náhľade na dáta použijeme nasledovné tri vektory kontrastov $\mathbf{a}_k$, nim prislúchajúce odhady efektov $\mathbf{a}_k^T \widehat{\boldsymbol{\mu}} = \sum_{j=1}^{J} a_{kj} \overline{y}_{j\cdot}$, ich rozptyly $s^2 \mathbf{a}_k^T (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{a}_k = s^2 \sum_{j=1}^{J} a_{kj}^2 / n_j$ a Scheffeho testovacie štatistiky $\sqrt{(J-1)F} = \left|\mathbf{a}_k^T \widehat{\boldsymbol{\mu}}\right| / \sqrt{s^2 \mathbf{a}_k^T (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{a}_k}$, kde $k = 1, 2$ a 3:

- $\mathbf{a}_1 = (0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1)^T$, $\mathbf{a}_1^T \widehat{\boldsymbol{\mu}} \doteq -9.7$, $s^2 \mathbf{a}_1^T (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{a}_1 = 1.47^2$, $\sqrt{4F} \doteq 11.20$,
- $\mathbf{a}_2 = (1, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, 0)^T$, $\mathbf{a}_2^T \widehat{\boldsymbol{\mu}} \doteq -16.5$, $s^2 \mathbf{a}_2^T (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{a}_2 = 1.47^2$, $\sqrt{4F} \doteq 6.60$,
- $\mathbf{a}_3 = (\frac{1}{2}, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{2})^T$, $\mathbf{a}_3^T \widehat{\boldsymbol{\mu}} \doteq 3.4$, $s^2 \mathbf{a}_3^T (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{a}_3 = 1.16^2$, $\sqrt{4F} \doteq 2.93$.

Scheffeho kritická hodnota je rovná $\sqrt{(J-1)\,F_{J-1,n-J}(\alpha)} = \sqrt{4F_{4,25}(0.05)} \doteq 3.32$. Potom H_{0k} : $\mathbf{a}_k^T \boldsymbol{\mu} = 0$ oproti $H_{1k} : \mathbf{a}_k^T \boldsymbol{\mu} \neq 0$ zamietame, ak $k = 1, 2$, a nezamietame, ak $k = 3$.

Tabuľka 8.4: Adjustované p-hodnoty pre Tukeyho HSD (THSD) metódu, Bonferroniho (B) metódu, Holm-Bonferroniho metódu (HB step-down), Hochberg-Bonferroniho metódu (HB step-up), Benjamini-Hochbergovu (BH) metódu a Benjamini-Yekutieliho (BY) metódu

	THSD	B	HB step-down	HB step-up	BH	BY
B-A	0.00112931	0.00130222	0.00052089	0.00052089	0.00018603	0.00054488
C-A	0.00000534	0.00000572	0.00000343	0.00000343	0.00000114	0.00000335
D-A	0.00033392	0.00037498	0.00018749	0.00018749	0.00006250	0.00018305
E-A	<0.00000001	<0.00000001	<0.00000001	<0.00000001	<0.00000001	<0.00000001
C-B	0.23762175	0.42841263	0.12852379	0.12852379	0.05355158	0.15685087
D-B	0.98848032	1.00000000	0.63514400	0.63514400	0.63514400	1.00000000
E-B	<0.00000001	<0.00000001	<0.00000001	<0.00000001	<0.00000001	<0.00000001
D-C	0.47910996	1.00000000	0.22143936	0.22143936	0.12302187	0.36032714
E-C	0.00000029	0.00000031	0.00000022	0.00000022	0.00000008	0.00000023
E-D	0.00000001	0.00000001	0.00000001	0.00000001	<0.00000001	0.00000001

8.1.3 Jednofaktorový ANOVA model s fixnými efektami pri rôznych rozptyloch

Nech $Y_{ji}(\mu_j, \sigma_j^2)$, kde existuje aspoň jedna dvojica $i \neq j$ taká, že $\sigma_i^2 \neq \sigma_j^2$ a zároveň σ_j^2 sú neznáme. Potom F_W štatistika nemá F rozdelenie s $J-1$ a $n-J$ stupňami voľnosti a musí byť modifikovaná nasledovne

$$F_W = \frac{\sum_{j=1}^J \widehat{w}_j \left(\overline{Y}_{j\cdot} - \overline{Y}_{\cdot\cdot}^{(w)}\right)^2}{J - 1 + \frac{J-2}{J+1}\sum_{j=1}^J \frac{\left(1-\widehat{h}_j\right)^2}{n_j-1}} \overset{\mathcal{D}}{\sim} F_{J-1, df_w},$$

$\widehat{w}_j = n_j/s_j^2$, $\widehat{h}_j = \frac{\widehat{w}_j}{\sum_{i=1}^J \widehat{w}_i}$, $j = 1, 2, \ldots, J$ a $\widehat{\mu} = \overline{Y}_{\cdot\cdot}^{(w)} = \sum_{j=1}^J \widehat{h}_j \overline{Y}_{j\cdot}$. Počet stupňov voľnosti

$$df_{w_1} = \frac{J^2 - 1}{3\sum_{j=1}^J \frac{\left(1-\widehat{h}_j\right)^2}{n_j-1}},$$

čo zaokrúhlime na najbližšie nižšie celé číslo, t.j. $df_w = \lfloor df_{w_1} \rfloor$. F_W sa nazýva **Fisherova testovacia štatistika** (alebo presnejšie **Welchova ANOVA F-štatistika**) a test **viacvýberový F-test s Welchovou aproximáciou stupňov voľnosti o rovnosti stredných hodnôt** $\mu_1, \mu_2, \ldots, \mu_J$ (alebo **Welchov ANOVA F-test**). Realizáciou F_W je F_{obs} a p-hodnota = $\Pr(F_W \geq F_{\text{obs}} | H_0)$. Na porovnanie ANOVA modelu pri rôznych rozptyloch s ANOVA modelom pri rovnakých rozptyloch – s^2 definujeme ako vážený priemer výberových rozptylov s_j^2, $j = 1, 2, \ldots, J$, teda

$$s^2 = \frac{\sum_{j=1}^J (n_j - 1)\, s_j^2}{\sum_{j=1}^J (n_j - 1)} = \frac{\sum_{j=1}^J (n_j - 1)\, s_j^2}{n - J}.$$

Potom $\mathbf{Y}_j \sim N_{n_j}(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_j)$, kde $\boldsymbol{\Sigma}_j = \sigma_j^2 \mathbf{I}_{n_j \times n_j}$, vektor chýb $\boldsymbol{\varepsilon}_j \sim N_{n_j}(\mathbf{0}, \boldsymbol{\Sigma}_j)$, vektor parametrov $\widehat{\boldsymbol{\beta}} \sim N_{J+1}(\boldsymbol{\beta}, (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1})$, kde $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_J^2)$ a maximálne vierohodný odhad $\widehat{\boldsymbol{\beta}}$ vypočítame pomocou **zovšeobecnenej metódy najmenších štvorcov**, t.j. $\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{Y}$.

Waldov $100 \times (1-\alpha)\%$ **empirický IS pre <u>nejakú</u> lineárnu kombináciu** $\sum_{j=1}^J a_j\mu_j$ (nazývaný aj **empirický IS Fisherovho typu**) ako

$$\left(\sum_{j=1}^J a_j\overline{y}_{j\cdot} - t_{df_w}(\alpha/2)\sqrt{\sum_{j=1}^J \frac{a_j^2 s_j^2}{n_j}}, \sum_{j=1}^J a_j\overline{y}_{j\cdot} + t_{df_w}(\alpha/2)\sqrt{\sum_{j=1}^J \frac{a_j^2 s_j^2}{n_j}}\right).$$

Waldove simultánne $100 \times (1-\alpha)\%$ **empirické intervaly spoľahlivosti Scheffeho typu** definujeme ako

$$\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - \sqrt{(J-1)F_{J-1,df_w}(\alpha)}\sqrt{\mathbf{a}^T(\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{a}}, \mathbf{a}^T\widehat{\boldsymbol{\mu}} + \sqrt{(J-1)F_{J-1,df_w}(\alpha)}\sqrt{\mathbf{a}^T(\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{a}}\right),$$

kde $\widehat{\mathbf{\Sigma}} = \mathbf{S} = \text{diag}(s_1^2, s_2^2, \ldots, s_J^2)$. **Waldove simultánne** $100 \times (1-\alpha)\%$ **empirické intervaly spoľahlivosti Tukeyho typu** definujeme ako

$$\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - q_{J,df_w}(\alpha)\sqrt{\frac{1}{2}\mathbf{a}^T(\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{a}}, \mathbf{a}^T\widehat{\boldsymbol{\mu}} + q_{J,df_w}(\alpha)\sqrt{\frac{1}{2}\mathbf{a}^T(\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{a}}\right).$$

Waldove simultánne $100 \times (1-\alpha)\%$ **empirické intervaly spoľahlivosti Fisherovho typu** definované nasledovne

$$\left(\mathbf{a}^T\widehat{\boldsymbol{\mu}} - t_{df_w}(\alpha_k)\sqrt{\mathbf{a}^T(\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{a}}, \mathbf{a}^T\widehat{\boldsymbol{\mu}} + t_{df_w}(\alpha_k)\sqrt{\mathbf{a}^T(\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{a}}\right).$$

ANOVA model v R

Funkciu oneway.test() je možné použiť aj v prípade, že rozptyly nie sú rovnaké, ak nastavíme argument var.equal=FALSE.

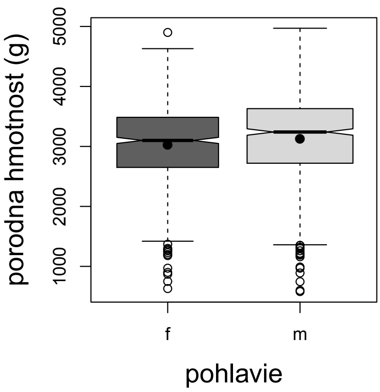
Mnohonásobné porovnávania v R

Na jednokrokové a viackrokové metódy (výpočet adjustovaných p-hodnôt) použijeme podobne ako predtým funkciu pairwise.t.test(y,x,p.adjust="metoda"), avšak argument pool.sd=FALSE.

Príklad 258 (Tukeyho HSD metóda) *Naprogramujte v R Tukeyho HSD metódu za predpokladu nerovnosti rozptylov.*

Príklad 259 (viacvýberový jednorozmerný prípad) *Majme dáta* anova-newborns.txt *a premennú pôrodná hmotnosť hodnoteného dieťaťa* weight.C *(v gramoch). V dáta sa nachádzajú odľahlé pozorovania, preto je potrebné ich pred výpočtami winsorizovať v každej skupine zvlášť (pozri kapitolu 3.1 a 3.2). Otestujte*
(a) sexuálny dimorfizmus (t.j. rozdiely stredných hodnôt – chlapci mínus dievčatá),
(b) sexuálny dimorfizmus u (1) prvorodených detí, (2) druhorodených detí a (3) detí narodených v treťom a ďalšom poradí, ako aj rozdiely medzi (1) a (2), (2) a (3) u každého pohlavia zvlášť (premenná prch.N*),*
(c) sexuálny dimorfizmus u detí, ktorých matky mali (1) základné vzdelanie, (2) vzdelanie stredné bez maturity, (3) vzdelanie stredné s maturitou a (4) vysokoškolské, ako aj rozdiely medzi (1) a (2), (2) a (3), (3) a (4) u každého pohlavia zvlášť (premenná edu.M*).*
Na testovanie použite
(1) dvojvýberové Studentove t-testy, kde $T_W \overset{\mathcal{D}}{\sim} t_{df}$*, (1.1)* $df = n_1 + n_2 - 2$ *a (1.2)* df *sú Welchove stupne voľnosti z dvojvýberového Studentovho t-testu s Welchovou aproximáciou stupňov voľnosti,*
(2) viacvýberové Fisherove LSD testy, kde $T_{LSD} \sim t_{df}$*, (2.1)* $df = n - J$ *a (2.2)* $df = df_w$ *sú Welchove stupne voľnosti z Welchovho ANOVA F-testu ((a)* $J = 2$ *a počet porovnávaní* $h = 1$*; (b)* $J = 6$ *a počet porovnávaní* $h = 7$*; (c)* $J = 8$ *a počet porovnávaní* $h = 10$*),*
(3) viacvýberové Tukeyho HSD testy pri homogenite a nehomogenite rozptylov ((a) $J = 2$*, (b)* $J = 6$ *a (c)* $J = 8$*).*
Ďalej vypočítajte
(4) Waldove $100 \times (1-\alpha)\%$ *empirické intervaly spoľahlivosti pre rozdiel stredných hodnôt s použitím kritických hodnôt* $t_{df}(\alpha/2)$*, kde* df *je z (1.1) a (1.2).*
(5) Waldove simultánne $100 \times (1-\alpha)\%$ *empirické intervaly spoľahlivosti Tukeyho typu pre rozdiel stredných hodnôt, kde* df *je z (2.1) a (2.2).*

Novorodenci mužského pohlavia majú systematicky o niečo vyššiu pôrodnú hmotnosť než novorodené dievčatá (259a, pozri graf 8.9; e.g. Kumar a kol., 2013). Súčasne existujú nezanedbateľné rozdiely v rade znakov medzi súrodencami narodenými v rôznom poradí. Napríklad prvorodení chlapci aj dievčatá majú v pupočnej krvi výrazne vyššie hladiny pohlavných hormónov než novorodenci narodení vo vyššom poradí (Maccoby a kol., 1979). Čo sa týka pôrodnej hmotnosti, prvorodení mávajú o niečo nižšiu pôrodnú hmotnosť než druhorodení. Rozdiely medzi pohlaviami sa vysvetľujú špecifickými faktormi z pohlavných chromozómov a odlišnými hladinami pohlavných hormónov, ktoré ovplyvňujú prenatálnu diferenciáciu a rast celého tela. Rozdiely medzi novorodencami narodenými v rôznom poradí sa hľadajú v mechanizmoch (imunitná interakcia matky s predchádzajúcimi potomkami, stres matky so zvyšujúcou sa paritou), ktoré cez matku ovplyvňujú embryo a plod v jej maternici (e.g. Bogaert a Skorska, 2011). V našom prípade máme k dispozícii pôrodnú hmotnosť (premenná weight.C) novorodených dievčat a chlapcov a chceme zistiť, ako tieto dva faktory (pohlavie a poradie narodenia) ovplyvňujú pôrodnú hmotnosť. Chceme vedieť, či sa líšia novorodenci narodení v inom poradí, u každého pohlavia zvlášť, v pôrodnej hmotnosti (premenná prch.N, f-1 prvorodené dievčatá, f-2 druhorodené dievčatá, f-3 dievčatá narodené v treťom a vyššom poradí; obdobne chlapci); a ďalej či existuje sexuálny dimorfizmus v pôrodnej hmotnosti v rámci každého poradia narodenia zvlášť (príklad 259b; obrázok 8.12).

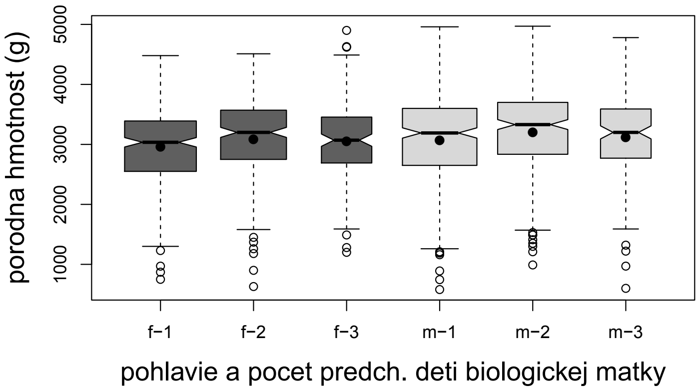


Obr. 8.9: Krabicové diagramy pôrodnej hmotnosti pre dievčatá a chlapcov

Celkovo vidíme väčšie rozptyly u chlapcov (pozri obrázok 8.10, tabuľku rozptylov neuvádzame), čo môže odpovedať faktu, že samčie pohlavie je všeobecne variabilnejšie. Preto je v tomto prípade potrebné vopred počítať s nehomogenitou rozptylov. Na základe krabicových diagramov (graf 8.10) a aritmetických priemerov je možné konštatovať, že dimorfizmus by mohol byť štatisticky významný až vo vyššom poradí. Organizmus matky je pravdepodobne plne prispôsobený reprodukčnej funkcii až pri druhom a treťom poradí. Podľa niektorých štúdií však poradie narodenia nie je jediným faktorom v hre, rolu hrá tiež vek matky (súvisí so vzdelaním – matky s vyšším vzdelaním sa v našej populácii v priemere začínajú reprodukovať neskôr; vzdelanie matiek je výrazným faktorom reprodukcie, pokiaľ sú matky stále vo vzdelávacom systéme, po jeho opustení už menej; Šťastná, 2010) a medzipôrodný interval; tehotenstvo po príliš krátkom medzipôrodnom intervale je stresujúce, po extrémne dlhom medzipôrodnom intervale sa môže fyziologicky blížiť situácii prvého tehotenstva (prehľad Rousso a kol., 2002). Pokiaľ budú väčšie rozptyly u jedného pohlavia než u druhého, toto pohlavie môže byť citlivejšie na niektorý z uvedených prenatálnych faktorov. Ešte výraznejším faktorom však môže byť skrytá heterogenita v dátach, spôsobená sociálnymi a etnickými rozdielmi. Pokiaľ sú v dátach skryté rôzne sociálne/etnické skupiny líšiace sa priemernou pôrodnou hmotnosťou a súčasne majúce odlišnú normu v počte detí v rodine (ktorý môže negatívne súvisieť i so vzdelaním matky), bude v rôznom poradí narodenia rôzne zastúpenie každej z týchto skupín, čo ovplyvní porovnávané stredné hodnoty a pravdepodobne aj rozptyly. Rozsiahle demografické štúdie pôrodnej hmotnosti založené na vzorcoch z moderných mestských populácií s veľkou pravdepodobnosťou (nekontrolovaného) mnohoetnického zloženia sú týmto faktorom pravdepodobne vždy ovplyvnené.

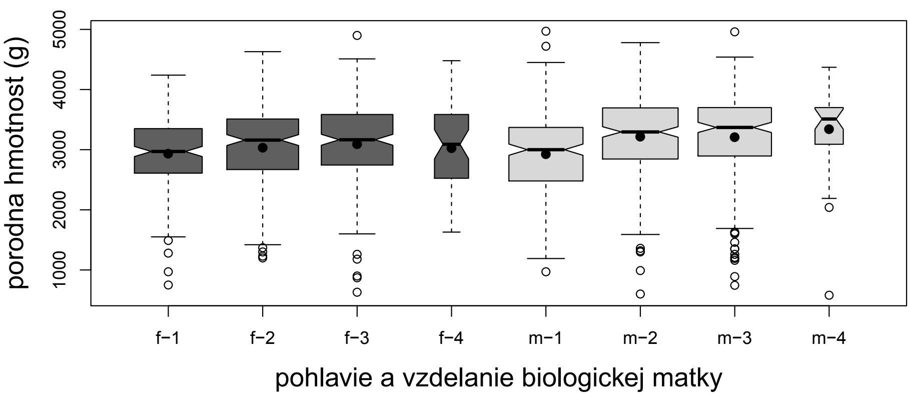
Ďalej chceme zistiť, či existujú rozdiely v pôrodnej hmotnosti (premenná weight.C) medzi deťmi narodenými matkám s rôznym stupňom vzdelania (premenná edu.M, pre novorodené dievčatá: f-1 základné, f-2 stredné bez maturity, f-3 stredné s maturitou a f-4 vysokoškolské; obdobne m-1, m-2, m-3 a m-4 pre chlapcov; príklad 259c; obrázok 8.12).

Biologické rozdiely medzi novorodencami jednotlivých skupín môžu byť dané napr. neskorším začiatkom reprodukcie u žien s vyšším vzdelaním (pozri graf 8.11). Vzhľadom na to, že dosiahnuté vzdelanie môže súvisieť aj s celkovým počtom detí (a dáta obsahujú ženy rôzneho veku, parity, atď.), je možné, že dve vyššie kategórie zahŕňajú vyššiu proporciu prvorodených, zatiaľ čo dve nižšie kategórie vyššiu proporciu detí narodených vo vyššom poradí (a tak hmotnejšiu než pri zložení odpovedajúcemu vyšším dvom kategóriám). Bolo by preto vhodné sledovať vplyv všetkých troch faktorov súčasne (pohlavie,



Obr. 8.10: Krabicové diagramy pôrodnej hmotnosti pre dievčatá a chlapcov podľa počtu predchádzajúcich detí biologickej matky

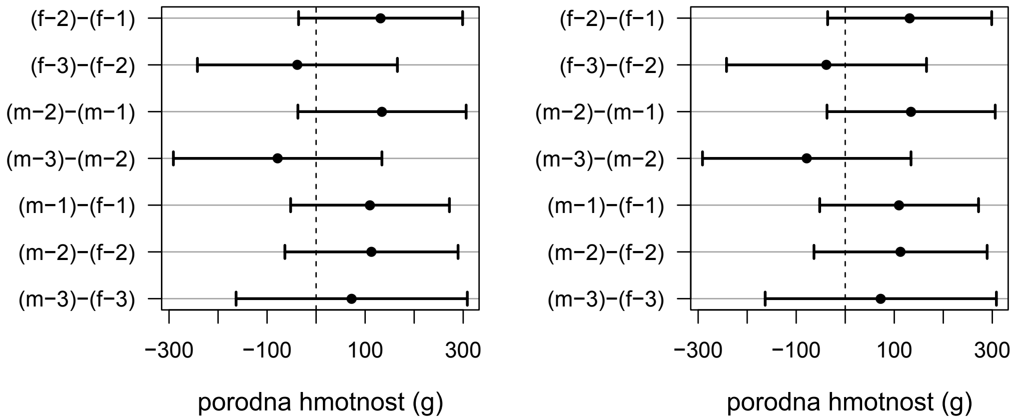
poradie narodenia, vzdelanie matky). Pokiaľ existuje nejaká súvislosť medzi dosiahnutým vzdelaním matky a počtom detí, a súčasne sa skryté sociálne/etnické skupiny líšia v počte detí i pôrodnej hmotnosti, je pravdepodobné, že výsledky budú ovplyvnené tiež sociálnym/etnickým zložením vzorky, ktoré ale v našich dátach nemáme pod kontrolou. Výsledky potom môžu byť rôzne ovplyvnené počtom detí, vzdelaním a priemernou hmotnosťou novorodencov v jednotlivých sociálnych/etnických skupinách.



Obr. 8.11: Krabicové diagramy pôrodnej hmotnosti pre dievčatá a chlapcov podľa vzdelania biologickej matky

Post-hoc testy ukazujú (pozri obrázok 8.12), že nemôžeme zamietnuť nulovú hypotézu v žiadnej porovnávanej dvojici populácií ani pri predpoklade homogenity rozptylov a ani pri nehomogenite rozptylov. Prejavuje sa len predpokladaná tendencia k sexuálnemu dimorfizmu s vyššími hodnotami u chlapcov a u oboch pohlaví tendencia k vyššej hmotnosti u druhorodených oproti prvorodeným, a tiež naopak tendencia k nižšej hmotnosti u detí vo vyššom poradí oproti druhorodeným.

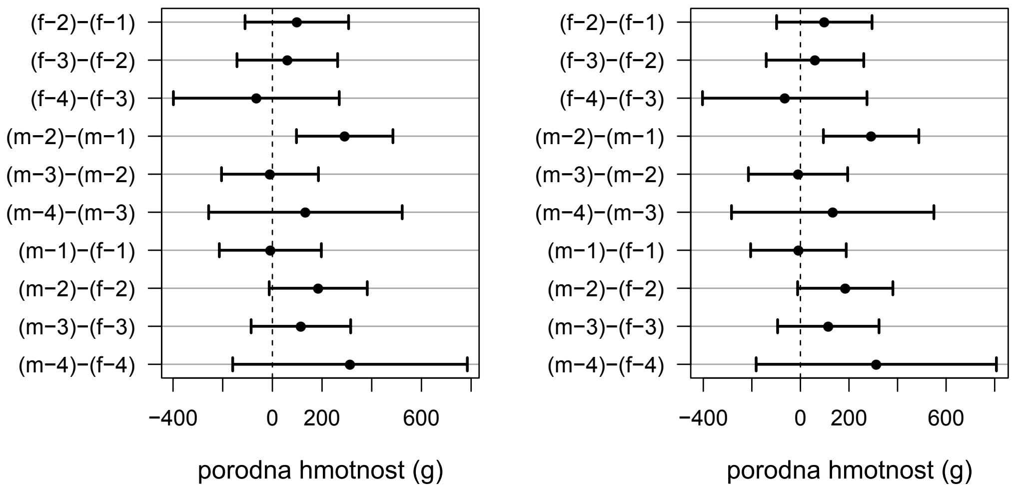
Jediný štatisticky významný rozdiel (pozri obrázok 8.13) bol zistený medzi synmi matiek so základným a stredným vzdelaním bez maturity (m-2 vs m-1), kedy vyššia priemerná hmotnosť bola zaznamenaná u matiek s vyšším vzdelaním. Celkovo je však z porovnania susedných skupín podľa vzdelania matiek v rámci pohlavia zrejmé, že pôrodná hmotnosť buď nesúvisí so vzdelaním matiek vôbec, alebo vidíme aspoň tendenciu k vyššej pôrodnej hmotnosti u novorodencov matiek so vzdelaním o stupeň vyšším (4 zo 6 rozdielov sú pozitívne). Čo sa týka sexuálneho dimorfizmu v žiadnej kategórii podľa vzdelania matiek nebola pôrodná hmotnosť štatisticky významne rozdielna medzi pohlaviami, aj keď je viditeľný trend k vyššej hodnote u chlapcov. Výnimkou sú deti matiek s najnižším vzdelaním, kde nie je ani žiadna tendencia k rozdielu. Detailné vysvetlenie výsledkov by bolo možné pri doplnení informácií v zmysle vyššie uvedených súvislostí s etnickým/sociálnym zložením vzorky.



Obr. 8.12: Waldove simultánne 95% empirické intervaly spoľahlivosti Tukeyho typu pre rozdiely stredných hodnôt pôrodnej hmotnosti (pohlavie dieťaťa vs počet predchádzajúcich detí biologickej matky) – za predpokladu homogenity rozptylov (vľavo) a nehomogenity rozptylov (vpravo); rozdiely v dĺžke toho istého intervalu spoľahlivosti pri rôznych predpokladoch sa líšia od 0.63 po 14.52 v závislosti od odchýlky hodnôt rozptylov použitých na ich výpočet a od rozdielov stupňov voľnosti df_w (newinsorizované dáta)

Príklad 260 (dvojvýberový mnohorozmerný prípad) *Majme dáta* anova-head.txt *a premenné (1) dĺžka hlavy* head.L*, (2) šírka hlavy* head.W*, (3) šírka dolnej čeľuste* bigo.W *a (4) šírka tváre* bizyg.W *(všetky v milimetroch). V dátach sa nachádzajú odľahlé pozorovania, preto je potrebné ich pred výpočtami winsorizovať v každej skupine zvlášť (pozri kapitolu 3.1 a 3.2). Otestujte sexuálny dimorfizmus (t.j. rozdiely stredných hodnôt premenných (1) až (4) – muži mínus ženy; premenná* sex*) nasledovne:*
(a) použite štyri dvojvýberové Studentove t-testy (nezávisle pre každú premennú), kde $T_W \overset{\mathcal{D}}{\sim} t_{df}$*, (1)* $df = n_1 + n_2 - 2$ *a (2) df sú Welchove stupne voľnosti z dvojvýberového Studentovho t-testu s Welchovou aproximáciou stupňov voľnosti,*
(b) použite štyri viacvýberové Fisherove LSD testy, kde $T_{LSD} \sim t_{df}$*, (1)* $df = n - J$ *a (2)* $df = df_w$ *sú Welchove stupne voľnosti z Welchovho ANOVA F-testu (*$J = 2$ *a počet porovnávaní* $h = 4 \times \frac{J(J-1)}{2} = 4$*).*
Vypočítajte
(c) Waldove $100 \times (1-\alpha)\%$ *empirické intervaly spoľahlivosti pre rozdiel stredných hodnôt s použitím kritických hodnôt* $t_{df}(\alpha/2)$*, kde df je z (1) a (2) z (a).*
(d) Waldove simultánne $100 \times (1-\alpha)\%$ *empirické intervaly spoľahlivosti Fisherovho typu pre rozdiel stredných hodnôt s použitím Benjamini-Yekutieliho kritických hodnôt* $t_{df}(\alpha_k)$*, kde* $\alpha_k = 1, 2, \ldots, h$*, df je z (1) a (2) z (b).*
(f) Waldove simultánne $100 \times (1-\alpha)\%$ *empirické intervaly spoľahlivosti Tukeyho typu pre rozdiel stredných hodnôt, kde* $J = 4$*.*

U človeka existuje výrazný dimorfizmus vo výške postavy (veľkosti tela), ktorý sa ale z veľkej časti utvára v puberte a adolescencii pod vplyvom odlišných hladín gonadálnych steroidných hormónov. Nervový systém a s ním i skelet neurokránia naopak rastie najvýraznejšie v období pred pubertou. Spolu s výrazne sa zvyšujúcimi priemernými hodnotami u mužov však môže v puberte a adolescencii narastať medzi mužmi viac než medzi ženami tiež variabilita, napríklad v prípade kľúčnej kosti môže byť rozsah hodnôt niektorých veľkostných ukazovateľov takmer dvojnásobný oproti ženám (e.g. Singh

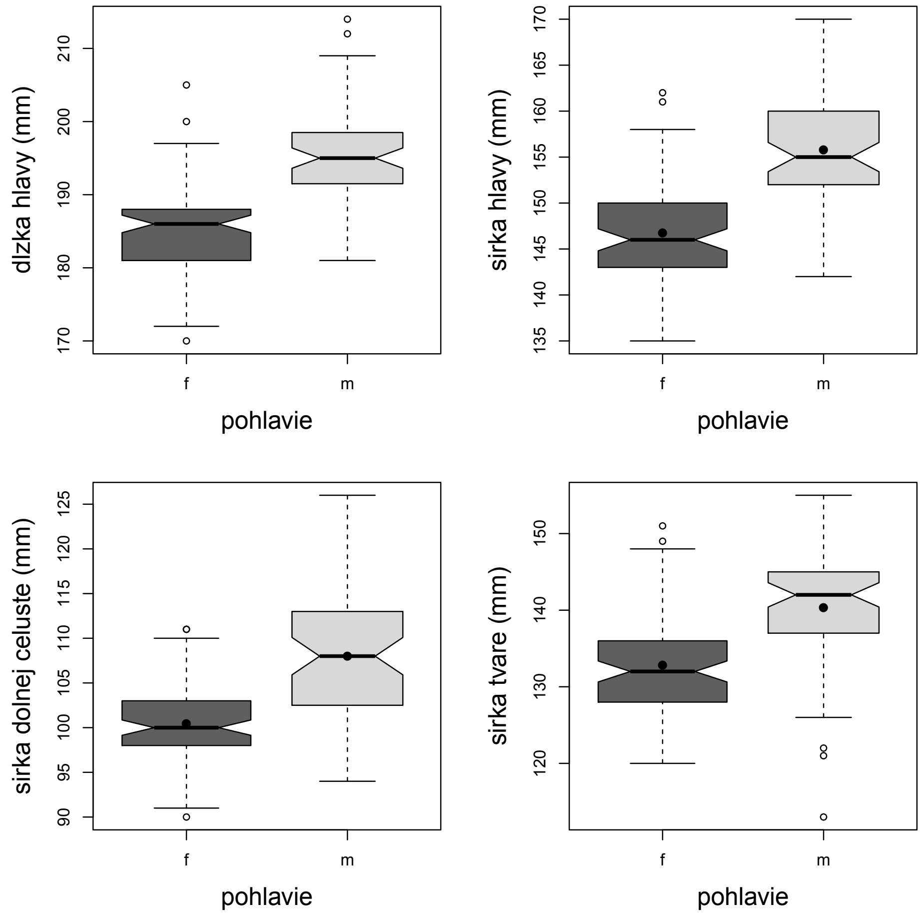


Obr. 8.13: Waldove simultánne 95% empirické intervaly spoľahlivosti Tukeyho typu pre rozdiely stredných hodnôt pôrodnej hmotnosti (pohlavie dieťaťa vs vzdelanie biologickej matky) – za predpokladu homogenity rozptylov (vľavo) a nehomogenity rozptylov (vpravo); rozdiely v dĺžke toho istého intervalu spoľahlivosti pri rôznych predpokladoch sa líšia od 2.51 po 54.40 v závislosti od odchýlky hodnôt rozptylov použitých na ich výpočet a od rozdielov stupňov voľnosti df_w (newinsorizované dáta)

a Jit, 1996). V rozmeroch lebky by tak vo výsledku mohli byť medzipohlavné rozdiely percentuálne väčšie v znakoch rastúcich prevažne v období puberty (t.j. skôr rozmery tvárovej časti, v dolnej polovici tváre), ale vzhľadom na rastúci rozptyl by to nemuselo mať vplyv na veľkosť efektu medzipohlavných rozdielov (štandardizovaný rozdiel stredných hodnôt, vztiahnutý na združenú smerodajnú odchýlku a teda aj štatistickú významnosť testu rozdielu stredných hodnôt).

Krabicové diagramy (pozri obrázok 8.14) naznačujú, že vo všetkých štyroch rozmeroch sa pravdepodobne bude nachádzať podľa predpokladu sexuálny dimorfizmus. Grafy tiež súčasne poukazujú na to, že v šírke dolnej čeľuste majú muži (vizuálne posudzované) skutočne väčší rozptyl než ženy, čo opäť odpovedá predpokladu väčšej variability pubertálneho/adolescentného rastu medzi mužmi než medzi ženami.

Príklad 261 (viacvýberový mnohorozmerný prípad) *Majme dáta* anova-head.txt *a premenné (1) dĺžka hlavy* head.L*, (2) šírka hlavy* head.W*, (3) šírka dolnej čeľuste* bigo.W *a (4) šírka tváre* bizyg.W *(všetky v milimetroch). V dátach sa nachádzajú odľahlé pozorovania, preto je potrebné ich pred výpočtami winsorizovať v každej skupine zvlášť (pozri kapitolu 3.1 a 3.2). Otestujte rozdiely v stredných hodnotách medzi (premenné* sex *a* sexor*)*
- f-op *– ženami so sexuálnou orientáciou výlučne na opačné pohlavie,*
- f-sa *– ženami s bisexuálnou alebo homosexuálnou orientáciou,*
- m-op *– mužmi so sexuálnou orientáciou výlučne na opačné pohlavie a*
- m-sa *– mužmi s bisexuálnou alebo homosexuálnou orientáciou.*
Testovanie vykonajte nasledovne:
(a) použite 24 dvojvýberových Studentových t-testov (nezávisle pre každú premennú), kde $T_W \overset{\mathcal{D}}{\sim} t_{df}$*, (1)* $df = n_1 + n_2 - 2$ *a (2) df sú Welchove stupne voľnosti z dvojvýberového Studentovho t-testu s Welchovou aproximáciou stupňov voľnosti,*
(b) použite 24 viacvýberových Fisherových LSD testov, kde $T_{LSD} \sim t_{df}$*, (1)* $df = n - 4$ *a (2)* $df = df_w$ *sú Welchove stupne voľnosti z Welchovho ANOVA F-testu (*$J = 4$*, počet porovnávaní* $h = 4 \times \frac{J(J-1)}{2} = 4 \times 6 = 24$*),*
(c) použite (1) ANOVA F-test a (2) Welchov ANOVA F-test pre všetky štyri premenné.

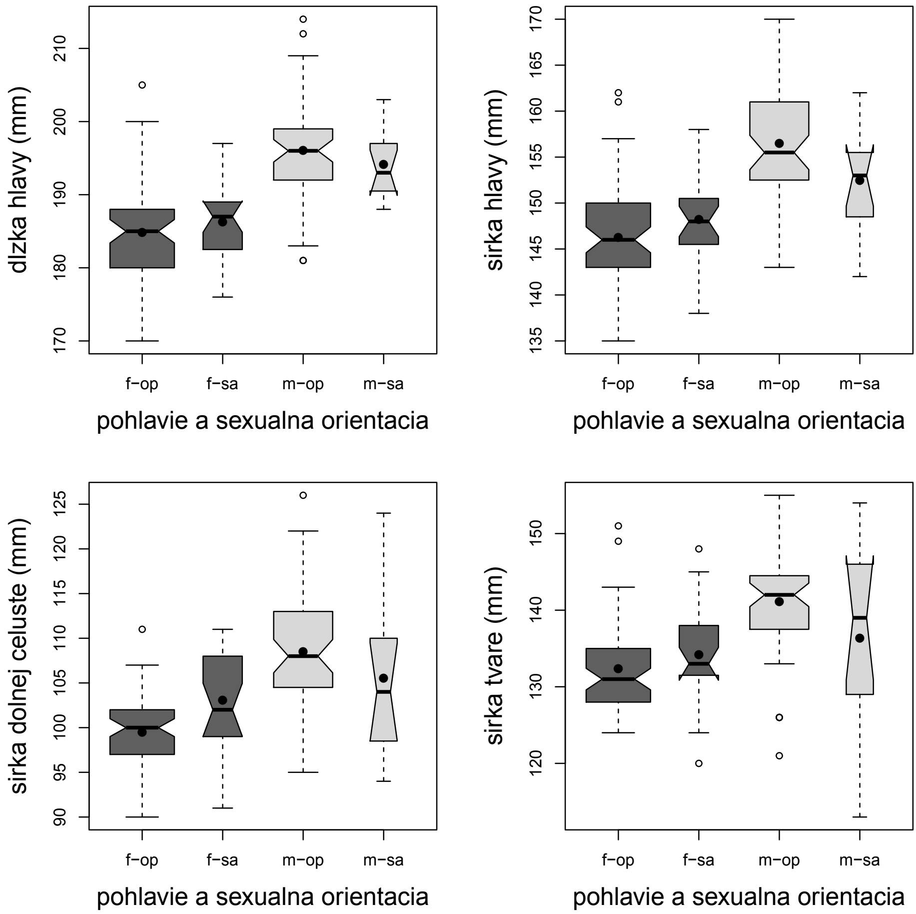


Obr. 8.14: Krabicové diagramy pre ženy a mužov – dĺžka hlavy (vľavo hore), šírka dolnej čeľuste (vľavo dole), šírka hlavy (vpravo hore) a šírka tváre (vpravo dole)

Vypočítajte
(d) Waldove $100 \times (1 - \alpha)\%$ *empirické intervaly spoľahlivosti pre rozdiel stredných hodnôt s použitím kritických hodnôt* $t_{df}(\alpha/2)$, *kde* df *je z (1) a (2) z (a).*
(e) Waldove simultánne $100 \times (1 - \alpha)\%$ *empirické intervaly spoľahlivosti Fisherovho typu pre rozdiel stredných hodnôt s použitím Benjamini-Yekutieliho kritických hodnôt* $t_{df}(\alpha_k)$, *kde* $\alpha_k = 1, 2, \ldots, h$, df *je z (1) a (2) z (b).*
(f) Waldove simultánne $100 \times (1 - \alpha)\%$ *empirické intervaly spoľahlivosti Tukeyho typu pre rozdiel stredných hodnôt, kde* $J = 4$.

Okrem sexuálneho dimorfizmu, t.j. rozdielov definovaných biologickým pohlavím (založených na odlišnej kombinácii pohlavných chromozómov a odlišných hladinách pohlavných hormónov prenatálne aj v puberte) ale existujú rozdiely spojené so *sexuálnou orientáciou* (pozri obrázok 8.15). Napriek tomu, že presný mechanizmus vzniku rozdielov v sexuálnej orientácii nie je doteraz známy, v súčasnej dobe sa teórie i nepriame empirické dôkazy klonia na stranu biologických hypotéz a k sexuálnej orientácii nezmeniteľne determinovanej už v prenatálnom období. Z odlišných frekvencií a ďalších súvislostí sexuálnej orientácie u mužov a u žien je zrejmé, že mechanizmy nastavenia finálnej sexuálnej orientácie môžu byť u mužov a žien odlišné. U žien (veľmi zjednodušene povedané) zaznamenané dôkazy (napr. vrodená adrenálna hyperplázia) nasvedčujú tomu, že na vznik ženskej homosexuality vplýva nerovnováha (zvýšené hladiny) prenatálnych androgénov. U mužov sú v tomto ohľade dôkazy ešte nejednoznačnejšie. Boli však nájdené dôkazy naznačujúce, že sexuálna orientácia mužov súvisí s dedičnými faktormi na X–chromozóme od matky. Je možné, že variant niektorého (zatiaľ hypotetického) génu na X–chromozóme je reprodukčne veľmi výhodný, pokiaľ sa nachádza v ženskom tele (t.j. zvyšuje reprodukčnú úspešnosť žien – nositeliek, čím prispieva k udržaniu svojej frekvencie v populácii), ale feminizuje telo a mozog mužských nositeľov, ak ho od matky zdedí syn (získa práve ten jej X–chromozóm, ktorý je nositeľom tejto výhodnej formy). Nepriame dôkazy (LeVay, 1991) nasvedčujú, že niektoré hypotalamické jadrá mužských homosexuálov majú feminizované kvantitatívne vlastnosti, iné štúdie tieto výsledky spochybňujú (Swaab a kol., 1992) a poukazujú na metodicky diskutabilné vzorkovanie a vzťah medzi zistenými asociáciami (súvislosťami, vzťahmi, koreláciami) a predpokladanými kauzálnymi závislosťami. Čo sa týka telesnej stavby, ako u ženských, tak aj u mužských homosexuálov nájdeme feminínne a maskulínne telesné formy (ženy: *butch* a *femme lesbians*). Obmedzené množstvo dostupných štúdií ukazuje u oboch pohlaví zmes maskulínnejších i feminínnejších foriem znakov oproti heterosexuálnym porovnávacím populáciám. Prevaha výsledkov však naznačuje, že feminínnejšie znaky u mužských homosexuálov a maskulínnejšie znaky u ženských homosexuálov celkovo prevládajú, t.j. homosexuálni jedinci sa zdajú byť v určitej miere posunutí smerom k norme opačného pohlavia (napr. Nedoma a Freund, 1961; Bogaert a Blanchard, 1996; Weill, 2009). V našom prípade chceme zistiť, či okrem pohlavia sledované 4 rozmery hlavy súvisia tiež s pohlavnou orientáciou. Ak platí vyššie uvedený predpoklad, dimorfizmus u heterosexuálnych jedincov by mal byť väčší než u jedincov homosexuálnych. Vzhľadom na to, že pohlavná orientácia by sa mala zakladať už prenatálne, zistené rozdiely by mohli mať raný ontogenetický pôvod (pokiaľ nie je po narodení rast týchto telesných partií ovplyvnený externými faktormi inak u /neskôr v dospelosti/ homosexuálnych a heterosexuálnych jedincov).

Výsledky (obrázok 8.16) sú do značnej miery v zhode s vyššie uvedenými predpokladmi. Vo všetkých štyroch rozmeroch je štatisticky významný dimorfizmus a číselne najvyšší medzipohlavný rozdiel nájdený medzi heterosexuálnymi mužmi a ženami. Všetky ostatné medzipohlavné rozdiely (heterosexuálni muži vs. homosexuálne ženy, heterosexuálne ženy vs. homosexuálni muži, homosexuálni muži vs. homosexuálne ženy) sú celkovo číselne menšie a v 4 z 12 týchto porovnaní nie sú štatisticky významné (čo sa ale dá očakávať, vzhľadom na početne menšiu vzorku homosexuálov). Rozdiely medzi homosexuálmi a heterosexuálmi v rámci každého pohlavia sú celkovo malé a štatisticky nevýznamné (až na hranične významný rozdiel v šírke dolnej čeľuste mezi homosexuálnymi a heterosexuálnymi ženami, kedy čeľusť homosexuálnych žien je širšia), aj tu platí vyššie uvedený komentár k veľkosti vzorky. Evidentne je však vo všetkých rozmeroch viditeľný trend k posunu rozdielu v zmysle posunu hodnôt homosexuálnej vzorky k hodnotám obvyklejším u opačného pohlavia (maskulinizácia homosexuálnych žien a feminizácia homosexuálnych mužov). Výsledky sa podobajú výsledkom zistených v iných antropometrických premenných (porovnanie Weill, 2009). Tento trend sa zdá najvýraznejší v šírke čeľusti u žien a šírke tváre u mužov. Výsledky teda naznačujú, že faktor sexuálnej orientácie súvisí viac s rastom tváre než mozgovne. Vzhľadom na výraznejší rast tejto časti hlavy v puberte a adolescencii možno usudzovať na nejakú prenatálnu súvislosť medzi sexuálnou orientáciou a pubertálnym/adolescentným rastom. U oboch rozmerov tváre súčasne vidíme (obrázok 8.15, i keď opäť s rezervou vzhľadom na veľkosť vzorky), že rozptyl vo vzorke homosexuálov sa zdá byť väčší, t.j. akoby zahŕňal ako formy obvyklé pre dané pohlavie, tak aj formy posunuté smerom k opačnému pohlaviu (v dôsledku čoho je však posunutá v priemere celá skupina). Obmedzením testovanej vzorky sú iste výrazne odlišné počty homosexuálnych a heterosexuálnych jedincov, predovšetkým malý počet prípadov vzorky homosexuálov. Na druhej strane však celé dáta predstavujú prirodzenú skupinu mladých dospelých ľudí, väčšinou vysokoškolských študentov, vrátane prirodzeného zastúpenia ne–heterosexuálnych jedincov, bez akýchkoľvek apriorných vplyvov sexuálnej orientácie na vznik vzorky (vrátane anonymného dotazníka sexuálnej orientácie). Pokiaľ by mal mať výber štatisticky príhodné vlastnosti (predovšetkým veľkosť), bolo by nutné zámerne hľadať jedincov oboch skupín a sexuálnu orientáciu dopredu zisťovať. Ďalej by sa museli zaistiť rovnaké kritéria výberu u oboch skupín z hľadiska sledovaných rozmerov (čo vzhľadom na celý rad (sub)kultúrnych a iných rozdielov nemusí byť práve triviálna úloha). Nemenej podstatnou oblasťou metodiky je samotná klasifikácia sexuálnej orientácie. Zlúčenie jedincov všetkých 6 kategórií Kinseyho škály (Kinsey a kol., 1948) s nejednoznačne heterosexuálnou orientáciou (všetci mimo výlučne heterosexuálnych jedincov) do jednej kategórie ne–heterosexuáli (pozri popis dát `anova-head.txt`) môže zakrývať ďalšiu variabilitu v rámci našej početne malej „homosexuálnej“ skupiny.



Obr. 8.15: Krabicové diagramy pre ženy a mužov podľa sexuálnej orientácie – dĺžka hlavy (vľavo hore), šírka dolnej čeľuste (vľavo dole), šírka hlavy (vpravo hore) a šírka tváre (vpravo dole)

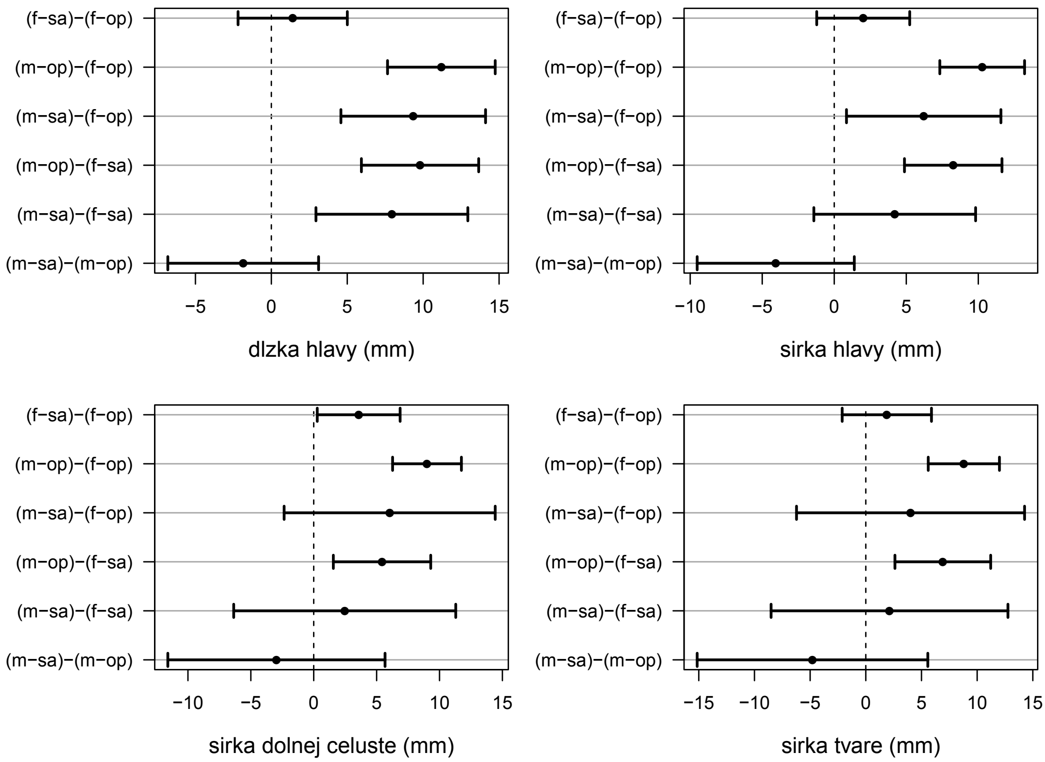
8.2 Asymptotické testy o rozptyloch

V tejto kapitole sa budeme zaoberať testami o homogenite (rovnosti) viacerých rozptylov. Najprv odvodíme testovaciu štatistiku pomerom vierohodnosti a následne Bartlettovu modifikáciu tejto štatistiky (Bartlett, 1937). Hypotézy definujeme nasledovne $H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_J^2 = \sigma^2$ oproti $H_1 : \sigma_i^2 \neq \sigma_j^2 \neq \sigma^2$ pre aspoň jedno $i < j, i = 1, 2, \ldots, J-1; j = 2, 3, \ldots, J$.

Test pomerom vierohodnosti o homogenite rozptylov.

Nech $Y_j \sim N(\mu_j, \sigma_j^2)$, kde $j = 1, 2, \ldots J$, $\boldsymbol{\theta} = (\mu_1, \mu_2, \ldots, \mu_J, \sigma_1^2, \sigma_2^2, \ldots, \sigma_J^2)^T$. Logaritmus funkcie vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_J) \quad = \quad -\sum_{j=1}^{J} \frac{n_j}{2} \ln(2\pi\sigma_j^2) - \sum_{j=1}^{J} \frac{1}{2\sigma_j^2} \left(\sum_{i=1}^{n_j} (y_{ji} - \mu_j)^2 \right).$$



Obr. 8.16: Waldove simultánne 95% empirické intervaly spoľahlivosti Tukeyho typu pre rozdiely stredných hodnôt – dĺžka hlavy (vľavo hore), šírka dolnej čeľuste (vľavo dole), šírka hlavy (vpravo hore) a šírka tváre (vpravo dole); newinsorizované dáta

MLE $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \mu_J, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2, \ldots, \widehat{\sigma}_J^2)^T$, kde

$$\widehat{\mu}_j = \overline{y}_j, \widehat{\sigma}_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} (y_{ji} - \overline{y}_j)^2,$$

t.j. $\Theta_1 = \{\boldsymbol{\theta} : \sigma_i^2 \neq \sigma_j^2 \neq \sigma^2; i < j, i = 1, 2, \ldots, J-1; j = 2, 3, \ldots, J\}$ pre aspoň jedno i a j. Za platnosti H_0 je $\boldsymbol{\theta}_0 = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_J, \widetilde{\sigma}^2, \widetilde{\sigma}^2, \ldots, \widetilde{\sigma}^2)^T$, kde (pozri Mood a kol., 1987)

$$\widetilde{\sigma}^2 = \frac{\sum_{j=1}^J n_j \widehat{\sigma}_j^2}{\sum_{j=1}^J n_j},$$

t.j. $\Theta_0 = \{\boldsymbol{\theta} : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_J^2 = \sigma^2\}$. Logaritmus funkcie vierohodnosti pre $\widehat{\boldsymbol{\theta}}$ bude mať tvar

$$\begin{aligned} l(\widehat{\boldsymbol{\theta}} | \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_J) &= -\sum_{j=1}^J \frac{n_j}{2} \ln(2\pi\widehat{\sigma}_j^2) - \sum_{j=1}^J \frac{1}{2\widehat{\sigma}_j^2} \left(\sum_{i=1}^{n_j} (y_{ji} - \mu_j)^2 \right) \\ &= -\sum_{j=1}^J \frac{n_j}{2} \ln(2\pi\widehat{\sigma}_j^2) - \sum_{j=1}^J \frac{n_j \widehat{\sigma}_j^2}{2\widehat{\sigma}_j^2}. \end{aligned}$$

Logaritmus funkcie vierohodnosti pre $\boldsymbol{\theta}_0$ bude mať tvar

$$\begin{aligned} l(\boldsymbol{\theta}_0|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_J) &= -\sum_{j=1}^{J}\frac{n_j}{2}\ln(2\pi\widetilde{\sigma}^2)-\sum_{j=1}^{J}\frac{1}{2\widetilde{\sigma}^2}\left(\sum_{i=1}^{n_j}(y_{ji}-\mu_j)^2\right) \\ &= -\sum_{j=1}^{J}\frac{n_j}{2}\ln(2\pi\widetilde{\sigma}^2)-\frac{1}{\widetilde{\sigma}^2}\sum_{j=1}^{J}\frac{n_j\widehat{\sigma}_j^2}{2} \end{aligned}$$

Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_J)) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_J)-l\left(\boldsymbol{\theta}_0|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_J\right) \\ &= -\sum_{j=1}^{J}\frac{n_j}{2}\ln\widehat{\sigma}_j^2-\sum_{j=1}^{J}\frac{n_j}{2}+\sum_{j=1}^{J}\frac{n_j}{2}\ln\left(\frac{\sum_{j=1}^{J}n_j\widehat{\sigma}_j^2}{\sum_{j=1}^{J}n_j}\right)+\sum_{j=1}^{J}\frac{n_j}{2} \\ &= \frac{1}{2}\sum_{j=1}^{J}n_j\ln\left(\frac{\sum_{j=1}^{J}n_j\widehat{\sigma}_j^2}{\widehat{\sigma}_j^2\sum_{j=1}^{J}n_j}\right)=\frac{1}{2}\sum_{j=1}^{J}n_j\ln\left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}_j^2}\right). \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}}=-2\ln\left(\lambda\left(\mathbf{Y}_1,\mathbf{Y}_2,\ldots,\mathbf{Y}_J\right)\right)\overset{\mathcal{D}}{\sim}\chi^2_{J-1}$.

Nech $J=2$ a $n_1=n_2$. Nech $F_{\mathrm{obs}}=\frac{\sum_{i=1}^{n}(y_{1i}-\overline{y}_1)^2}{\sum_{i=1}^{n}(y_{2i}-\overline{y}_2)^2}$, čo je špeciálny prípad F_{obs} z kapitoly 7.2. Potom

$$\frac{\widetilde{\sigma}^2}{\widehat{\sigma}_1^2}=\frac{1}{2}\left(1+\frac{1}{F_{\mathrm{obs}}}\right)\ \text{a}\ \frac{\widetilde{\sigma}^2}{\widehat{\sigma}_2^2}=\frac{1}{2}\left(1+F_{\mathrm{obs}}\right)$$

a testovacia štatistika pomerom vierohodnosti

$$u_{\mathrm{LR}}=n\ln\left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}_1^2}\right)+n\left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}_2^2}\right)=2n\ln\frac{1}{2}+n\left(\ln\left(1+\frac{1}{F_{\mathrm{obs}}}\right)+\ln\left(1+F_{\mathrm{obs}}\right)\right),$$

čo je veľké číslo, ak $1/F_{\mathrm{obs}}$ alebo F_{obs} je veľké, t.j. F_{obs} je malé alebo veľké. Prvá derivácia U_{LR} podľa F je rovná $\frac{n}{1+F}\left(-\frac{1}{F}+1\right)$, čo je rovné nule, ak $F_{\mathrm{obs}}=1$. Keďže funkcia je rastúca pre malé a veľké F_{obs}, dosahuje u_{LR} v bode $F_{\mathrm{obs}}=1$ minimum. Ak $F_{\mathrm{obs}}=0$, potom $u_{\mathrm{LR}}=0$.

Bartlett (1937) modifikoval testovaciu štatistiku pomerom vierohodnosti nasledovne

$$U_{\mathrm{B}}=\frac{U_{\mathrm{LR}}^{(\mathrm{alt})}}{C}\overset{\mathcal{D}}{\sim}\chi^2_{J-1},$$

kde

$$U_{\mathrm{LR}}^{(\mathrm{alt})}=\sum_{j=1}^{J}(n_j-1)\ln\left(\frac{\widetilde{S}^2}{S_j^2}\right),\widetilde{s}^2=\frac{\sum_{j=1}^{J}\left(n_j-1\right)S_j^2}{\sum_{j=1}^{J}\left(n_j-1\right)},$$

S_j^2 sú výberové rozptyly a

$$C=1+\frac{1}{3(J-1)}\left(\sum_{j=1}^{J}\frac{1}{n_j-1}-\frac{1}{\sum_{j=1}^{J}\left(n_j-1\right)}\right).$$

U_{B} konverguje ku χ^2_{J-1} rozdeleniu rýchlejšie ako U_{LR}.

Príklad 262 (testy o homogenite rozptylov) *Naprogramujte v* R
(a) test pomerom vierohodnosti o homogenite rozptylov a
(b) Bartlettovu modifikáciu testu pomerom vierohodnosti o homogenite rozptylov. Funkciu nazvite *testy.o.rozpty*
Test (b) porovnajte s výsledkom funkcie *bartlett.test().*

Riešenie v

```
"testy.o.rozptyloch" <- function(vec.n,vec.s2){
J <- length(vec.n)
mle.s2 <- (vec.n-1)*vec.s2/vec.n
pooled.s2 <- sum((vec.n-1)*vec.s2)/sum(vec.n)
u.LR <- sum(vec.n*log(pooled.s2/mle.s2 ))
p.hodn <- 1-pchisq(u.LR,df=J-1)
VYSL.LR.test <- c(pooled.s2,u.LR,p.hodn)
names(VYSL.LR.test) <- c("s2","test-stat","p-hodnota")
pooled.s2 <- sum((vec.n-1)*vec.s2)/sum(vec.n-1)
u.LR.alt <- sum((vec.n-1) * log(pooled.s2/vec.s2))
menovatel <- 1+(sum(1/(vec.n-1))-1/sum(vec.n-1))/(3*(J-1))
u.B <- u.LR.alt/menovatel
p.hodn <- 1-pchisq(u.B,df=J-1)
VYSL.B.test <- c(pooled.s2,u.B,p.hodn)
names(VYSL.B.test) <- c("s2","test-stat","p-hodnota")
VYSL <- list(VYSL.LR.test,VYSL.B.test)
return(VYSL)
}
```

Príklad 263 (testy o homogenite rozptylov; simulačná štúdia) *Na základe simulačnej štúdie preverte, že ak (a)* $Y_j \sim N(\mu_j, \sigma^2)$*, kde* $\mu_j = j$*,* $\sigma^2 = 1/4$*,* $j = 1, 2, \ldots, J$ *a (b)* $Y_j \sim [(1-p)N(\mu_j, \sigma^2) + pN(\mu_j, \sigma_1^2)]$*, kde* $p = 0.05$ *a* $\sigma_1^2 = 1$*, potom testovacia štatistika (1)* U_{LR} *a (2)* U_B *má* χ^2_{J-1} *rozdelenie s* $J-1$ *stupňami voľnosti. Použite rozsahy náhodných výberov* $n_j = 10$ *a* $n_j = 50$*. Pre každú simuláciu vypočítajte* $u_{LR,m}$ *a* $u_{B,m}$*,* $m = 1, 2, \ldots, M$*, kde* $M = 10000$*. Superponujte histogram vygenerovaných testovacích štatistík v relatívnej škále s teoretickou krivkou hustoty (1)* U_{LR} *a (2)* U_B*.*

Riešenie (riešenie (a) pozri na obrázku 8.17)
Na základe simulačnej štúdie z príkladu 263 je lepšie používať Bartlettovu modifikáciu testu pomerom vierohodnosti o homogenite rozptylov ako test pomerom vierohodnosti pre $n \leq 50$, pretože konvergencia testovacej štatistiky testu pomerom vierohodnost k χ^2_{J-1} rozdeleniu je pomalšia.

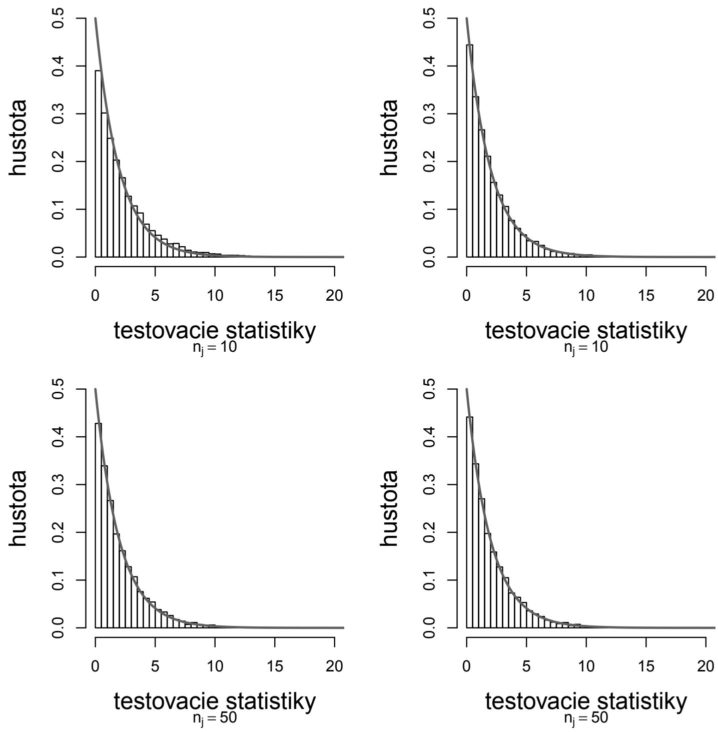
Príklad 264 (testy o homogenite rozptylov) *Majme dáta* more-samples-variances-clavicle.txt *a premennú dĺžka kľúčnej kosti u mužov* cla.L *v mm. Otestujte rovnosť rozptylov štyroch populácií pomocou (1) testu pomerom vierohodnosti a (2) Bartlettovej modifikácie testu pomerom vierohodnosti o homogenite rozptylov. Pokiaľ* $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma^2$ *zamietnete, zistite medzi ktorými rozptylmi je štatisticky signifikantný rozdiel na hladine významnosti* α *použitím* χ^2_{J-1} *(Scheffeho metóda; správny spôsob) a* χ^2_1 *(nesprávny spôsob).*

Riešenie v
$u_{\text{LR}} \doteq 5.086$ a p-hodnota $\doteq 0.166$
$u_{\text{B}} \doteq 5.075$ a p-hodnota $\doteq 0.166$.

```
DATA <- read.table("more-samples-variances-clavicle.txt",header=TRUE)
names(DATA)
## [1] "population" "sex"          "cla.L"
DATA <- na.omit(DATA)
attach(DATA)
n <- tapply(cla.L,population,length)
# eng  gre  ind1  ind2
#  70   94   120    81
rozptyly <- tapply(cla.L,population,var)
#       eng        gre       ind1       ind2
#110.46149   83.15546   76.27283   67.57184
testy.o.rozptyloch(n,rozptyly)
#         s2   test-stat   p-hodnota
#81.7466134   5.0857782   0.1656227
#   s2   test-stat   p-hodnota
#82.6523930   5.0747268   0.1664064
```

Predpokladané rozdiely mohli byť teoreticky vyvolané odlišnými faktormi postnatálneho rastu kľúčnej kosti v každej z populácií alebo odlišne diverzným zložením každej populácie z hľadiska vnútorných rozdielov v raste kľúčnej kosti v puberte a adolescencii. I keď v hodnotách rozptylov vidíme relatívne veľké číselné rozdiely, viacvýberový test homogenity rozptylov nezamietol nulovú hypotézu zhody rozptylov dĺžky kľúčnej kosti pravej strany medzi mužmi pochádzajúcimi zo štyroch rôznych populácií. Číselné rozdiely v rozptyle sú teda spôsobené nejakými náhodnými faktormi (vzorkovanie, meranie, ap.). Významné rozdiely by sa však mohli prejaviť u kostí ľavej strany, ktorá je u väčšiny ľudí v priebehu postnatálneho rastu zaťažovaná menej a tiež rastie v priemere dlhšie než kľúčna kosť pravej strany

(Mays a kol., 1999), priestor pre vznik systematických rozdielov v rozptyle je tak u nej o niečo väčší. Významné rozdiely by sa tiež mohli prejaviť pri porovnaní mužov a žien (napr. Singh a Jit, 1996), keďže vzorce postnatálneho rastu sú u mužov a žien odlišné a v rade metrických znakov sú muži variabilnejší než ženy (Marini a kol., 1999). Vzhľadom na relatívne mierny veľkostný dimorfizmus a podobné stredné hodnoty metrických znakov u mužov a žien (najbežnejšie od cca 5 do 15 percent) je vhodné porovnávať variabilitu medzi pohlaviami na základe rozptylov v jednotkách odpovedajúcich druhej mocnine jednotiek pôvodného znaku. Pokiaľ by však bol rozdiel medzi strednými hodnotami pre mužov a ženy väčší (napr. v niektorých rozmeroch zápästných kostí; Mastrangelo a kol., 2011), rozdiely stredných hodnôt by mohli mechanicky ovplyvňovať aj rozdiely v rozptyloch. V takom prípade je vhodnejšie porovnanie relatívneho rozptýlenia hodnôt, napr. koeficientu variácie (pozri kapitolu 3.2).



Obr. 8.17: Histogramy vygenerovaných testovacích štatistík (za platnosti nulovej hypotézy) v relatívnej škále superponované s teoretickými krivkami hustoty U_{LR} (vľavo) a U_{B} (vpravo)

8.3 Asymptotické testy o korelačných koeficientoch

V tejto kapitole sa budeme zaoberať testami o homogenite (rovnosti) viacerých korelačných koeficientov. Najprv odvodíme χ^2-testovaciu štatistiku testujúcu rovnosť Fisherových Z-transformácií (zovšeobecnenie dvojvýberovej Z-štatistiky z kapitoly 7.3) a následne testovaciu štatistiku pomerom vierohodnosti. Hypotézy definujeme nasledovne $H_0 : \rho_1 = \rho_2 = \ldots = \rho_J = \rho$ oproti $H_1 : \rho_i \neq \rho_j \neq \rho$ pre aspoň jedno $i < j, i = 1, 2, \ldots, J-1; j = 2, 3, \ldots, J$.

Z kapitol 6.3 a 7.3 vieme, že výberový korelačný koeficient R má asymptoticky normálne rozdelenie, ale konvergencia k nemu je pomalá (rozdelenie je zošikmené) a tiež, že je lepšie použiť normalizujúcu a rozptyl stabilizujúzu Fisherovu Z-transformáciu korelačného koeficientu, ktorá má asymptoticky normálne rozdelenie. Potom (Donner a Rosner, 1974)

$$\chi^2_{\mathrm{F}} = \sum_{j=1}^{J} \left(\frac{Z_j - \xi_0}{\frac{1}{\sqrt{n_j - 3}}} \right)^2 = \sum_{j=1}^{J} (n_j - 3)(Z_j - \xi_0)^2 \overset{\mathcal{D}}{\sim} \chi^2_J.$$

kde $Z_j = \frac{1}{2} \ln \frac{1+R_j}{1-R_j}$, $j = 1, 2, \ldots, J$a $\xi_0 = \frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0}$ a ξ_0 je potrebné odhadnúť nasledovne

$$\widetilde{\xi}_0 = \overline{Z} = \frac{\sum_{j=1}^{J} (n_j - 3) Z_j}{\sum_{j=1}^{J} (n_j - 3)},$$

z čoho vyplýva že

$$R_F = \tanh \overline{Z} = \frac{\exp\left(2\overline{Z}\right) - 1}{\exp\left(2\overline{Z}\right) + 1}.$$

Alternatívna verzia výpočtu $\overline{Z}$ je uvedená v práci Hotelling (1953). Potom bude platiť

$$\chi^2_{\mathrm{F}} = \sum_{j=1}^{J} (n_j - 3)(Z_j - \overline{Z})^2 \overset{\mathcal{D}}{\sim} \chi^2_{J-1}.$$

Test pomerom vierohodnosti o rovnosti korelačných koeficientov.

Nech $(X_j, Y_j)^T \sim N_2(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, kde $(X_j, Y_j)^T$ sú nezávislé, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_J)^T$, $\boldsymbol{\theta}_j = (\mu_{j1}, \mu_{j2}, \sigma^2_{j1}, \sigma^2_{j2}, \rho_j)^T$ a $j = 1, 2, \ldots, J$. Logaritmus funkcie vierohodnosti bude mať tvar (Paul, 1989)

$$\begin{aligned} l(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_J, \mathbf{y}_J) &= -n \ln(2\pi) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(\sigma^2_{j1}\sigma^2_{j2}\right) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \rho^2_j\right) \\ &- \sum_{j=1}^{J} \frac{1}{2(1-\rho^2_j)} \left(\frac{\sum_{i=1}^{n_j} (x_{ji} - \mu_{j1})^2}{\sigma^2_{j1}} - 2\rho_j \frac{\sum_{i=1}^{n_j} (x_{ji} - \mu_{j1})(y_{ji} - \mu_{j2})}{\sigma_{j1}\sigma_{j2}} \right. \\ &+ \left. \frac{\sum_{i=1}^{n_j} (y_{ji} - \mu_{j2})^2}{\sigma^2_{j2}} \right), \end{aligned}$$

kde $n = \sum_{j=1}^{J} n_j$. MLE $\widehat{\boldsymbol{\theta}}_j = (\widehat{\mu}_{j1}, \widehat{\mu}_{j2}, \widehat{\sigma}^2_{j1}, \widehat{\sigma}^2_{j2}, \widehat{\rho}_j)^T$, kde

$$\widehat{\mu}_{j1} = \overline{x}_j, \widehat{\mu}_{j2} = \overline{y}_j, \widehat{\sigma}^2_{j1} = \frac{1}{n_j} \sum_{i=1}^{n_j} (x_{ji} - \overline{x}_j)^2, \widehat{\sigma}^2_{j2} = \frac{1}{n_j} \sum_{i=1}^{n_j} (y_{ji} - \overline{y}_j)^2,$$

$$\widehat{\rho}_j = \frac{\sum_{i=1}^{n_j} (x_{ji} - \overline{x}_j)(y_{ji} - \overline{y}_j)}{n_j s_{j1} s_{j2}}$$

a $\Theta_1 = \{\boldsymbol{\theta} : \rho_i \neq \rho_j \neq \rho, i < j, i = 1, 2, \ldots, J-1; j = 2, 3, \ldots, J\}$ pre aspoň jedno i a j. Za platnosti H_0 je $\boldsymbol{\theta}_{j0} = (\widehat{\mu}_{j1}, \widehat{\mu}_{j2}, \widetilde{\sigma}^2_{j1}, \widetilde{\sigma}^2_{j2}, \widehat{\rho})^T$, kde

$$\widetilde{\sigma}^2_{j1} = \frac{\widehat{\sigma}^2_{j1}(1 - \widehat{\rho}\widehat{\rho}_j)}{1 - \widehat{\rho}^2}, \widetilde{\sigma}^2_{j2} = \frac{\widehat{\sigma}^2_{j2}(1 - \widehat{\rho}\widehat{\rho}_j)}{1 - \widehat{\rho}^2}$$

maximálne vierohodný odhad $\widehat{\rho}$ je iteračným riešením rovnice $\sum_{j=1}^{J} \frac{n_j(\widehat{\rho}_j - \widehat{\rho})}{1 - \widehat{\rho}_j\widehat{\rho}} = 0$ (pozri Pearson,

1933) a $\Theta_0 = \{\boldsymbol{\theta} : \rho_1 = \rho_2 = \rho\}$. Logaritmus funkcie vierohodnosti pre $\widehat{\boldsymbol{\theta}}$ bude mať tvar

$$\begin{aligned} l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_J, \mathbf{y}_J) &= -n \ln(2\pi) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(\widehat{\sigma}_{j1}^2 \widehat{\sigma}_{j2}^2\right) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \widehat{\rho}_j^2\right) \\ &- \sum_{j=1}^{J} \frac{1}{2(1-\widehat{\rho}_j^2)} \left(\frac{n_j \widehat{\sigma}_{j1}^2}{\widehat{\sigma}_{j1}^2} - 2\widehat{\rho}_j \frac{n_j \widehat{\sigma}_{j1}^2 \widehat{\sigma}_{j2}^2 \widehat{\rho}_j}{\widehat{\sigma}_{j1}^2 \widehat{\sigma}_{j2}^2} + \frac{n_j \widehat{\sigma}_{j2}^2}{\widehat{\sigma}_{j2}^2} \right) \\ &= -n \ln(2\pi) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(\widehat{\sigma}_{j1}^2 \widehat{\sigma}_{j2}^2\right) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \widehat{\rho}_j^2\right) - n. \end{aligned}$$

Logaritmus funkcie vierohodnosti pre $\boldsymbol{\theta}_0$ bude mať tvar

$$\begin{aligned} l(\boldsymbol{\theta}_0|\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_J, \mathbf{y}_J) &= -n \ln(2\pi) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(\widetilde{\sigma}_{j1}^2 \widetilde{\sigma}_{j2}^2\right) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \widehat{\rho}^2\right) \\ &- \sum_{j=1}^{J} \frac{1}{2(1-\widehat{\rho}^2)} \left(\frac{n_j \widehat{\sigma}_{j1}^2}{\widetilde{\sigma}_{j1}^2} - 2\widehat{\rho} \frac{n_j \widehat{\sigma}_{j1}^2 \widehat{\sigma}_{j2}^2 \widehat{\rho}_j}{\widetilde{\sigma}_{j1}^2 \widetilde{\sigma}_{j2}^2} + \frac{n_j \widehat{\sigma}_{j2}^2}{\widetilde{\sigma}_{j2}^2} \right) \\ &= -n \ln(2\pi) - \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(\widehat{\sigma}_{j1}^2 \widehat{\sigma}_{j2}^2\right) - \sum_{j=1}^{J} n_j \ln \frac{(1 - \widehat{\rho}\widehat{\rho}_j)^2}{(1-\widehat{\rho}^2)^2} \\ &- \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \widehat{\rho}^2\right) - n \end{aligned}$$

Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude potom rovný

$$\begin{aligned} -\ln(\lambda(\boldsymbol{\theta}|\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_J, \mathbf{y}_J)) &= l(\widehat{\boldsymbol{\theta}}|\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_J, \mathbf{y}_J) - l\left(\boldsymbol{\theta}_0|\mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_J, \mathbf{y}_J\right) \\ &= -\sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \widehat{\rho}_j^2\right) + \sum_{j=1}^{J} \frac{n_j}{2} \ln\left(1 - \widehat{\rho}^2\right) + \sum_{j=1}^{J} n_j \ln \frac{(1 - \widehat{\rho}\widehat{\rho}_j)^2}{(1-\widehat{\rho}^2)^2} \\ &= \frac{1}{2} \sum_{j=1}^{J} n_j \ln \frac{(1 - \widehat{\rho}\widehat{\rho}_j)^2}{(1-\widehat{\rho}^2)(1-\widehat{\rho}_j^2)}. \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2 \ln\left(\lambda\left(\mathbf{X}_1, \mathbf{Y}_1, \ldots, \mathbf{X}_J, \mathbf{Y}_J\right)\right) \stackrel{\mathcal{D}}{\sim} \chi^2_{J-1}$.

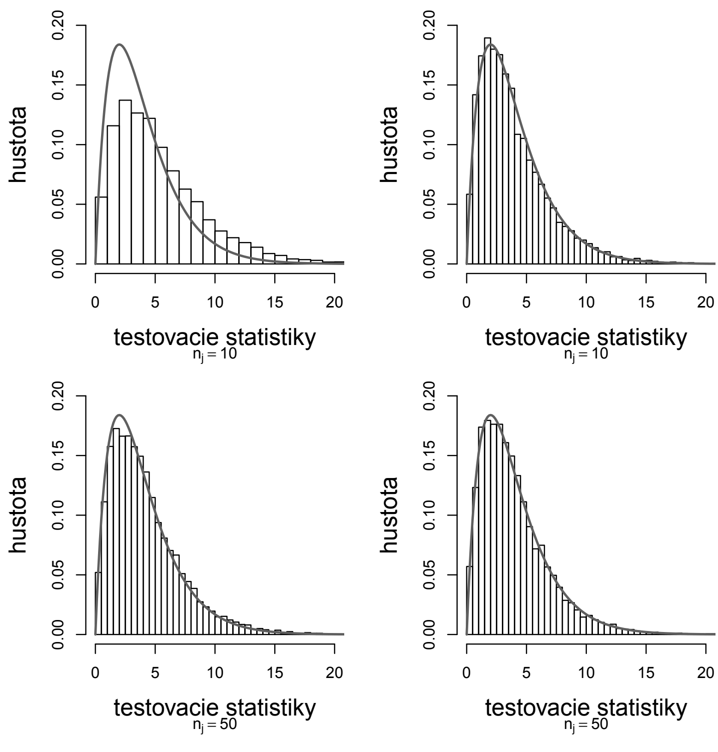
Príklad 265 (testy o homogenite korelačných koeficientov) *Naprogramujte v R (a) test pomerom vierohodnosti a (b) χ^2-test o homogenite korelačných koeficientov. Funkciu nazvite* testy.o.kor.koef().

Riešenie v R

```
"rho.hat" <- function(r,vec.r,vec.n) sum(vec.n*(vec.r-r)/(1-vec.r*r))
"testy.o.kor.koef" <- function(vec.n,vec.r){
J <- length(vec.n)
kor <- uniroot(rho.hat,interval=c(-1,1),vec.n=vec.n,vec.r=vec.r)$root
u.LR <- sum(vec.n*log((1-vec.r*kor)^2/((1-kor^2)*(1-vec.r^2))))
p.hodn <- 1-pchisq(u.LR,df=J-1)
VYSL.LR.test <- c(u.LR,p.hodn)
names(VYSL.LR.test) <- c("test-stat","p-hodnota")
xi.j <- atanh(vec.r)
z.bar <- sum((vec.n-3)*xi.j)/sum(vec.n-3)
kor <- (exp(2*z.bar)-1)/(exp(2*z.bar)+1)
#kor <- tanh(z.bar)
chisq.obs <- sum((vec.n-3)*(xi.j-z.bar)^2)
p.hodn <- 1-pchisq(chisq.obs,df=J-1)
VYSL.chisq.test <- c(kor,chisq.obs,p.hodn)
names(VYSL.chisq.test) <- c("rho.hat","test-stat","p-hodnota")
VYSL <- list(VYSL.LR.test,VYSL.chisq.test)
return(VYSL)
}
```

Príklad 266 (testy o homogenite korelačných koeficientov; simulačná štúdia) *Na základe simulačnej štúdie preverte, že ak (a)* $(X_j, Y_j)^T \sim N_2(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, *kde* $\mu_{j1} = 50, \mu_{j2} = 25$, $\sigma_{j1}^2 = \sigma_{j2}^2 = 9$, $\rho = 0$, $j = 1, 2, \ldots, J$ *a (b)* $pN_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + (1-p)N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, *kde* $p = 0.05$ *a* $\sigma_1^2 = 1$, *potom testovacia štatistika (1)* U_{LR} *a (2)* χ_F^2 *má* χ_{J-1}^2 *rozdelenie s* $J-1$ *stupňami voľnosti. Použite rozsahy náhodných výberov* $n_j = 10$ *a* $n_j = 50$. *Pre každú simuláciu vypočítajte* $u_{LR,m}$ *a* $\chi^2_{F,obs,m}$, $m = 1, 2, \ldots, M$, *kde* $M = 10000$. *Superponujte histogram vygenerovaných testovacích štatistík v relatívnej škále s teoretickou krivkou hustoty (1)* U_{LR} *a (2)* χ_F^2.

Riešenie (riešenie (a) pozri na obrázku 8.18)
Na základe simulačnej štúdie z príkladu 266 je lepšie používať χ^2-test o homogenite korelačných koeficientov ako test pomerom vierohodnosti pre $n \leq 50$, pretože konvergencia testovacej štatistiky testu pomerom vierohodnost k χ^2_{J-1} rozdeleniu je pomalšia.



Obr. 8.18: Histogramy vygenerovaných testovacích štatistík (za platnosti nulovej hypotézy) v relatívnej škále superponované s teoretickými krivkami hustoty U_{LR} (vľavo) a χ^2_{F} (vpravo)

Príklad 267 (testy o homogenite korelačných koeficientov) *Majme dáta* more-samples-correlations-skull.txt *a premenné výška nosa (*nose.H*) a šírka nosa (*nose.B*) v mm. (a) Otestujte rovnosť rozptylov piatich populácií pomocou (1) testu pomerom vierohodnosti a (2)* χ^2*-testu o homogenite korelačných koeficientov. (b) Pokiaľ* $H_0 : \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho$ *zamietnete, zistite medzi ktorými korelačnými koeficientami je štatisticky signifikantný rozdiel na hladine významnosti* α *použitím* χ^2_{J-1} *(Scheffeho metóda; správny spôsob) a* χ^2_1 *(nesprávny spôsob). (c) Vypočítajte aj aritmetické priemery a rozptyly jednotlivých populácií. (d) Otestujte nulovosť jednotlivých korelačných koeficientov v každej z piatich populácií na hladine významnosti* α *použitím* χ^2_{J-1}.

Riešenie v R

$u_{\mathrm{LR}} \doteq 9.820$ a p-hodnota $\doteq 0.044$

$\chi^2_{\mathrm{F,obs}} \doteq 8.428$ a p-hodnota $\doteq 0.077$.

```
DATA <- read.table("more-samples-correlations-skull.txt",header=TRUE)
names(DATA)
#[1] "id"        "pop"        "sex"        "nose.H"     "nose.B"     "intorb.B"
DATA <- na.omit(DATA)
attach(DATA)
hladiny <- levels(pop)
#[1] "ban" "cin" "mal" "nem" "per"
vec.n <- tapply(nose.H,pop,length)
vec.n <- table(pop)
#ban cin mal nem per
# 14  19  73  20  46
vec.r <- numeric(5)
for (j in 1:5) vec.r[j] <- cor(DATA[pop==hladiny[j],4],DATA[pop==hladiny[j],5])
names(vec.r) <- hladiny
vec.r
#        ban        cin        mal        nem        per
#0.67190649 0.01149325 0.18891009 0.59573830 0.13706907
testy.o.kor.koef(vec.n,vec.r)
#   rho.hat   test-stat   p-hodnota
#0.25355680 9.81976444 0.04357567
#  rho.hat  test-stat  p-hodnota
#0.2501608 8.4281651 0.0770948
# zakladne charakteristiky
vec.cov <- matrix(0,5,2)
vec.cov[,1] <- tapply(DATA[,4],pop,var)
vec.cov[,2] <- tapply(DATA[,5],pop,var)
dimnames(vec.cov)[[1]] <- hladiny
vec.mean <- matrix(0,5,2)
vec.mean[,1] <- tapply(DATA[,4],pop,mean)
vec.mean[,2] <- tapply(DATA[,5],pop,mean)
dimnames(vec.mean)[[1]] <- hladiny
TAB.zakl.char <- cbind(vec.mean,vec.cov)
dimnames(TAB.zakl.char)[[2]] <- c("mean(nose.H)","mean(nose.B)",
                                  "var(nose.H)","var(nose.B)")
#     mean(nose.H)  mean(nose.B)  var(nose.H)  var(nose.B)
#ban      48.92857      26.92857    17.456044     9.917582
#cin      53.00000      25.21053     8.555556     2.730994
#mal      51.82192      26.06849    13.953957     4.509132
#nem      51.85000      24.20000     8.450000     5.852632
#per      50.45652      22.97826     9.364734     2.777295
rho0 <- 0
tW.obs <- sqrt(vec.n-2)*vec.r/(sqrt(1-vec.r^2))
uLR.obs <- -vec.n*log(1-tW.obs^2/(tW.obs^2+(vec.n-2)))
#         ban          cin          mal          nem          per
#8.406888343 0.002509969 2.652774161 8.767126374 0.872466639
p.hodn.LR <- 1-pchisq(uLR.obs,df=5-1) # spravny sposob
#        ban        cin        mal        nem        per
#0.07776037 0.99999921 0.61750345 0.06719122 0.92847727
p.hodn.LR <- 1-pchisq(uLR.obs,df=1) # nespravny sposob
#         ban          cin          mal          nem          per
#0.003738019 0.960043029 0.103369360 0.003067083 0.350273343
```

Z výsledkov príkladu 267 je zreteľné, že použitie testu pomerom vierohodnosti nie je kvôli nízkym početnostiam vhodné, čo je aj dôvod rozdielu v signifikancii výsledkov použitých testov. Ďalším problémom môžu byť aj veľké rozdiely v rozptyloch (t.j. ich nehomogenita).

Z popisnej štatistiky je zrejmá malá vzorka u niektorých populácií a tiež veľké rozdiely v rozptyloch. To môže byť príčinou jednak číselných rozdielov v korelačných koeficientoch samotných a jednak nemožnosti zamietnuť zhodu medzi nimi viacvýberovým testom. Ak by sme testovali nulovosť korelačných koeficientov nezávisle pre každú populáciu, zamietli by sme nulovú hypotézu u dvoch populácií (ban a nem), ale nezamietli u dvoch najpočetnejších populácií (mal a per). Keďže ide o viacvýberový problém, žiadnu nulovú hypotézu nezamietame. Na základe dát teda nemôžeme tvrdiť, že medzi dvoma sledovanými rozmermi existuje nejaký vzťah. Súčasne ale vidíme, že (a) všetky korelačné koeficienty sú pozitívne a (b) pozitívna, i keď slabá, je aj závislosť medzi oboma rozmermi u dvoch najpočetnejšie zastúpených populácií (malajská a peruánska). Výsledok správne použitého Scheffeho testu síce na 5% hladine významnosti nezamietol zhodu korelačných koeficientov, významnosť je ale hraničná (p-hodnota = 0.077), čo naznačuje, že pri dostatočne veľkej vzorke by sme rozdiely medzi populáciami mohli nájsť a vysvetlenie tohto rozdielu by bolo treba hľadať v odlišnostiach biologických procesov medzi populáciami a/alebo v procesoch vzniku vzoriek. Táto situácia je žiaľ typická pre metrické porovnanie kostrových vzoriek, ktoré sú (nevyhnutne) výsledkom spracovania populácií zomretých ľudí z pohrebísk (Wood a kol., 1992), kde charakter analyzovaných dát je ovplyvnený celým radom faktorov, od hľadaných biologických procesov a selektívnej mortality, cez pohrebný rítus a tafonomické procesy (zachovalosť a fragmentárnosť pozostatkov), až po rozsah a metodiku archeologického výskumu (čiastkový záchranný výskum – neúplné pohrebiská).

8.4 Asymptotické testy o pravdepodobnostiach

V tejto kapitole sa budeme zaoberať testami o zhode vektora pravdepodobností s nejakým teoretickým vektorom pravdepodobností. Najprv odvodíme testovaciu štatistiku pomerom vierohodnosti, potom Waldovu a skóre testovaciu štatistiku. Ukážeme, že v posledných dvoch prípadoch ide o χ^2-test dobrej zhody. Hypotézu definujeme nasledovne $H_0 : \mathbf{p} = \mathbf{p}_0$ oproti $H_1 : \mathbf{p} \neq \mathbf{p}_0$, kde $\mathbf{p} = (p_1, p_2, \ldots, p_J)^T$ a $\mathbf{p}_0 = (p_{01}, p_{02}, \ldots, p_{0J})^T$. Ďalej sa budeme zaoberať testami o homogenite (rovnosti) viacerých pravdepodobností. Hypotézy definujeme nasledovne $H_0 : p_1 = p_2 = \ldots = p_I = p$ oproti $H_1 : p_i \neq p_j \neq p$ pre aspoň jedno $i < j, i = 1, 2, \ldots, I-1; j = 2, 3, \ldots, I$ (ide o nezávislé pravdepodobnosti z I binomických rozdelení). Nakoniec budeme diskutovať testy homogenity viacerých vektorov pravdepodobností. Pre dva vektory pravdepodobností hypotézy definujeme nasledovne $H_0 : \mathbf{p}_1 = \mathbf{p}_2$ oproti $H_1 : \mathbf{p}_1 \neq \mathbf{p}_2$, kde $\mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{iJ})^T$.

V prvom prípade môžeme realizácie y_j usporiadať do nasledovnej **kontingenčnej tabuľky** $1 \times J$ **pre početnosti**

odpoveď 1	odpoveď 2	...	odpoveď J	spolu
y_1	y_2	...	y_J	N

kde $\sum_{j=1}^{J} y_j = N$ a **kontingenčnej tabuľky** $1 \times J$ **pre pravdepodobnosti**

odpoveď 1	odpoveď 2	...	odpoveď J	spolu
p_1	p_2	...	p_J	1

V druhom prípade ide o špeciálny prípad súčinového multinomického rozdelenia (I nezávislých binomických rozdelení) a nezávislé realizácie y_j môžeme usporiadať do nasledovnej **kontingenčnej tabuľky** $2 \times I$ **pre početnosti**

	skupina 1	skupina 2	...	skupina I
má znak	y_1	y_2	...	y_I
nemá znak	$N_1 - y_1$	$N_2 - y_2$	...	$N_I - y_I$
	N_1	N_2	...	N_I

a **kontingenčnej tabuľky** $2 \times I$ **pre pravdepodobnosti**

	skupina 1	skupina 2	...	skupina I
má znak	p_1	p_2	...	p_I
nemá znak	$1 - p_1$	$1 - p_2$	...	$1 - p_I$
spolu	1	1	...	1

V treťom prípade ide o špeciálny prípad súčinového multinomického rozdelenia (2 nezávislé multinomické rozdelenia) a realizácie y_{ij} môžeme usporiadať do nasledovnej **kontingenčnej tabuľky** $2 \times J$ **pre početnosti**

	kategória 1	kategória 2	...	kategória J	spolu
skupina 1	y_{11}	y_{12}	...	y_{1J}	N_1
skupina 2	y_{21}	y_{22}	...	y_{2J}	N_2

a **kontingenčnej tabuľky** $2 \times J$ **pre pravdepodobnosti**

	kategória 1	kategória 2	...	kategória J	spolu
skupina 1	$p_{1\|1}$	$p_{2\|1}$	...	$p_{J\|1}$	1
skupina 2	$p_{1\|2}$	$p_{2\|2}$	...	$p_{J\|2}$	1

Predchádzajúcu tabuľku kvôli zjednodušeniu značenia preznačme nasledovne

	kategória 1	kategória 2	...	kategória J	spolu
skupina 1	p_{11}	p_{12}	...	p_{1J}	1
skupina 2	p_{21}	p_{22}	...	p_{2J}	1

Test pomerom vierohodnosti v multinomickom rozdelení.

Nech $Y \sim Mult_J(N, \mathbf{p})$, kde $\boldsymbol{\theta} = \mathbf{p} = (p_1, p_2, \dots, p_J)^T$. Logaritmus jadra vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{y}) \quad = \quad \ln \prod_{j=1}^{J} p_j^{y_j}.$$

MLE $\boldsymbol{\theta}$ je rovný $\widehat{\boldsymbol{\theta}} = \widehat{\mathbf{p}} = (\widehat{p}_1, \widehat{p}_2, \dots, \widehat{p}_J)^T$, kde $\widehat{p}_j = y_j/N$, t.j. $\Theta_1 = \{\boldsymbol{\theta} : p_j \neq p_{0j}, j = 1, 2, \dots, J\}$ pre aspoň jedno j. Za platnosti H_0 je $\boldsymbol{\theta}_0 = \mathbf{p}_0$ a $\Theta_0 = \{\boldsymbol{\theta} : p_j = p_{0j}, j = 1, 2, \dots, J\}$ pre všetky j. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{y})) &= l\left(\widehat{\boldsymbol{\theta}}|\mathbf{y}\right) - l\left(\boldsymbol{\theta}_0|\mathbf{y}\right) = \sum_{j=1}^{J} y_j \ln \frac{y_j}{N} - \sum_{j=1}^{J} y_j \ln p_{0j} \\ &= \sum_{j=1}^{J} y_j \ln \frac{\frac{y_j}{N}}{p_{0j}} = \sum_{j=1}^{J} y_j \ln \frac{y_j}{N p_{0j}} = \sum_{j} \text{pozorované}_j \times \ln \frac{\text{pozorované}_j}{\text{očakávané}_j}. \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\text{LR}} = -2\ln\left(\lambda\left(\mathbf{Y}\right)\right) \overset{\mathcal{D}}{\sim} \chi^2_{J-1}$.

Waldov test v multinomickom rozdelení.
Na základe definície multinomického rozdelenia, definície $\mathcal{I}(\mathbf{p})$ z kapitoly 2, asymptotickej normality $\sqrt{N}\left(\frac{1}{N}\mathbf{X} - \mathbf{p}_0\right) \sim N_{J-1}\left(\mathbf{0}_{J-1}, \boldsymbol{\Sigma}_0\right)$, rozdelenia kvadratickej formy $N\left(\frac{1}{N}\mathbf{X} - \mathbf{p}_0\right)^T \boldsymbol{\Sigma}_0^{-1}\left(\frac{1}{N}\mathbf{X} - \mathbf{p}_0\right) \sim \chi^2_{J-1}$ a rovnosti $p_{0J} = 1 - \sum_{j=1}^{J-1} p_{0j}$ (kde vektory pravdepodobností sú $(J-1)$-rozmerné), môžeme písať

$$\begin{aligned} u_W &= N\sum_{j=1}^{J-1}\left(\widehat{p}_j - p_{0j}\right)^2/p_{0j} + N\sum_{i=1}^{J-1}\sum_{j=1}^{J-1}\left(\widehat{p}_i - p_{0i}\right)\left(\widehat{p}_j - p_{0j}\right)/p_{0J} \\ &= N\sum_{j=1}^{J-1}\left(\widehat{p}_j - p_{0j}\right)^2/p_{0j} + N\left[\sum_{j=1}^{J-1}\left(\widehat{p}_j - p_{0j}\right)\right]^2/p_{0J} \\ &= N\sum_{j=1}^{J-1}\left(\widehat{p}_j - p_{0j}\right)^2/p_{0j} + N\left(\widehat{p}_J - p_{0J}\right)^2/p_{0J} \\ &= N\sum_{j=1}^{J}\left(\widehat{p}_j - p_{0j}\right)^2/p_{0j} = \sum_{j=1}^{J}\frac{\left(y_j - Np_{0j}\right)^2}{Np_{0j}} \\ &= \sum_{j=1}^{J}\frac{(\text{pozorované}_j - \text{očakávané}_j)^2}{\text{očakávané}_j}. \end{aligned}$$

Vieme, že **Waldova testovacia štatistika** $U_W \overset{\mathcal{D}}{\sim} \chi^2_{J-1}$.

Skóre test v multinomickom rozdelení.
Na základe definície multinomického rozdelenia, definície $\mathcal{I}^{-1}(\mathbf{p})$ z kapitoly 2 a rovnosti $p_{0J} = 1 - \sum_{j=1}^{J-1} p_{0j}$ (kde vektory pravdepodobností sú $(J-1)$-rozmerné), môžeme písať

$$\begin{aligned} u_S &= N\sum_{j=1}^{J-1}\left(\frac{\widehat{p}_j}{p_{0j}} - \frac{\widehat{p}_J}{p_{0J}}\right)^2 p_{0j} - N\sum_{i=1}^{J-1}\sum_{j=1}^{J-1}\left(\frac{\widehat{p}_i}{p_{0i}} - \frac{\widehat{p}_J}{p_{0J}}\right)\left(\frac{\widehat{p}_j}{p_{0j}} - \frac{\widehat{p}_J}{p_{0J}}\right)p_{0i}p_{0j} \\ &= N\sum_{j=1}^{J-1}(B)^2 p_{0j} + A = C + A, \end{aligned}$$

kde $\sum_{j=1}^{J-1}(\widehat{p}_j - p_{0j}) = -(\widehat{p}_J - p_{0J})$,

$$A = -N\left[\sum_{j=1}^{J-1}\left(\frac{\widehat{p}_j}{p_{0j}} - \frac{\widehat{p}_J}{p_{0J}}\right)p_{0J}\right]^2 = -N\left(\frac{\widehat{p}_J}{p_{0J}} - 1\right)^2$$

$$B = \frac{\widehat{p}_j p_{0J} - \widehat{p}_J p_{0j} \pm p_{0J}p_{0j}}{p_{0j}p_{0J}} = \left[p_{0J}(\widehat{p}_j - p_{0j}) - p_{0j}(\widehat{p}_J - p_{0J})\right]/(p_{0j}p_{0J}),$$

$$\begin{aligned} C &= N\left[\sum_{j=1}^{J-1}\frac{1}{p_{0j}}(\widehat{p}_j - p_{0j})^2 + \frac{2}{p_{0J}}(\widehat{p}_J - p_{0J})^2 + \frac{1}{p_{0J}^2}(1 - p_{0J})(\widehat{p}_J - p_{0J})^2\right] \\ &= N\left[\sum_{j=1}^{J}\frac{1}{p_{0j}}(\widehat{p}_j - p_{0j})^2 + \frac{1}{p_{0J}^2}(\widehat{p}_J - p_{0J})^2\right]. \end{aligned}$$

Potom

$$u_S = N\sum_{j=1}^{J}\frac{1}{p_{0j}}(\widehat{p}_j - p_{0j})^2 = \sum_{j=1}^{J}\frac{(y_j - Np_{0j})^2}{Np_{0j}} = \sum_{j=1}^{J}\frac{(\text{pozorované}_j - \text{očakávané}_j)^2}{\text{očakávané}_j}.$$

Vieme, že **skóre testovacia štatistika** $U_S \overset{\mathcal{D}}{\sim} \chi^2_{J-1}$.

Sila testu. Sila χ^2 testu (Waldovho a skóre testu v multinomickom rozdelení) $1-\beta$ je definovaná ako

$$1 - \beta = \Pr\left(\chi^2_{J-1,\lambda^2} \geq \chi^2_{J-1}(\alpha)\right),$$

kde $\chi^2_{J-1}(\alpha)$ je kritická hodnota χ^2_{J-1} rozdelenia s $J-1$ stupňami voľnosti a χ^2_{J-1,λ^2} náhodná premenná z necentrálneho χ^2_{J-1} rozdelenia s $J-1$ stupňami voľnosti a parametrom necentrality λ^2. Potom **minimálny rozsah** N pre nejaký vektor pravdepodobností $\mathbf{p}$ za platnosti alternatívnej hypotézy bude rovný

$$N = \lambda^2\left(\sum_{j=1}^{J}N\frac{(p_j - p_{0j})^2}{p_{0j}}\right)^{-1},$$

kde

$$\lambda^2 = \lim_{N\to\infty}\sum_{j=1}^{J}N\frac{(p_j - p_{0j})^2}{p_{0j}}$$

a tento parameter necentrality vypočítame iteračným riešením rovníc $1-\beta = \Pr\left(\chi^2_{J-1,\lambda^2} \geq \chi^2_{J-1}(\alpha)\right)$.

Príklad 268 (minimálny rozsah) *Naprogramujte v R funkciu na výpočet minimálneho rozsahu súboru N. Nazvite ju* minimalny.rozsah.multinom()*. Aplikujte na vektory* $\mathbf{p} = (0.4, 0.3, 0.2, 0.1)^T$ *a* $\mathbf{p}_0 = (0.420, 0.366, 0.160, 0.054)^T$.

Riešenie

Parameter necentrality $\lambda^2 = 11.48$ (necentrálneho χ^2_{J-1} rozdelenia) a minimálny rozsah $N = 185$.

Krvné A a B sú dominanté nad 0 a kodominantné navzájom. Preto pre fenotyp 0 existuje len jeden genotyp 00; fenotyp krvnej skupiny A pozostáva z genotypov AA alebo $A0$, fenotyp krvnej skupiny B z genotypov BB alebo $B0$; pre fenotyp AB existuje len jeden genotyp AB. Nech p je pravdepodobnosť výskytu alely 0, q je pravdepodobnosť výskytu alely A a r je pravdepodobnosť výskytu alely B. Potom $p+q+r=1$. Za platnosti Hardy-Weinbergovho (HW) zákona môžeme písať

$$(p+q+r)^2 = p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1, \text{ kde } p = 1 - q - r.$$

Príklad 269 (populačná genetika; krvné skupiny) *Majme dáta* multinom-blood-groups.txt. *(a) Pomocou χ^2-testu dobrej zhody otestuje zhodu početností krvných skupín pre Prahu s Hardy-Weinbergovým zákonom. (b) Nakreslite logaritmus štandardizovanej funkcie vierohodnosti v parametroch q a r pomocou funkcie* contour(). *Dokreslite do obrázku jej maximum v bode $\widehat{\boldsymbol{\theta}} = (\widehat{q}, \widehat{r})^T$. (c) V obrázku zvýraznite vierohodnostnú 95% empirickú elipsu spoľahlivosti pomocou funkcie* polygon(). *(c) Nakreslite stĺpcový diagram relatívnych početností krvných skupín (v %) pre Košice a Prahu.*

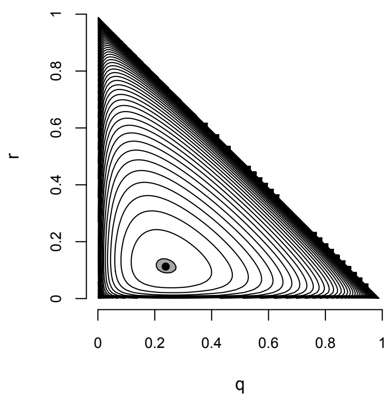
Riešenie (pozri tabuľku 8.5, obrázky 8.19)
Nech $(Y_1, Y_2, Y_3, Y_4)^T \sim Mult_4(N, \mathbf{p})$, kde $N = 500$ a $\mathbf{p} = (p_1, p_2, p_3, p_4)^T$. Za platnosti HW zákona $(Y_1, Y_2, Y_3, Y_4)^T \sim Mult_4(N, \mathbf{p}_0)$, kde $N = 500$, $\mathbf{p}_0 = (p_{01}, p_{02}, p_{03}, p_{04})^T$, $p_{01} = p^2$ (genotyp 00), $p_{02} = q^2 + 2qp$ (genotyp AA a $A0$), $p_{03} = r^2 + 2rp$ (genotyp BB a $B0$) a $p_{04} = 2qr$ (genotyp AB). Pravdepodobnosti q a r odhadneme pomocou maximalizácie logaritmu funkcie vierohodnosti parametra $\boldsymbol{\theta} = (q, r)^T$, kde po dosadení do jadra logaritmu funkcie vierohodnosti dostaneme

$$l(\boldsymbol{\theta}|\mathbf{x}) = \sum_{j=1}^{4} y_j \ln p_j = y_1 \ln\left(p^2\right) + y_2 \ln((q^2+2qp)^2) + y_3 \ln((r^2+2rp)^2) + y_4 \ln((2qr)^2), \text{ kde} p = 1-q-r.$$

Odhady pravdepodobností výskytu aliel sú $\widehat{q} = 0.239$, $\widehat{r} = 0.114$ a $\widehat{p} = g(\widehat{\boldsymbol{\theta}}) = b + \mathbf{a}^T\widehat{\boldsymbol{\theta}} = 1 - \widehat{q} - \widehat{r} = 0.648$ (čo sme mohli urobiť na základe invariantnosti odhadu), kde $\mathbf{a} = (-1, -1)^T$ a $b = 1$. Rozptyly parametrov $\widehat{q}$ a $\widehat{r}$ vypočítame z hesiánu, t.j. $\widehat{Var[\widehat{q}]} = 0.0002$ a $\widehat{Var[\widehat{r}]} = 0.0001$. $\widehat{Var[\widehat{p}]} = \mathbf{a}^T(\mathcal{I}(\widehat{\boldsymbol{\theta}}))^{-1}\mathbf{a} = 0.0003$. Vierohodnostnú 95% empirickú elipsu spoľahlivosti je možné vypočítať pomocou cut-off logaritmu štandardizovanej funkcie vierohodnosti v bode $\log(\exp(-\frac{1}{2}\chi^2_2(0.05))) = \log(0.05) = -2.995732$ (pozri obrázok 8.19).

Tabuľka 8.5: Pozorované početnosti a pravdepodobnosti a očakávané početnosti a pravdepodobnosti (vypočítané pomocou funkcie vierohodnosti) pre Prahu

krvné skupiny	0	A	B	AB
pozorované početnosti y_j	209	184	81	26
pozorované pravdepodobnosti y_j/N	0.4180	0.3680	0.1620	0.0520
očakávané početnosti Np_{0j}	210	183	80	27
očakávané pravdepodobnosti p_{0j}	0.4197	0.3661	0.1600	0.0542



Obr. 8.19: Logaritmus štandardizovanej funkcie vierohodnosti s maximom označeným • a vierohodnostnou 95% empirickou elipsou spoľahlivosti pre $\boldsymbol{\theta}$

1. **Hypotézy**
   - **slovná formulácia** – H_0 : početnosti výskytu krvných skupín sú v zhode s HW zákonom oproti H_1 : početnosti výskytu krvných skupín nie sú v zhode s HW zákonom.
   - **matematická formulácia** – $H_0 : \mathbf{p} = \mathbf{p}_0$ oproti $H_1 : \mathbf{p} \neq \mathbf{p}_0$.
2. **Testovacia štatistika** – $\chi^2_{\text{obs}} \doteq 0.065$.

```
chi.obs <- sum((poz.poc-ocak.poc)^2/ocak.poc) # 0.06484433
```

3. **Zamietacia oblasť** – kritická hodnota $\chi^2_3(\alpha) = \chi^2_3(0.05) = 5.991$, potom $\mathcal{W} = (5.991, \infty)$.

```
qchisq(1-0.05,3) # 7.814728
```

4. **Štatistický záver** – p-hodnota$= 0.664$. Kedže p-hodnota nie je menšia ako $\alpha = 0.05$ (ekvivalentne χ^2_{obs} nepatrí zamietacej oblasti), H_0 nezamietame na hladine významnosti α.

```
1-pchisq(chi.obs,df=3) # 0.9956928
```

5. **Slovný záver** – Nemáme dostatok dôkazov na zamietnutie nulovej hypotézy o tom, že početnosti výskytu krvných skupín sú v zhode s HW zákonom.
6. **Antropologický záver** – Empiricky zistené početnosti krvných skupin $AB0$ systému vo vzorke pražského obyvateľstva zodpovedajú teoretickým početnostiam vyplývajúcim z HW zákona a daného modelu dedičnosti (alely A a B kodominantné a alela 0 recesívna).

Testovanie na základe hlavných alel $AB0$ systému nemusí kompletne vyčerpať problematiku variability $AB0$ systému a rozdielov medzi populáciami vo frekvenciách krvných skupín tohoto systému, pretože alela A sa ďalej delí na imunitne odlišné formy A_1 a A_2, čím vzniká šesť rôznych fenotypov (0, A_1, A_2, B, A_1B, A_2B), a ďalej existuje celý rad menej početných foriem alely A i B. Navyše je molekulárny podklad antigénov krvných skupín $AB0$ systému (oligosacharidov) prepojený s antigénmi krvných skupín systémov H a Lewis (Mielke a kol., 2011). V prípade štúdia adaptívnych vlastností $AB0$ systému (vzťah k mikrobiálnym patogénom) u rôznych populácií je potrebné vyššie uvedené vziať do úvahy a adekvátne tomu prispôsobiť i štatistický model.

Test pomerom vierohodnosti o homogenite pravdepodobností z I binomických rozdelení.

Nech $Y_i \sim Bin\,(N_i, p_i)$, kde $\theta_i = p_i$. Označme $\boldsymbol{\theta} = \mathbf{p} = (p_1, p_2, \ldots, p_I)^T$. Logaritmus vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_I) \quad = \quad \ln \prod_{j=1}^{I} p_j^{y_j} (1-p_j)^{N_j - y_j}.$$

MLE θ_i je rovný $\widehat{\theta}_i = \widehat{p}_i$, kde $\widehat{p}_i = y_i/N_j$, t.j. $\Theta_1 = \{\boldsymbol{\theta} : p_i \neq p_j, i < j, i = 1, 2, \ldots, I-1; j = 2, 3, \ldots, I\}$ pre aspoň jedno i a j. Za platnosti H_0 je $\boldsymbol{\theta}_0 = (\widehat{p}, \widehat{p}, \ldots, \widehat{p})^T$, kde $\widehat{p} = \sum_{i=1}^I N_i p_i / \sum_{i=1}^I N_i$ a $\Theta_0 = \{\boldsymbol{\theta} : p_1 = p_2 = \ldots = p_I = p\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_I)) \quad &= \quad l\left(\widehat{\boldsymbol{\theta}}|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_I\right) - l\left(\boldsymbol{\theta}_0|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_I\right) = \sum_{i=1}^I y_i \ln \widehat{p}_i \\ &+ \quad \sum_{i=1}^I (N_i - y_i) \ln(1-\widehat{p}_i) - \sum_{i=1}^I y_i \ln \widehat{p} - \sum_{i=1}^I (N_i - y_i) \ln(1-\widehat{p}) \\ &= \quad \sum_{i=1}^I y_i \ln \frac{\widehat{p}_i}{\widehat{p}} + \sum_{i=1}^I (N_i - y_i) \ln \frac{1-\widehat{p}_i}{1-\widehat{p}}. \end{aligned}$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\text{LR}} = -2\ln\left(\lambda\left(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_I\right)\right) \overset{\mathcal{D}}{\sim} \chi^2_{I-1}$. Preznačme p_i na p_{i1} a $1-p_i$ na p_{i2}. Ekvivalentne preznačíme y_i na y_{i1}, $N_i - y_i$ na y_{i2}.

Potom $\boldsymbol{\theta} = \mathbf{p} = (p_{11}, p_{12}, p_{21}, p_{22}, \ldots, p_{I1}, p_{I2})^T$, $\widehat{\boldsymbol{\theta}} = \widehat{\mathbf{p}} = (\widehat{p}_{11}, \widehat{p}_{12}, \widehat{p}_{21}, \widehat{p}_{22}, \ldots, \widehat{p}_{I1}, \widehat{p}_{I2})^T$ a $\boldsymbol{\theta}_0 = \mathbf{p}_0 = (p_{0,11}, p_{0,12}, p_{0,21}, p_{0,22}, \ldots, p_{0,I1}, p_{0,I2})^T = (\widehat{p}, 1-\widehat{p}, \widehat{p}, 1-\widehat{p}, \ldots, \widehat{p}, 1-\widehat{p})^T$, kde $p_{i1} + p_{i2} = 1$, $y_{i1} + y_{i2} = N_i$ a $i = 1, 2, \ldots, I$. Potom dostaneme

$$u_{\mathrm{LR}} = 2\sum_{i=1}^{I}\sum_{j=1}^{2} y_{ij} \ln \frac{\widehat{p}_{ij}}{p_{0,i}} = 2\sum_{i=1}^{I}\sum_{j=1}^{2} \text{pozorované}_{ij} \times \ln \frac{\text{pozorované}_{ij}}{\text{očakávané}_{ij}}.$$

Waldov a skóre test o homogenite pravdepodobností z I binomických rozdelení. Tieto testy odvodíme podobne ako Waldov a skóre test v multinomickom rozdelení. Dostaneme

$$u_W = u_S \quad = \quad \sum_{i=1}^{I}\sum_{j=1}^{2} \frac{(y_{ij} - Np_{0,ij})^2}{Np_{0,ij}} = \sum_{i=1}^{I}\sum_{j=1}^{2} \frac{(\text{pozorované}_{ij} - \text{očakávané}_{ij})^2}{\text{očakávané}_{ij}}.$$

Vieme, že **Waldova a skóre testovacia štatistika** $U_W = U_S \overset{\mathcal{D}}{\sim} \chi^2_{I-1}$.

Príklad 270 (rovnosť I pravdepodobností) *Ak $I = 2$, ktorej z testovacích štatistík Z_W alebo $Z_W^{(alt)}$ testu rozdielu pravdepodobností p_1 a p_2 z kapitoly 7.4 je rovná testovacia štatistika $U_W^{1/2}$?*

Príklad 271 (rovnosť I pravdepodobností) *Ukážte, že $u_W = \frac{1}{\widehat{p}(1-\widehat{p})}\left(\sum_{i=1}^{I} y_{ij}\widehat{p}_i^2 - N\widehat{p}^2.\right)$*

Príklad 272 (rovnosť I pravdepodobností) *Majme dáta* more-samples-probabilities-pubis.txt *a premennú výskyt zmien kostného reliéfu (súčet početností premenných* trace.to.small *a* moderate.to.large*) u troch populácií žien. (a) Pomocou (1) χ^2-testu dobrej zhody (Waldovho a skóre testu) a (2) testu pomerom vierohodnosti otestuje rovnosť v početnostiach žien s výskytom zmien kostného reliéfu u všetkých troch populácií. (b) Nakreslite stĺpcový diagram relatívnych početností absencie a výskytu zmien kostného reliéfu u žien (v %) pre všetky populácie.*

Test pomerom vierohodnosti o homogenite dvoch vektorov pravdepodobností.

Nech $Y_i \sim Mult_J(N_i, \mathbf{p}_i)$, kde $i = 1, 2$, a $\boldsymbol{\theta}_i = \mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{iJ})^T$. Logaritmus jadra vierohodnosti má tvar

$$l(\boldsymbol{\theta}|\mathbf{y}_1, \mathbf{y}_2) \quad = \quad \ln \prod_{i=1}^{2}\prod_{j=1}^{J} p_{ij}^{y_{ij}}.$$

MLE θ_{ij} je rovný $\widehat{\theta}_{ij} = \widehat{p}_{ij}$, kde $\widehat{p}_{ij} = y_{ij}/N_i$, t.j. $\Theta_1 = \{\boldsymbol{\theta} : p_{1j} \neq p_{2j}, j = 1, 2, \ldots, J\}$ pre aspoň jedno j. Za platnosti H_0 je $\boldsymbol{\theta}_0 = (\widehat{p}_1, \widehat{p}_2, \ldots, \widehat{p}_J, \widehat{p}_1, \widehat{p}_2, \ldots, \widehat{p}_J)^T$, kde $p_{0,ij} = \widehat{p}_j = \sum_{i=1}^{2} N_i p_{ij}/N$, kde $N = \sum_{i=1}^{2} N_i$ a $\Theta_0 = \{\boldsymbol{\theta} : p_{1j} = p_{2j}, j = 1, 2, \ldots, J\}$. Potom -1 krát prirodzený logaritmus pomeru vierohodnosti bude rovný

$$\begin{aligned} -\ln(\lambda(\mathbf{y}_1, \mathbf{y}_2)) \quad &= \quad l\left(\widehat{\boldsymbol{\theta}}|\mathbf{y}_1, \mathbf{y}_2\right) - l(\boldsymbol{\theta}_0|\mathbf{y}_1, \mathbf{y}_2) = \sum_{i=1}^{2}\sum_{j=1}^{J} y_{ij}\ln\widehat{p}_{ij} - \sum_{i=1}^{2}\sum_{j=1}^{J} y_{ij}\ln\widehat{p}_j \\ &= \quad \sum_{i=1}^{2}\sum_{j=1}^{J} y_{ij}\ln\frac{\widehat{p}_{ij}}{\widehat{p}_j}. \end{aligned}$$

Potom dostaneme

$$u_{\mathrm{LR}} = 2\sum_{i=1}^{2}\sum_{j=1}^{J} y_{ij} \ln \frac{\widehat{p}_{ij}}{p_{0,ij}} = 2\sum_{i=1}^{2}\sum_{j=1}^{J} \text{pozorované}_{ij} \times \ln \frac{\text{pozorované}_{ij}}{\text{očakávané}_{ij}}.$$

Vieme, že **testovacia štatistika pomerom vierohodnosti** $U_{\mathrm{LR}} = -2\ln\left(\lambda\left(\mathbf{Y}_1, \mathbf{Y}_2\right)\right) \overset{\mathcal{D}}{\sim} \chi^2_{J-1}$.

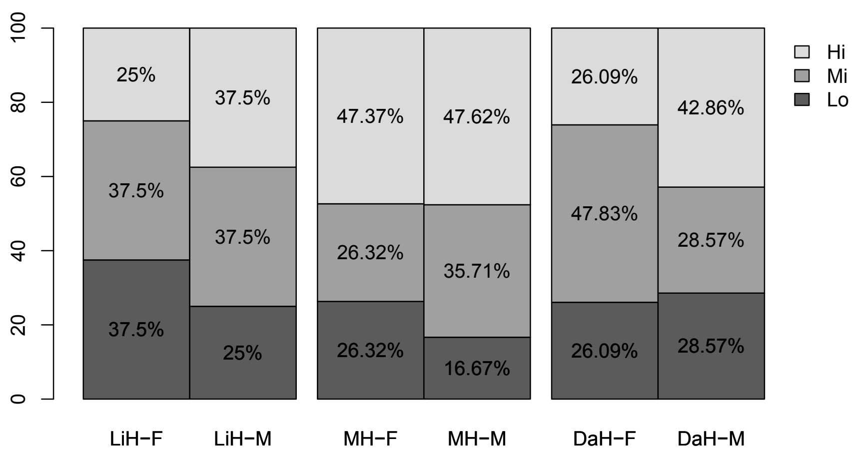
Waldov a skóre test o homogenite dvoch vektorov pravdepodobností.
Tieto testy odvodíme podobne ako Waldov a skóre test v multinomickom rozdelení. Dostaneme

$$u_W = u_S = \sum_{i=1}^{2}\sum_{j=1}^{J}\frac{(y_{ij} - Np_{0,ij})^2}{Np_{0,ij}} = \sum_{i=1}^{2}\sum_{j=1}^{J}\frac{(\text{pozorované}_{ij} - \text{očakávané}_{ij})^2}{\text{očakávané}_{ij}}.$$

Vieme, že **Waldova a skóre testovacia štatistika** $U_W = U_S \overset{\mathcal{D}}{\sim} \chi^2_{J-1}$.

Príklad 273 (homogenita dvoch vektorov pravdepodobností) *Majme dáta* multinom-palmar-lines.txt *pre 200 študentov (100 mužov a 100 žien), ktorí boli zaradení do troch kategórií podľa farby vlasov (*LiH, MH, DaH*) a súčasne do troch kategórií podľa zakončenia troch hlavných dlaňových línií (*Hi, Mi, Lo*). (a) Pomocou (1) χ^2-testu dobrej zhody (Waldovho a skóre testu) a (2) testu pomerom vierohodnosti otestuje pre každú farbu vlasov zvlášť pohlavný dimorfizmus v početnostiach zakončení dlaňových línií. (b) Nakreslite stĺpcový diagram relatívnych početností zakončení dlaňových línií (v %) pre každú farbu vlasov.*

Riešenie (čiastočné; pozri obrázok 8.20)



Obr. 8.20: Stĺpcový diagram relatívnych početností zakončení dlaňových línií (v %) pre každú farbu vlasov.

9 Antropologické dátové subory

9.1 Dátový súbor – jednovýberový test o strednej hodnote

Hodnotený súbor: Z archívnych materiálov (Schmidt, 1888) máme k dispozícii pôvodné kraniometrické údaje o dĺžke a šírke lebky zo starovekej egyptskej populácie (pozri obrázok 9.1). Súčasne máme k dispozícii priemerné hodnoty oboch rozmerov, hodnoty smerodajnej odchýlky a počty prípadov vzorky z novovekej egyptskej populácie (dĺžka lebky: $\overline{x}_{\mathrm{m}} = 177.568$ mm, $\overline{x}_{\mathrm{f}} = 171.962$ mm; s_{m}=7.526 mm, s_{f}=7.052 mm; n_{m}=88, $n_{\mathrm{f}} = 52$ a šírka lebky: $\overline{x}_{\mathrm{m}} = 136.402$ mm, $\overline{x}_{\mathrm{f}} = 131.038$ mm; $s_{\mathrm{m}} = 6.411$ mm, $s_{\mathrm{f}} = 5.361$ mm; $n_{\mathrm{m}} = 87$, $n_{\mathrm{f}} = 52$).

Súbor dát: one-sample-mean-skull-mf.txt

Popis premenných:
id – poradové číslo;
pop – populácie (egant – egyptská staroveká);
sex – pohlavie (m – muž, f – žena);
skull.L – najväčšia dĺžka mozgovne (mm), t.j. priama vzdialenosť kraniometrických bodov *glabella* a *opisthocranion*;
skull.B – najväčšia šírka mozgovne (mm), t.j. vzdialenosť oboch kraniometrických bodov *euryon*.

Biologické súvislosti: Výrazná zmena určitého biologického znaku v populácii po uplynutí dlhšieho časového obdobia sa označuje ako sekulárny trend (z latinského *saeculum* – generácia, vek, storočie). Brachycefalizácia, resp. debrachycefalizácia, t.j. relatívne skracovanie či predlžovanie lebky je jedným z príkladov sekulárneho trendu. Tieto zmeny lebky/hlavy korelujú so zmenami kostí končatín a dávajú sa do súvislostí so zmenami vonkajších životných podmienok i genetického zloženia populácie. Napriek tomu, že pomer šírky a dĺžky lebky závisí od oboch rozmerov, ukazuje sa, že zmeny v tvare lebky ovplyvňujú predovšetkým zmeny v jej šírke.

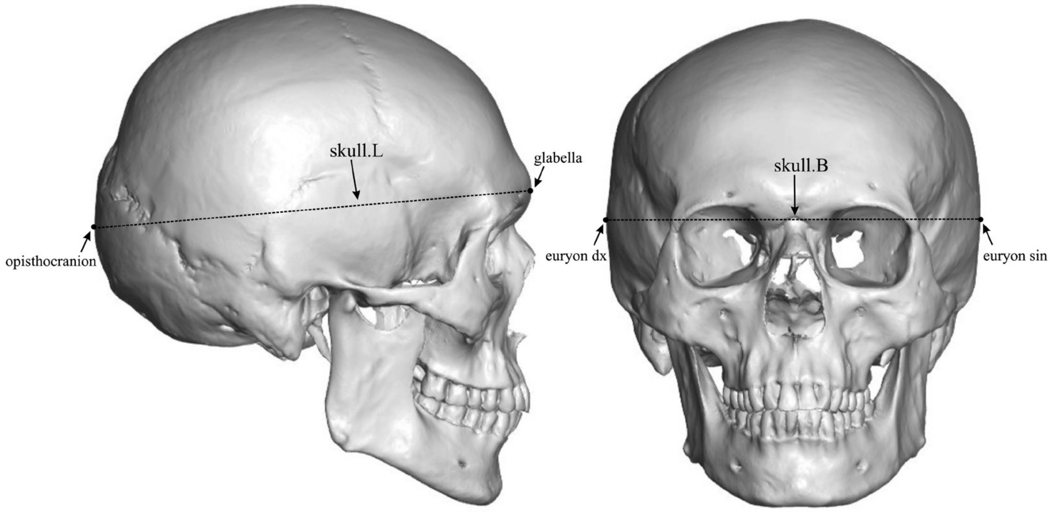
Ciele:
(a) zistiť, či sa dĺžka lebky starovekej egyptskej populácie líši v strednej hodnote od novovekej egyptskej populácie (zvlášť u mužov a u žien);
(b) zistiť, či sa šírka lebky starovekej egyptskej populácie líši v strednej hodnote od novovekej egyptskej populácie (zvlášť u mužov a u žien).

9.2 Dátový súbor – párový test o strednej hodnote

Hodnotený súbor: Hodnotený súbor predstavujú osteometrické dáta, konkrétne hodnoty vertikálneho priemeru stredu dĺžky tela kľúčnej kosti (*clavicula*) z pohrebiska u Sv. Jakuba v Brne, prevažne z obdobia stredoveku. K dispozícii máme hodnoty rozmeru 40 vybraných jedincov na pravej aj ľavej strane tela z pôvodného merania (Živný, 2010) a z dvoch nových opakovaných meraní (Hupková, nepublikované dáta).

Súbor dát: paired-means-clavicle.txt

Popis premenných:
id – poradové číslo jedinca;
sex – pohlavie (M3 – veľmi pravdepodobne muž, M2 – pravdepodobne muž, M1 – skôr muž, I – indiferentný, F1 – skôr žena, F2 – pravdepodobne žena, F3 – veľmi pravdepodobne žena);
side – strana (R – pravá, L – ľavá);
simd – vertikálny priemer v strede dĺžky tela kľúčnej kosti (*superior–inferior midshaft diameter*), 1.



Obr. 9.1: Znázornenie premenných najväčšia dĺžka mozgovne (skull.L) a najväčšia šírka mozgovne (skull.B)

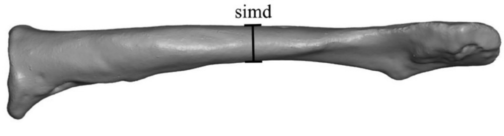
meranie (mm) prvého výskumníka (pozri obrázok 9.2);

simd.1 – vertikálny priemer v strede dĺžky tela kľúčnej kosti (*superior–inferior midshaft diameter*), 1. meranie (mm) druhého výskumníka;

simd.2 – vertikálny priemer v strede dĺžky tela kľúčnej kosti, 2. meranie (mm) druhého výskumníka.

Biologické súvislosti: Stranové rozdiely v hrúbke tela kľúčnej kosti môžu odrážať rozdielne zaťažovanie každej zo strán. U malých rozmerov, ako je vertikálny priemer stredu tela kľúčnej kosti, môže hrať významnú úlohu skreslenie skutočnej hodnoty rozmeru vplyvom chyby merania (interindividuálna chyba – viac výskumníkov a intraindividuálna chyba – jeden výskumník, systematická chyba – správnosť merania a náhodná chyba – presnosť merania). Keďže chyba merania môže výrazne ovplyvniť hodnotenie subtílnych biologických trendov, ako sú napr. stranové rozdiely ((a)symetria), je veľmi dôležité (ešte pred samotnými analýzami stranových rozdielov) chybu merania kvantifikovať.

Ciele:
(a) zistiť, či je stredná hodnota vertikálneho priemeru stredu tela kľúčnej kosti prvého a druhého merania zhodná, t.j. zhodnotiť intraindividuálnu chybu merania;
(b) vypočítať (1) technickú chybu merania, (2) relatívnu technickú chybu merania a (3) koeficient reliability merania (pozri kap. 6.1)
(c) zistiť, či sa stredná hodnota priemeru prvého (*simd.1*) a druhého opakovaného merania (*simd.2*) líši od strednej hodnoty prvého merania (*simd*), t. j. zhodnotiť interindividuálnu chybu merania;
(d) zistiť, či je kľúčna kosť v mieste vertikálneho priemeru stredu diafýzy na pravej strane tela vyvinutá inak než na strane ľavej, t.j. či existuje systematická stranová asymetria;
(e) overiť, či je stredná hodnota stranového rozdielu u oboch pozorovateľov zhodná.



Obr. 9.2: Znázornenie premennej vertikálny priemer v strede dĺžky tela kľúčnej kosti (simd)

9.3 Dátový súbor – párový test o strednej hodnote

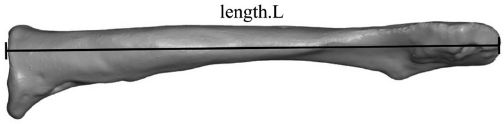
Hodnotený súbor: Hodnotený súbor predstavujú osteometrické dáta kľúčnej kosti (*clavicula*) anglického súboru dokumentovaných skeletov (Parsons, 1916). Konkrétne ide o dĺžku kľúčnej kosti z pravej a ľavej strany tela (pozri obrázok 9.3) v párovom usporiadaní. Jednotlivé kosti bez druhostrannej kosti neboli do súboru zaradené.

Súbor dát: paired-means-clavicle2.txt

Popis premenných:
id – poradové číslo jedinca;
sex – pohlavie (m – muž, f – žena);
length.R – dĺžka kosti z pravej strany (mm);
length.L – dĺžka kosti z ľavej strany (mm).

Biologické súvislosti: V prípade kľúčnej kosti sa ukazuje, že kosť z ľavej strany je v priemere tenšia a súčasne dlhšia než kosť z pravej strany. Rozdiel pravdepodobne vzniká v dôsledku odlišnej doby ukončenia rastu oboch strán v priebehu dospievania (ľavá kosť väčšinou rastie dlhšie; Mays a kol. (1999)) v kombinácii s odlišným zaťažením oboch strán v priebehu života.

Ciele:
(a) zistiť (pre každé pohlavie zvlášť), či je dĺžka kľúčnej kosti na ľavej strane tela dlhšia než na strane pravej.



Obr. 9.3: Znázornenie premennej dĺžka kosti kľúčnej z ľavej strany (length.L)

9.4 Dátový súbor – jednovýberový test o rozptyle

Hodnotený súbor: Z archívnych materiálov (Schmidt, 1888) máme k dispozícii pôvodné kraniometrické údaje o dĺžke lebky zo starovekej egyptskej populácie. Súčasne máme k dispozícii priemerné hodnoty oboch rozmerov, hodnoty smerodajnej odchýlky a počty prípadov vzorky z novovekej egyptskej populácie (dĺžka lebky: $\overline{x}_{\mathrm{m}} = 177.568$ mm, $\overline{x}_{\mathrm{f}} = 171.962$ mm; $s_{\mathrm{m}} = 7.526$ mm, $s_{\mathrm{f}} = 7.052$ mm; $n_{\mathrm{m}} = 88$, $n_{\mathrm{f}} = 52$ a šírka lebky: $\overline{x}_{\mathrm{m}}$=136.402 mm, $\overline{x}_{\mathrm{f}}$=131.038 mm; $s_{\mathrm{m}} = 6.411$ mm, $s_{\mathrm{f}} = 5.361$ mm; $n_{\mathrm{m}} = 87$, $n_{\mathrm{f}} = 52$).

Súbor dát: one-sample-variance-skull-mf.txt

Popis premenných:
id – poradové číslo;
pop – populácie (egant – egyptská staroveká);
sex – pohlavie (m – muž, f – žena);
skull.L – najväčšia dĺžka mozgovne (mm), t.j. priama vzdialenosť kraniometrických bodov *glabella* a *opisthocranion* (pozri obrázok 9.1);
skull.B – najväčšia šírka mozgovne (mm), t.j. vzdialenosť oboch kraniometrických bodov *euryon* (pozri obrázok 9.1).

Biologické súvislosti: Nezamietnutie zhody rozptylov dvoch porovnávaných súborov môže znamenať podobné biologické vlastnosti alebo obdobné vzorkovanie (výber vzorky z populácie).

Ciele:
(a) zistiť, či sa dĺžka lebky starovekej egyptskej populácie líši v rozptyle od novovekej egyptskej populácie (zvlášť u mužov a u žien).
(b) zistiť, či sa šírka lebky starovekej egyptskej populácie líši v rozptyle od novovekej egyptskej populácie (zvlášť u mužov a u žien).

9.5 Dátový súbor – jednovýberový test o korelačnom koeficiente

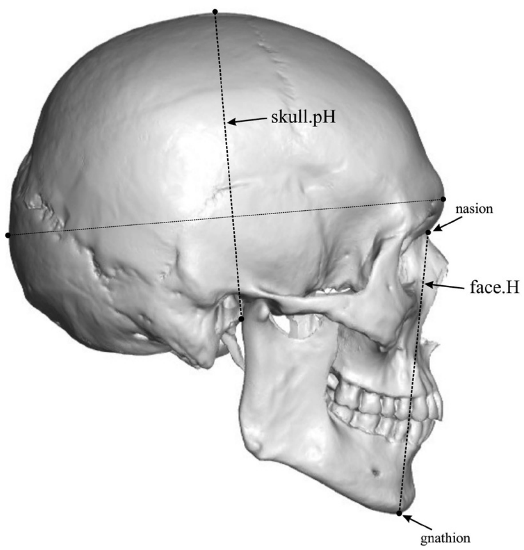
Hodnotený súbor: Vo vyššie uvedenom súbore (Schmidt, 1888) starovekej egyptskej populácie ďalej sledujeme súvislosti dvoch rozmerov tvárovej časti lebky (*splanchnocranium*) a mozgovej časti lebky (*neurocranium*); pozri obrázok 9.4. Súčasne máme k dispozícii hodnoty korelačného koeficientu medzi oboma rozmermi a údaje o počtoch prípadov zo vzorky novovekej egyptskej populácie ($r_{\mathrm{m}} = 0.251$, $r_{\mathrm{f}} = 0.144$; $n_{\mathrm{m}} = 30$, $n_{\mathrm{f}} = 19$).

Súbor dát: one-sample-correlation-skull-mf.txt

Popis premenných:
id – poradové číslo;
pop – populácie (egant – egyptská staroveká);
sex – pohlavie (m – muž, f – žena);
skull.pH – najväčšia výška mozgovne, projekčná vzdialenosť najvyššieho bodu mozgovne k najnižšiemu bodu lebečnej bázy v strednej rovine, kolmá na najväčšiu dĺžku mozgovne (mm);
face.H – morfologická výška tváre, vzdialenosť bodu *nasion* a *gnathion* (mm).

Biologické súvislosti: Rozmery oboch hlavných častí lebky sú počas vývinu riadené inými faktormi. Rast mozgovej časti lebky je spojený s rastom mozgu a prebieha predovšetkým v prvých siedmich či ôsmich rokoch po narodení, potom už rastie len málo. Intenzívny rast tvárovej časti lebky pokračuje aj v priebehu puberty a adolescencie. Napriek tomu spolu obe časti tvoria komplexný funkčný celok (Lieberman, 2011).

Ciele:
(a) zistiť, či výška mozgovej časti lebky súvisí s výškou tvárovej časti lebky (zvlášť u mužov a žien);
(b) porovnať, či je v tejto súvislosti rozdiel medzi starovekou a novovekou egyptskou populáciou.



Obr. 9.4: Znázornenie premenných najväčšia výška mozgovne (skull.pH) a morfologická výška tváre (face.H)

9.6 Dátový súbor – jednovýberový test o lineárno-uhlovom korelačnom koeficiente

Hodnotený súbor: Máme k dispozícii hodnoty troch lineárnych rozmerov a jedného uhla popisujúce výšku a šírku lebky a lebečnej bázy (pozri obrázok 9.5) vypočítané z pôvodných x, y a z súradníc štyroch význačných bodov (*bregma*, *basion*, *porion dx* a *porion sin*) digitalizovaných na 60 vybraných lebkách dospelých jedincov (40 mužov a 20 žien) z kostrovej zbierky z archeologickej lokality Pohansko – Pohřebiště okolo kostela (Jurda, 2008).

Súbor dát: lin-uhl-fm.txt

Popis premenných:
id – poradové číslo;
sex – pohlavie (m – muž, f – žena);
skull.H – výška lebky, vzdialenosť bodov *basion* a *bregma* (mm);
base.H – výška lebečnej bázy, minimálna vzdialenosť bodu *basion* k spojnici pravostranného a ľavostranného bodu *porion* (mm);
base.B – šírka lebečnej bázy na spojnici oboch bodov *porion* (mm);
base.A – uhol, ktorý zvierajú línie prechádzajúce bodom *basion* a pravostranným a ľavostranným bodom *porion* (stupne).

Biologické súvislosti: Sploštenie lebečnej bázy (*platybázia*) je jedným zo znakov používaných v bioarcheologických štúdiách ako indikátor stresu a horších životných podmienok pri rekonštruovaní a hodnotení životných podmienok minulých populácií. Výška lebečnej bázy je u ľudí trpiacich v priebehu vývinu nutričným stresom nižšia než u ľudí, ktorí boli nutričnému stresu vystavení v menšej miere. Podľa práce Angel (1982) ide o dôsledok deformácie nedostatočne nutrične zásobených kostí lebečnej bázy pod hmotnosťou vyššie položených oblastí hlavy a mozgu. Larsen (1997, 19–20) ale uvádza, že chrupavkové elementy lebečnej bázy majú vnútornú rastovú kapacitu dostatočne odolnú voči tlaku a uvedené vysvetlenie sploštovania bázy v dôsledku nutričného stresu nepovažuje za dostatočné.

Ciele:
(a) zistiť, či veľkosť uhla, ktorý zvierajú línia prechádzajúca bodom *basion* a pravostranným bodom *porion* s líniou prechádzajúcou bodom *basion* a ľavostranným bodom *porion* súvisí s niektorým zo sledovaných rozmerov na lebke (zvlášť u mužov a u žien);
(b) zistiť, či sa zistené závislosti líšia u mužov a u žien;
(c) zistiť, či veľkosť tohto uhla súvisí viac s výškou alebo šírkou bázy; ak s výškou bázy, zistiť, ako veľmi súvisí výška bázy s basion-bregmatickou výškou lebky.

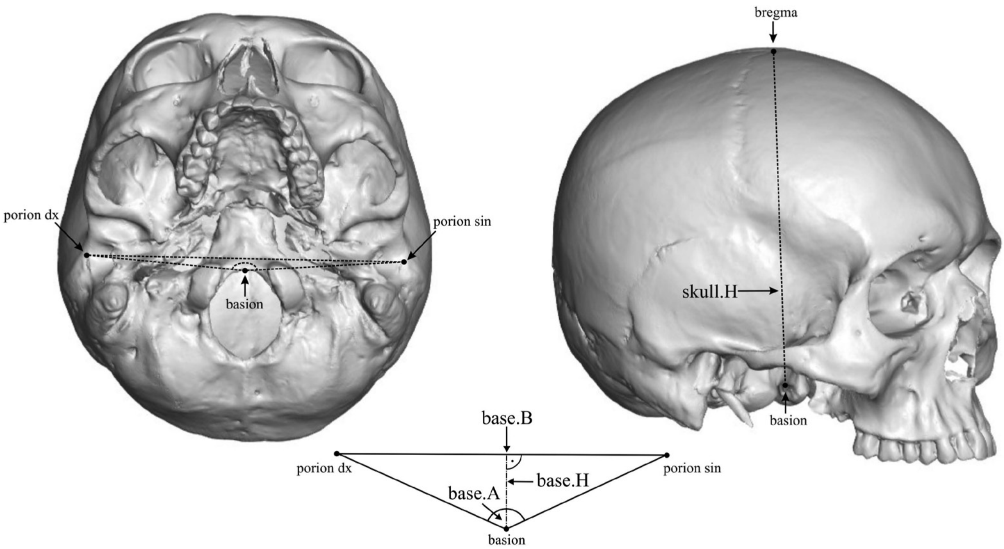
9.7 Dátový súbor – jednovýberový test o uhlovom korelačnom koeficiente

Hodnotený súbor: Máme k dispozícii číselné hodnoty dvoch uhlov na lebke (pozri obrázok 9.6) vypočítané z pôvodných x, y a z súradníc štyroch význačných bodov (*bregma*, *basion*, *nasion* a *prosthion*) digitalizovaných na 60 vybraných lebkách dospelých jedincov (40 mužov a 20 žien) z kostrovej zbierky z archeologickej lokality Pohansko – Pohřebiště okolo kostela (Jurda, 2008).

Súbor dát: uhl-uhl-fm.txt

Popis premenných:
id – poradové číslo;
sex – pohlavie (m – muž, f – žena);
front.A – uhol v bode *nasion*; uhol, ktorý zviera línia prechádzajúca bodmi *bregma* a *nasion* s líniou prechádzajúcou bodmi *nasion* a *basion* (stupne);
prog.A – uhol tvárového trojuholníka v bode *prosthion*; uhol, ktorý zviera línia prechádzajúca bodmi *basion* a *prosthion* s líniou prechádzajúcou bodmi *prosthion* a *nasion* (stupne).

Biologické súvislosti: Uhol nasia popisuje oblasť predného neurokrania, uhol tvárového trojuholníka v prosthiu vyjadruje stupeň alveolárneho prognatizmu hornej čeľuste a stupeň vývoja hornej časti



Obr. 9.5: Znázornenie premenných výška lebky (skull.H), výška lebečnej bázy (base.H), šírka lebečnej bázy (base.B) a uhol, ktorý zvierajú línie prechádzajúce oboma bodmi *porion* s vrcholom v bode *basion* (base.A)

splanchnokrania. I keď mozgová (*neurocranium*) a tvárová (*splanchnocranium*) časť lebky spolu tvoria komplexný funkčný celok, počas vývinu sú riadené inými faktormi (Lieberman, 2011).

Ciele:
(a) zistiť, či súvisí veľkosť uhla nasia s veľkosťou uhla tvárového trojuholníka v prosthiu (zvlášť u mužov a u žien);
(b) zistiť, či sa zistená závislosť líši u mužov a u žien.

9.8 Dátový súbor – jednovýberový test o pravdepodobnosti

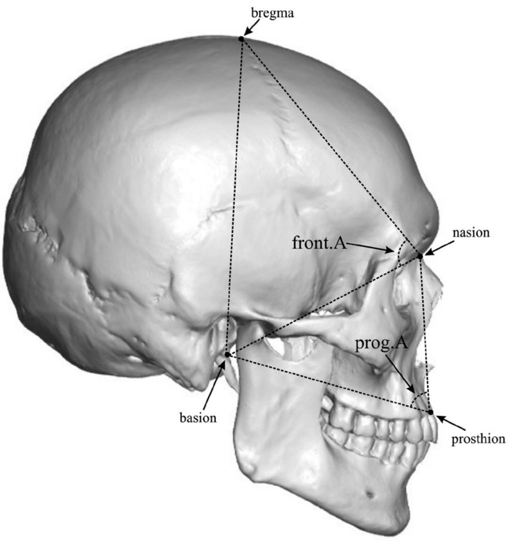
Hodnotený súbor: Sekundárny pomer pohlaví (šanca narodenia chlapca; anglicky *secondary sex ratio*, *odds ratio*) je pomer počtu novorodencov mužského pohlavia a počtu novorodencov ženského pohlavia. V priemere sa udáva hodnota 1.06 pre Českú republiku a 1.05 pre celosvetovú populáciu (Cheng a kol., 2007); sumárny údaj pre 220 národov z celého sveta). V priebehu jedného roka sa v krajskej nemocnici narodilo 729 chlapcov a 674 dievčat (Alánová, 2008).

Súbor dát: one-sample-probability-sexratio.txt

Popis premenných:
sex – pohlavie novorodenca (m – mužské, f – ženské).

Biologické súvislosti: Je známe, že v ľudských populáciách je sekundárny pomer pohlaví systematicky vychýlený v prospech chlapcov. Nerovná sa teda presne jednej (1.00), ako by to vyplývalo z princípu genetického určenia pohlavia pomocou pohlavných chromozómov. O príčinách a mechanizme vychýlenia sekundárneho pomeru pohlaví, rovnako ako o význame medzipopulačných rozdielov v pomere pohlaví sa diskutuje (James, 2006, 2008).

Ciele:
(a) testovať, či sa pomer pohlaví líši od vyrovnaného pomeru (1.00);
(b) či sa pomer pohlaví líši od celosvetového priemeru (1.05);
(c) či sa pomer pohlaví líši od tabelového priemeru pre Českú republiku (1.06).



Obr. 9.6: Znázornenie premenných uhol v bode *nasion* (front.A) a uhol tvárového trojuholníka v bode *prosthion* (prog.A)

9.9 Dátový súbor – jednovýberový test o pravdepodobnosti

Hodnotený súbor: Na základe publikovanej štúdie (Hauser, De Stefano a kol., 1989, str. 42–43) máme k dispozícii údaje o frekvencii výskytu epigenetického znaku *sutura metopica* (binárny znak) na lebkách Ainov z ostrova Hokkaido – v súbore 184 lebiek bol metopizmus zaznamenaný v 3.3 % prípadov. Súčasne máme k dispozícii percentuálne zastúpenie *sutura metopica* v japonskej populácii, ktorá dnes v Japonsku prevláda – metopizmus bol zaznamenaný v súbore 241 lebiek v 9.1 % prípadov (Mouri, 1976; Hauser, De Stefano a kol., 1989, str. 42–43).

Súbor dát: one-sample-probability-met.txt

Popis premenných:
origin – pôvod jedincov kostrového súboru (Ain – Ainovia, Jap – japonská populácia);
n – počet jedincov v jednotlivých súboroch;
met – počet jedincov s výskytom *sutura metopica*.

Biologické súvislosti: *Sutura metopica* (metopizmus) je epigenetický znak. Predstavuje v dospelosti perzistujúci šev (*sutura interfrontalis*) uprostred čelovej kosti, ktorý v priebehu ontogenézy oddeľuje jej pravú a ľavú časť a normálne sa uzatvára asi do dvoch rokov postnatálneho vývinu. Funkčne môže pretrvávanie švu spôsobovať dlhší rast čelovej kosti do šírky. Podľa niektorých štúdií majú lebky s metopizmom dlhšie transverzálne rozmery (Hauser, De Stefano a kol., 1989, str. 41). Na výskyte sa podieľa v určitej miere dedičnosť a medzi rôznymi populáciami existujú odlišnosti vo frekvencii tohoto znaku (od 0 do 16 %). Ainovia, najstarší obyvatelia severnej oblasti japonských ostrovov, ktorých pôvod je dodnes predmetom výskumu, predstavujú subpopuláciu v mnohých znakoch odlišnú od majoritnej japonskej populácie.

Ciele:
(a) zistiť, či sa či sa frekvencia výskytu *sutura metopica* u Ainov líši od frekvencie v prevládajúcej japonskej populácii.

9.10 Dátový súbor – dvojvýberový test o rozdiele stredných hodnôt

Hodnotený súbor: Máme k dispozícii údaje o pôrodnej hmotnosti prvorodených a druhorodených chlapcov, novorodencov narodených v krajskej nemocnici v priebehu jedného roka (Alánová, 2008).

Novorodencov narodených vo vyššom poradí sme z tohto porovnania vylúčili.

Súbor dát: two-samples-means-birth.txt

Popis premenných:
o.sib.N – počet starších súrodencov (0 – žiadny, 1 – jeden);
birth.W – pôrodná hmotnosť (g).

Biologické súvislosti: Z niektorých štúdií vyplýva, že medzi prvorodenými a druhorodenými novorodencami môžu byť rozdiely v pôrodnej hmotnosti. Prvorodení by potom mali mať nižšiu pôrodnú hmotnosť než deti narodené ako druhé v poradí (Seidman a kol., 1988; Swamy a kol., 2012).

Ciele:
(a) zistiť, či sa pôrodná hmotnosť prvorodených a druhorodených chlapcov z jednej pôrodnice a sezóny v priemere líši.

9.11 Dátový súbor – dvojvýberový test o rozdiele stredných hodnôt

Hodnotený súbor: Z archívnych materiálov (Schmidt, 1888) máme k dispozícii pôvodné kraniometrické údaje 215 dospelých mužov a 107 dospelých žien zo starovekej egyptskej populácie o basion–bregmatickej výške lebky (pozri obrázok 9.7). Súčasne máme k dispozícii priemerné hodnoty basion–bregmatickej výšky ($\overline{x}_{\mathrm{m}} = 133.977$ mm; $\overline{x}_{\mathrm{f}} = 126.942$ mm), hodnoty smerodajnej odchýlky ($s_{\mathrm{m}} = 5.171$ mm; $s_{\mathrm{f}} = 4.430$ mm) a počty prípadov ($n_{\mathrm{m}} = 87$, $n_{\mathrm{f}} = 52$) vzorky z novovekej egyptskej populácie.

Súbor dát: two-samples-means-skull.txt

Popis premenných:
id – poradové číslo;
pop – populácie (egant – egyptská staroveká);
sex – pohlavie (m – muž, f – žena);
skull.H – výška lebky, vzdialenosť bodov *basion* a *bregma* (mm).

Biologické súvislosti: Rozdiely medzi dospelými mužmi a ženami vo veľkosti lebky bývajú v ľudských populáciách systematicky posunuté v prospech mužov (väčšie hodnoty u mužov), i keď väčšinou nedosahujú takej miery ako v puberte a adolescencii výrazne rozvinutý dimorfizmus rozmerov postkraniálneho skeletu (Dadejová a kol., 2011). Miera dimorfizmu môže tiež kolísať medzi populáciami, pričom zmeny/rozdiely môžu byť spôsobené zmenami u jedného, druhého alebo oboch pohlaví.

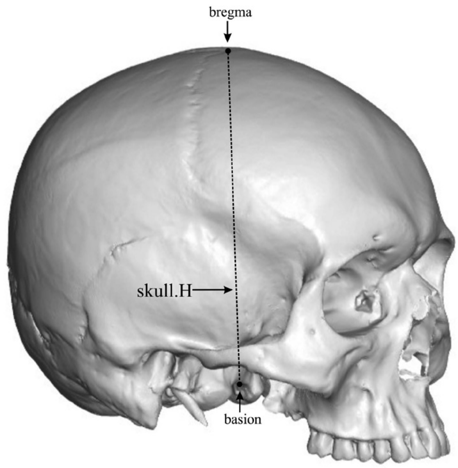
Ciele:
(a) vyjadriť mieru medzipohlavných rozdielov pomocou **indexu sexuálneho dimorfizmu**

$$\mathrm{ISD} = \frac{\overline{x}_{\mathrm{m}} - \overline{x}_{\mathrm{f}}}{\overline{x}_{\mathrm{m}}};$$

(b) zistiť, či sa stredné hodnoty basion–bregmatickej výšky lebky u mužov a u žien líšia;
(c) zistiť, či sa líšia stredné hodnoty basion–bregmatickej výšky starovekej od stredných hodnôt novovekej, zvlášť u mužov a zvlášť u žien.

9.12 Dátový súbor – dvojvýberový test o podiele rozptylov

Hodnotený súbor: Z archívnych materiálov (Schmidt, 1888) máme k dispozícii pôvodné kraniometrické údaje 215 dospelých mužov a 107 dospelých žien zo starovekej egyptskej populácie o basion–bregmatickej výške lebky (pozri obrázok 9.7). Súčasne máme k dispozícii priemerné hodnoty ($\overline{x}_{\mathrm{m}} = 133.977$ mm; $\overline{x}_{\mathrm{f}} = 126.942$ mm), hodnoty smerodajnej odchýlky ($s_{\mathrm{m}} = 5.171$ mm; $s_{\mathrm{f}} = 4.430$ mm) a počty prípadov ($n_{\mathrm{m}} = 87$, $n_{\mathrm{f}} = 52$) tohto znaku u vzorky z novovekej egyptskej populácie.



Obr. 9.7: Znázornenie premennej výška lebky (skull.H)

Súbor dát: two-samples-variances-skull.txt

Popis premenných:
id – poradové číslo;
pop – populácie (egant – egyptská staroveká);
sex – pohlavie (m – muž, f – žena);
skull.H – výška lebky, vzdialenosť bodov *basion* a *bregma* (mm).

Biologické súvislosti: Rozdiel v rozptyle rovnakého znaku v dvoch populáciách môže znamenať odlišnosť vo variabilite alebo odlišnosti v spôsobe výberu subjektov.

Ciele:
(a) zistiť, či sa muži a ženy zo starovekej egyptskej populácie líšia v rozptyle výšky lebky;
(b) zistiť, či sa muži a ženy z novovekej egyptskej populácie líšia v rozptyle výšky lebky;
(c) zistiť, či sa muži zo starovekej a novovekej egyptskej populácie líšia v rozptyle výšky lebky;
(d) zistiť, či sa ženy zo starovekej a novovekej egyptskej populácie líšia v rozptyle výšky lebky.

9.13 Dátový súbor – dvojvýberový test o rozdiele korelačných koeficientov

Hodnotený súbor: Máme k dispozícii súbor hodnôt dĺžky trupu (rozdiel akromiálnej a spinálnej výšky tela) a dĺžky dolnej končatiny (spinálna výška tela) mladých dospelých jedincov (pozri obrázok 9.8; zdroj [1]), prevažne študentov vysokých škôl z Brna a Ostravy (Králík, nepublikované dáta).

Súbor dát: two-samples-correlations-trunk.txt

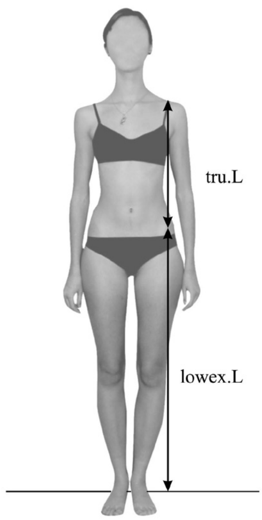
Popis premenných:
sex – pohlavie (m – muž, f – žena);
lowex.L – dĺžka dolnej končatiny (mm);
tru.L – dĺžka trupu (mm).

Biologické súvislosti: Rozmery trupu sú u žien priestorovo zásadné z hľadiska tehotenstva a miesta pre rastúci plod. Muži takéto obmedzenia nemajú. Preto napríklad u žien prevláda dýchanie hrudné,

[1] Atlas somatoskopických znaků člověka (http://www.sci.muni.cz/somatoskopie), upravené

zatiaľ čo u mužov brušné. So zväčšujúcou sa veľkosťou tela v priebehu rastu v puberte a adolescencii by preto mala u dospelých žien narastať úmernejšie aj veľkosť trupu. U mužov by tendencia nemusela byť tak silná. Možno teda predpokladať, že závislosť veľkosti oboch častí tela bude u žien silnejšia než u mužov.

Ciele:
(a) zistiť, či u mladých dospelých ľudí, bez ohľadu na pohlavie, súvisí dĺžka dolnej končatiny s dĺžkou trupu;
(b) zistiť, či sa ženy a muži v miere tejto závislosti líšia.



Obr. 9.8: Znázornenie premenných dĺžka dolnej končatiny (lowex.L) a dĺžka trupu (tru.L)

9.14 Dátový súbor – dvojvýberový test o rozdiele pravdepodobností

Hodnotený súbor: Máme k dispozícii údaje o počte prvorodených a druhorodených chlapcov a dievčat, novorodencoch narodených v krajskej nemocnici v priebehu jedného roka (Alánová, 2008). Novorodencov narodených vo vyššom poradí sme z tohto porovnania vylúčili.

Súbor dát: two-samples-probabilities-sexratio.txt

Popis premenných:
sex – pohlavie novorodenca (m – mužské, f – ženské);
o.sib.N – počet starších súrodencov (0 – žiadny, 1 – jeden).

Biologické súvislosti: Sekundárny pomer pohlaví je u človeka systematicky vychýlený v prospech chlapcov. Podľa niektorých evolučných teórií nie je pohlavie potomkov u vtákov a cicavcov náhodné, pretože naň nenáhodne vplýva prostredie a pohlavie predchádzajúcich potomkov regulujúce reprodukčnú úspešnosť rodičov. U človeka niektoré štúdie zaznamenali súvislosti sekundárneho pomeru pohlaví s paritou (Juntunen a kol., 1997), iné však nie (Almagor a kol., 1998).

Ciele:
(a) zistiť, či je podiel chlapcov a dievčat u skupiny prvorodených a druhorodených detí odlišný.

9.15 Dátový súbor – viacvýberový test o stredných hodnotách, nominálna premenná

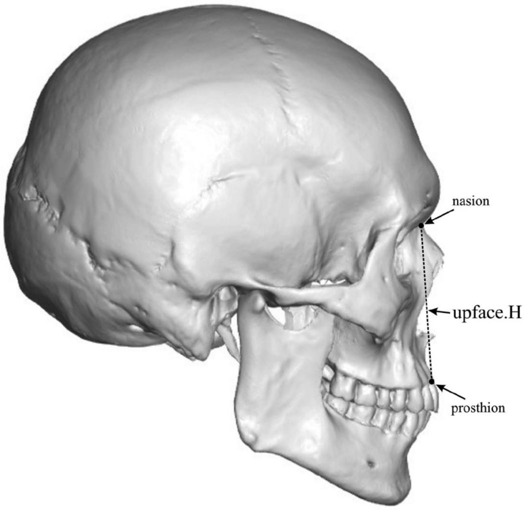
Hodnotený súbor: Z archívnych materiálov (Schmidt, 1888) máme k dispozícii pôvodné kraniometrické údaje o výške hornej časti tváre mužov (pozri obrázok 9.9) z piatich populácií – nemeckej (19 jedincov), malajskej (69 jedincov), čínskej (18 jedincov), peruánskej (44 jedincov) a bantuskej (13 jedincov).

Súbor dát: anova-means-skull.txt

Popis premenných:
id – poradové číslo;
pop – populácie (nem – nemecká, mal – malajská, cin – čínska, per – peruánska, ban – bantuská);
sex – pohlavie (m – muž);
upface.H – výška hornej časti tváre, priama vzdialenosť medzi bodmi *nasion* a *prosthion* (mm).

Biologické súvislosti: Výška hornej časti tváre je jedným z rozmerov, ktoré sa uplatňujú pri hodnotení populačnej afinity (odhade etnickej príslušnosti) vo forenzných aplikáciách, kde sa obvykle rozlišujú tri hlavné kategórie – kaukazoidná, negroidná a mongoloidná. I keď hlavným ukazovateľom populačnej afinity na lebke sú jej morfologické rysy a nie metrické znaky, môžeme vo všeobecnosti povedať, že pre kaukazoidný fenotyp je charakteristická stredná až vysoká a často široká lebka s rôzne širokou a stredne vysokou tvárou; pre negroidný fenotyp dlhá, úzka a nízka lebka s širokou a nízkou tvárou; a pre mongoloidný fenotyp zas dlhá, široká a stredne vysoká lebka s veľmi širokou a vysokou tvárou.

Ciele:
(a) zistiť, či sú u mužov zo sledovaných populácií rozdiely v strednej hodnote výšky hornej časti tváre;
(b) zistiť, či prípadné rozdiely odpovedajú tradične uvádzaným rozdielom (najnižšia tvár u negroidných populácií, stredne vysoká u kaukazoidných populácií a vysoká u mongoloidných populácií).



Obr. 9.9: Znázornenie premennej výška hornej časti tváre (upface.H)

9.16 Dátový súbor – viacvýberový test o stredných hodnotách, nominálna premenná

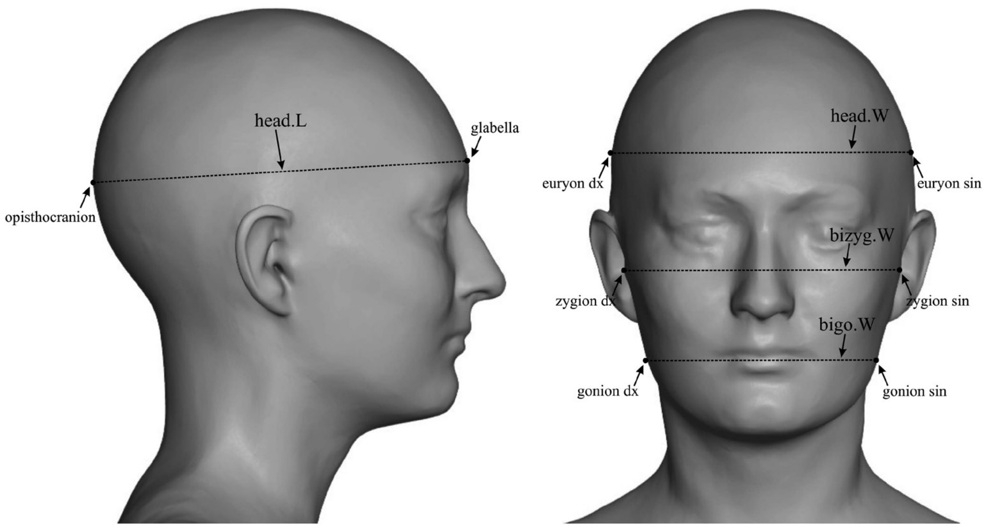
Hodnotený súbor: Máme k dispozícii antropometrické údaje mladých dospelých ľudí, prevažne študentov vysokých škôl z Brna a Ostravy (Králík, nepublikované dáta), konkrétne údaje o rozmeroch hlavy (pozri obrázok 9.10; zdroj[2]), a súčasne zaradenie prípadov do kategórií podľa pohlavia, sexuálnej orientácie a počtu vlastných súrodencov.

Súbor dát: anova-head.txt

Popis premenných:
sex – pohlavie (m – muž, f – žena);
sexor – sexuálna orientácia (op – výlučne na opačné pohlavie, sa – ostatné, t.j. iné než výlučne na opačné pohlavie (bisexuálna, homosexuálna));
obra – existencia staršieho biologického brata (yes – jedinec má staršieho brata, no – jedinec nemá staršieho brata);
body.H – výška postavy (mm);
head.L – dĺžka hlavy, vzdialenosť medzi bodmi *glabella* a *opisthocranion* (mm);
head.W – šírka hlavy, vzdialenosť oboch bodov *euryon* (mm);
bigo.W – šírka dolnej čeľuste, vzdialenosť oboch bodov *gonion* (mm);
bizyg.W – šírka tváre, vzdialenosť oboch bodov *zygion* (mm).

Biologické súvislosti: Pohlavia sa obvykle líšia celým radom telesných rozmerov. Okrem niektorých reprodukčne významných rozmerov v oblasti panvy majú ženy obvykle väčšinu rozmerov menšiu než muži odpovedajúceho veku z rovnakej populácie. Je však známe, že modifikujúci vplyv na tie isté rozmery má celý rad ďalších faktorov, napríklad sexuálna orientácia, poradie narodenia, počet starších súrodencov atď. (i keď sa v miere vplyvu týchto modifikujúcich faktorov rôzne štúdie líšia).

Ciele:
(a) zistiť, či sú jednotlivé rozmery hlavy sexuálne dimorfné, t.j. či sa stredné hodnoty jednotlivých rozmerov medzi mužmi a ženami líšia;
(b) zistiť, či existuje vplyv pohlavia na jednotlivé rozmery hlavy modifikovaný vplyvom sexuálnej orientácie;
(c) zistiť, či existuje vplyv pohlavia na jednotlivé rozmery hlavy modifikovaný vplyvom existencie biologického staršieho brata.



Obr. 9.10: Znázornenie premenných dĺžka hlavy (head.L), šírka hlavy (head.W), šírka dolnej čeľuste (bigo.W) a šírka tváre bizyg.W

[2] Atlas somatoskopických znaků člověka (http://www.sci.muni.cz/somatoskopie), upravené

9.17 Dátový súbor – viacvýberový test o stredných hodnotách, ordinálna premenná

Hodnotený súbor: Máme k dispozícii dáta o pôrodnej hmotnosti v súbore novorodencov z okresnej nemocnice za jeden rok a súčasne údaje o parite matiek (poradie narodenia jednotlivých novorodencov) a vzdelaní matiek (Alánová, 2008).

Súbor dát: anova-newborns.txt

Popis premenných:
edu.M – vzdelanie matky (1 – základné, 2 – stredné bez maturity, 3 – stredné s maturitou, 4 – vysokoškolské);
prch.N – počet predchádzajúcich detí biologickej matky hodnoteného jedinca (hodnoty od 0 po 8, sledujeme však iba tri kategórie: 1 – prvorodené, 2 – druhorodené a 3 – spoločne deti narodené v treťom a ďalšom poradí);
sex.C – pohlavie hodnoteného dieťaťa (m – mužské, f – ženské);
weight.C – pôrodná hmotnosť hodnoteného dieťaťa (g).

Biologické súvislosti: Hmotnosť novorodenca je dôležitý údaj z hľadiska perspektívy jeho ďalšieho vývinu. Závisí na celom rade faktorov, modifikujúci vplyv môže mať napr. parita a vzdelanie matky.

Ciele:
(a) zistiť, či závisí pôrodná hmotnosť chlapca na poradí, v ktorom sa biologickej matke narodilo;
(b) zistiť, či súvisí pôrodná hmotnosť chlapca so vzdelaním matky;
(c) zistiť, či závisí pôrodná hmotnosť dievčaťa na poradí, v ktorom sa biologickej matke narodilo;
(d) zistiť, či súvisí pôrodná hmotnosť dievčaťa so vzdelaním matky.

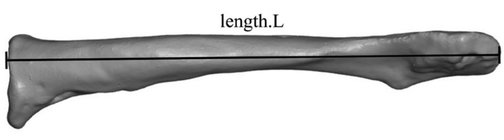
9.18 Dátový súbor – viacvýberový test o rozptyloch

Hodnotený súbor: Máme k dispozícii osteometrické dáta o dĺžke kľúčnej kosti (*clavicula*) z pravej strany u štyroch súborov – výberov zo štyroch rôznych populácií: anglickej, dvoch indických a gréckej (pozri obrázok 9.11).

Súbor dát: more-samples-variances-clavicle.txt

Popis premenných:
population – populácia, z ktorej kostrový súbor pochádza: eng – anglická populácia (dáta Parsons, 1916), ind1 – indická populácia z Amritsaru (dáta Jit a Singh, 1966), ind2 – indická populácia z Varanasi (dáta Singh a Gangrade, 1968), gre – grécka populácia z Atén (dáta Králík a kol., 2014);
sex – pohlavie (m – muž);
cla.L – najväčšia dĺžka kosti kľúčnej z pravej strany (mm).

Biologické súvislosti: Osteologické vzorky porovnávaných kostí kľúčnych vznikali v rôznych obdobiach a predstavujú rôznym spôsobom vybraných jedincov z rôzne rozsiahlych populácií. Je preto možné, že sa súbory budú líšiť svojim rozptylom.
Ciele:
(a) zistiť, či sa štyri porovnávané populácie líšia v rozptyloch dĺžky kľúčnej kosti z pravej strany.



Obr. 9.11: Znázornenie premennej najväčšia dĺžka kosti kľúčnej z pravej strany (cla.L)

9.19 Dátový súbor – viacvýberový test o korelačných koeficientoch

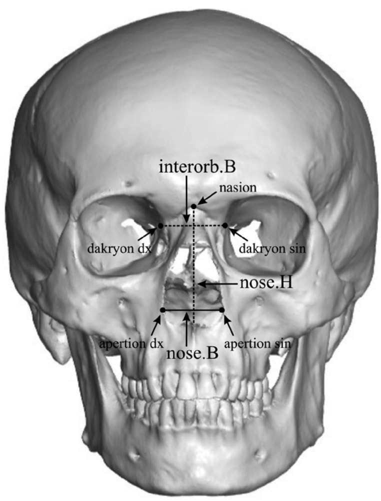
Hodnotený súbor: Vo vyššie uvedenom súbore (Schmidt, 1888) mužov z piatich rôznych populácií – nemeckej (20 jedincov), malajskej (73 jedincov), čínskej (19 jedincov), peruánskej (46 jedincov) a bantuskej (14 jedincov) – sledujeme súvislosti medzi rozmermi tvárovej časti lebky (*splanchnocranium*), konkrétne medzi výškou a šírkou nosa a medzi šírkou nosa a interorbitálnou šírkou (pozri obrázok 9.12).

Súbor dát: more-samples-correlations-skull.txt

Popis premenných:
id – poradové číslo;
pop – populácie (nem – nemecká, mal – malajská, cin – čínska, per – peruánska, ban – bantuská);
sex – pohlavie (m – muž);
nose.H – výška nosa, t.j. priama vzdialenosť medzi bodom *nasion* a najnižším bodom *apertura piriformis* (mm);
nose.B – šírka nosa, t.j. najväčšia šírka *apertura piriformis*, medzi pravým a ľavým bodom *apertion* (mm);
interorb.B – interorbitálna šírka odpovedajúca priamej vzdialenosti medzi pravým a ľavým bodom *dakryon* (mm).

Biologické súvislosti: Výška a šírka nosa a interorbitálna šírka patria medzi rozmery, ktoré sa uplatňujú pri hodnotení populačnej afinity (odhade etnickej príslušnosti) vo forenzných aplikáciách, kde sa obvykle odlišujú tri hlavné kategórie – kaukazoidná, negroidná a mongoloidná. I keď hlavným ukazovateľom populačnej afinity na lebke sú jej morfologické rysy a nie metrické znaky, vo všeobecnosti možno povedať, že pre kaukazoidný fenotyp je charakteristický vysoký a úzky nosný otvor; pre negroidný fenotyp široký a nízky nosný otvor a široká interorbitálna oblasť; a pre mongoloidný fenotyp široký nosný otvor.

Ciele:
(a) zistiť, či u jednotlivých skupín súvisí výška so šírkou nosa;
(b) zistiť, či u jednotlivých skupín súvisí interorbitálna šírka so šírkou nosa;
(c) zistiť, či sa sledované populácie v miere týchto súvislostí líšia.



Obr. 9.12: Znázornenie premenných výška nosa (nose.H), šírka nosa (nose.B) a interorbitálna šírka (interorb.B)

9.20 Dátový súbor – viacvýberový test o pravdepodobnostiach

Hodnotený súbor: Na základe publikovanej štúdie (Stewart, 1970) máme k dispozícii údaje o frekvencii výskytu troch stupňov zmien kostného reliéfu na vnútornej strane lonovej kosti (*os pubis*) v blízkosti lonovej spony (*symphysis pubica*) u žien z troch kostrových súborov: európskeho pôvodu, afrického pôvodu a Inuitov (pozri obrázok 9.13). Súčasne máme k dispozícii počty jedincov každého z týchto troch súborov.

Súbor dát: more-samples-probabilities-pubis.txt

origin	absence	trace.to.small	moderate.to.large	number.of.cases
European	30	20	10	60
African	56	37	17	110
Inuits	16	6	13	35

Popis premenných:
origin – pôvod jedincov kostrového súboru (European – európsky, African – africký, Inuits – Inuiti;
absence – početnosti prípadov s neprítomnosťou zmien kostného reliéfu;
trace.to.small – početnosti prípadov so stopami zmien až malými zmenami kostného reliéfu;
moderate.to.large – početnosti prípadov so strednými až výraznými zmenami kostného reliéfu;
number.of.cases – početnosti prípadov v jednotlivých kostrových vzorkách.

Biologické súvislosti: Popôrodné zmeny na kostre sú spornou otázkou, ktorej riešenie má viac než 140 rokov trvajúcu históriu (cf. Mikešová, 2008). Predovšetkým sú s tehotenstvom a/alebo pôrodom spájané zmeny kostného reliéfu na kosti panvovej – tzv. *sulcus praeauricularis* pod dolným okrajom *facies auricularis* a obdobné zmeny na vnútornej strane lonovej kosti v blízkosti *symphysis pubica*. Etiológia týchto zmien však nie je dosiaľ známa. Zreteľne vytvorený *sulcus praeauricularis* vo forme špecifických jamkovitých depresií sa vyskytuje iba u žien, ktoré rodili. Neprítomnosť týchto zmien ale neznamená, že ide o ženu, ktorá nerodila. Môže ísť o muža, o ženu, ktorá nerodila, o ženu, ktorá rodila a zmeny sa u nej nevytvorili alebo sa vytvorili, ale časom došlo k ich vymiznutiu v dôsledku remodelácie kosti. Tieto útvary kostného reliéfu teda nemožno použiť ani na odhad pohlavia (prítomnosť vs. neprítomnosť), ani na stanovenie parity (miera prejavu). Okrem toho sa ženy z rôznych populácií môžu vo výskyte týchto útvarov líšiť.

Ciele:
(a) zistiť, či sa tri sledované populácie líšia vo výskyte každej z troch foriem kostného reliéfu na zadnej (vnútornej) strane kosti lonovej.

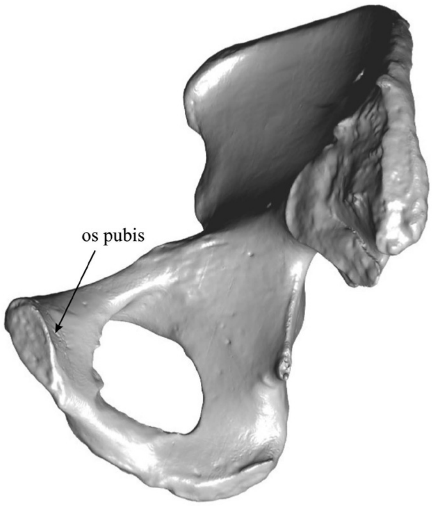
9.21 Dátový súbor – viacvýberový test o pravdepodobnostiach

Hodnotený súbor: Pri vyšetrení očnej dúhovky 60 mužov a 60 žien (Horníčková, 1992) bol hodnotený farebný odtieň dúhovky a výskyt rôznych typov útvarov v štruktúre dúhovky. Farba očí bola posudzovaná pomocou lupy podľa vzorkovnice R. Martina (Martin, 1914/1928), ktorá obsahuje 16 farebných odtieňov, a následne kategorizovaná do jednej zo štyroch skupín. Súčasne bol podľa schémy Ziegelmayera (Martin a Saller, 1957–1966, str. 400) hodnotený výskyt koncentrických, kryptovitých a lúčovitých útvarov v štruktúre dúhovky. K dispozícii máme početnosti jednotlivých farebných kategórií pre obe pohlavia a početnosti dúhoviek s lúčovitými útvarmi v jednotlivých farebných kategóriách.

Súbor dát: multinom-iris-color.txt

	H	HZ	ME	M	L.H	L.HZ	L.ME	L.M
m	13	11	24	12	7	5	9	2
f	16	12	26	6	9	7	2	1
sum	29	23	50	18	16	12	11	3

Popis premenných:
sex – pohlavie jedincov vo vzorke (m – muži, f – ženy, sum – obe pohlavia);
iris.C – farba očnej dúhovky (H – hnedé oči (odtiene 2 – 4), HZ – hnedozelené oči (odtieň 5), ME – melírované oči (odtiene 6 – 8 a 12 – 15), M – modré oči (odtiene 9 – 11));



Obr. 9.13: Kostný reliéf na vnútornej strane *os pubis*

iris.S – štruktúra dúhovky (L.H – hnedé oči s lúčovitými útvarmi, L.HZ – hnedozelené oči s lúčovitými útvarmi, L.ME – melírované oči s lúčovitými útvarmi, L.M – modré oči s lúčovitými útvarmi).

Biologické súvislosti: Farba očí (dúhovky) je daná z veľkej časti dedične. U svetlejších očí (modrých, šedých a zelených) závisí farba očí predovšetkým od množstva pigmentu v dúhovkovej časti sietnice (*pars iridica retinae*), u hnedých očí obsahuje melanocyty aj dúhovková trámčina (*stroma iridis*) a u afrických populácií aj predná hraničná vrstva dúhovky, a to vo veľmi veľkom množstve. Farba dúhovky teda súvisí s obsahom dúhovkovej trámčiny. Jej vnútorná štruktúra však nie je homogénna. Vyskytujú sa v nej rôzne koncentrické, kryptovité (otvory) a radiálne (lúčovité) útvary, ktorých prítomnosť môže do určitej miery súvisieť s pigmentáciou dúhovkovej trámčiny.

Ciele:
(a) zistiť, či medzi farebnými kategóriami existujú štatisticky významné rozdiely vo výskyte lúčovitých útvarov dúhovky;
(b) zistiť, či existujú rozdiely vo výskyte lúčovitých útvarov dúhovky medzi pohlaviami.

9.22 Dátový súbor – viacvýberový test o pravdepodobnostiach

Hodnotený súbor: Vo vzorke, ktorú tvorilo 200 študentov (100 mužov a 100 žien), boli štandardnou dermatoglyfickou metodikou snímané dermatoglyfy dlane (Býmová, 1990). Na odtlačkoch bolo hodnotené zakončenie troch hlavných dlaňových línií (D, C a B); pozri obrázok 9.14. Prípady boli podľa vzorca zakončenia (vyústenia proximálnych radiant digitálnych trirádií na štandardne číslovaných polohách okraja dlane) rozdelené do troch kategórií. Súčasne bola hodnotená farba vlasov podľa štandardnej Fischer-Sallerovej (Martin a Saller, 1957–1966, s. 391) stupnice 30 odtieňov, ktoré boli rozdelené do 3 skupín. K dispozícii máme početnosti jedincov v jednotlivých kategóriách, zvlášť pre mužov a ženy.

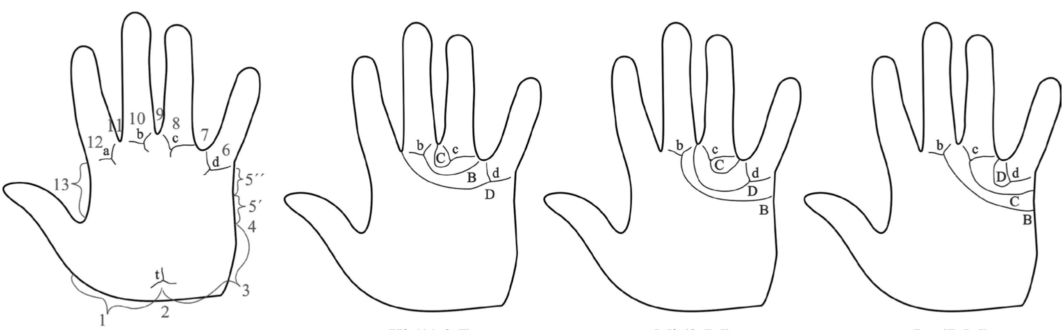
Súbor dát: multinom-palmar-lines.txt

m	Hi	Mi	Lo	f	Hi	Mi	Lo
LiH	6	6	4	LiH	4	6	6
MH	20	15	7	MH	18	10	10
DaH	18	12	12	DaH	12	22	12

Popis premenných:
sex – pohlavie (m – muži, f – ženy);
palmar.lines – zakončenie troch dlaňových línií (Hi – vysoké (najčastejší vzorec 11 9 7), Mi – stredné (najčastejší vzorec 9 7 5), Lo – nízke (najčastejší vzorec 7 5 5));
hair.C – farba vlasov (LiH – svetlé, MH – stredné, DaH – tmavé).

Biologické súvislosti: Zakončenie hlavných dlaňových línií na okraji dlane odráža celkový smer priebehu epidermálnych líšt na dlani, ktorý môže byť niekde medzi transverzálnym a longitudinálnym. Vysoké polohy odpovedajú prevažne transverzálnemu priebehu, nižšie polohy reprezentujú prevažne šikmý až longitudinálny priebeh epidermálnych líšt dlaní. Keďže sa epidermálne lišty vytvárajú prenatálne od 3. mesiaca tehotenstva, finálny priebeh epidermálnych líšt v dospelosti odráža epigenetické procesy z tohto obdobia. Približne v rovnakej dobe prebieha migrácia melanocytov z neurálnej lišty do pokožky (*epidermis*) a rozvíjajú sa vlasové folikuly, ktoré predstavujú deriváty kože na iných miestach tela (mimo papilárny terén). Nie je však jasné, akým spôsobom spolu epigenetické procesy v papilárnom teréne kože a mimo papilárneho terénu súvisia. Súčasne je známe, že rozdielne ľudské populácie sa líšia ako v priebehu dlaňových línií (obyvatelia Indie a Predného východu majú skôr vysoké zakončenie, pôvodní obyvatelia rovníkovej Afriky prevažne nízke a obyvatelia východnej Ázie prechodné), tak aj vo farbe kože. Medzipopulačné rozdiely však nemusia byť plne obsiahnuté v rozdieloch medzi ľuďmi v rámci jednej populácie, akokoľvek sú variabilní v študovaných znakoch. Je teda otázne, či spolu môžu súvisieť priebeh hlavných dlaňových línií a farba vlasov u Európanov s veľmi variabilnou farbou vlasov.

Ciele:
(a) zistiť, či existuje vzťah medzi farbou vlasov a zakončením hlavných dlaňových línií (zvlášť u mužov a u žien);
(b) zistiť, či sa zakončenie troch dlaňových línií líši medzi skupinami s odlišnou farbou vlasov;
(c) zistiť, či sa pohlavia líšia vo farbe vlasov;
(d) zistiť, či sa početnosti zakončenia hlavných dlaňových línií líšia;
(e) zistiť, pri akom minimálnom počte jedincov vo vzorke by rovnaké proporcionálne zastúpenie v jednotlivých kategóriách a jednotlivých porovnaniach dosiahlo štatistickú významnosť na hladine významnosti $\alpha = 0.05$ a pri sile testu $1 - \beta = 0.8$.



Obr. 9.14: Číslovanie pozíc na okraji dlane a príklady vysokého (Hi), stredného (Mi) a nízkeho (Lo) zakončenia troch hlavných dlaňových línií (D, C a B)

9.23 Dátový súbor – viacvýberový test o pravdepodobnostiach

Hodnotený súbor: Vo vzorke, ktorú tvorilo 120 študentov vysokej školy (60 mužov a 60 žien) bola okrem iného hodnotená priľahlosť ušnice (Štouračová, 1992) a sledované medzipohlavné a stranové rozdiely v tomto znaku. Prípady boli podľa priľahlosti ušnice rozdelené do troch kategórií (Titlbachová, 1967, s. 97) zvlášť na pravej a ľavej strane. K dispozícii máme početnosti jedincov v jednotlivých kategóriách.

Súbor dát: multinom-earlobe.txt

m	R	L	f	R	L
1	22	22	1	38	36
2	34	34	2	21	24
3	4	4	3	1	0

Popis premenných:
sex – pohlavie (m – muži, f – ženy);
side – strana tela (R – pravá, L – ľavá);
earlobe – priľahlosť ušnice (1 – priľahlé, 2 – stredne priľahlé, 3 – odstávajúce).

Biologické súvislosti: Vo veľkosti, tvare a ďalších vlastnostiach ušnice (*auricula*) sa človek líši od ostatných primátov. Už u vyšších primátov má ušnica len rudimentárnu pohyblivosť (za zvukom sa otáča celá hlava), stále sa však uplatňuje pri zachytávaní akustických vĺn. Ušnica človeka má skrátenú pozdĺžnu os, okraj zvinutý v tzv. *helix*, rozšírené pripojenie k hlave a vytvorený ušný lalôčik (*lobulus auriculae*). Podľa niektorých výskumov (Hulanicka, 1973) sa však uši mužov a žien v niektorých vlastnostiach systematicky líšia. Uši napríklad odstávajú viac mužom než ženám. Nie je však jasné, či je tento rozdiel vedľajším dôsledkom dimorfizmu ontogenézy a výsledného tvaru hlavy alebo odráža funkčné rozdiely v sluchovej percepcii medzi pohlaviami (napr. obvyklý charakter a smer sluchových signálov). Vylúčený ale nie je ani kultúrne podmienený dlhodobý vplyv odlišnej pokrývky hlavy (šatky u žien a klobúky u mužov). Pri hľadaní odpovede treba najskôr zistiť, či je dimorfizmus v priľahlosti ušnice univerzálny jav platný aj v iných ľudských populáciách, t.j. overiť existenciu dimorfizmu v priľahlosti ušnice na iných súboroch.

Ciele:
(a) zistiť, či existuje sexuálny dimorfizmus v priľahlosti ušnice (zvlášť na pravej a ľavej strane);
(b) zistiť, pri akom minimálnom počte jedincov vo vzorke by rovnaké proporcionálne zastúpenie v jednotlivých kategóriách dosiahlo štatistickú významnosť na hladine významnosti $\alpha = 0.05$ a pri sile testu $1 - \beta = 0.8$.

9.24 Dátový súbor – homogenita vektorov pravdepodobností

Hodnotený súbor: V nezávislých výskumoch boli zistené frekvencie krvných skupín AB0 systému (skupina 0, A, B a AB) u 400 obyvateľov Košíc (Slovensko) a 500 obyvateľov Prahy (pozri Vondrušková, 1983). K dispozícii máme početnosti v jednotlivých kategóriách.

Súbor dát: multinom-blood-groups.txt

	0	A	B	AB
K	138	147	84	31
P	209	184	81	26

Popis premenných:
city – mesto (K – Košice, P – Praha);
blood – krvná skupina AB0 systému (0, A, B, AB).

Biologické súvislosti: Krvné skupiny sú vrodené vlastnosti povrchu erytrocytov (proteíny, glykoproteíny alebo glykolipidy), detegovateľné pomocou aloprotilátok. Do dnešnej doby bolo rozlíšených minimálne 30 systémov krvných skupín, z ktorých väčšina je polymorfných (ľudia sa vzájomne líšia kombináciou alel podmieňujúcich príslušné antigény). Polymorfizmus krvných skupín je podľa dnešných poznatkov udržovaný prírodným výberom a predstavuje výsledok adaptácií minulých populácií na pôsobenie bakteriálnych patogénov. Napríklad relatívne nízka frekvencia skupiny 0 (AB0 systému) v mnohých oblastiach Európy a Ázie môže byť dôsledkom devastujúcich epidémií (*Yersiniapestis*). Antigén na povrchu tohto mikroorganizmu je podobný H antigénu krvnej skupiny 0. Ľudia s krvnou skupinou 0 neprodukujúci žiadne protilátky anti-H pravdepodobne tento patogén nerozpoznali, čo u nich mohlo viesť k horšiemu priebehu choroby a vyššej úmrtnosti (Mielke a kol., 2011). Frekvencie foriem v rámci daného systému (napr. AB0 systému) sa naprieč geografickým územím mení *klinálne*, t.j. existujú geografické gradienty v zastúpení jednotlivých foriem (A, B, AB a 0).

Ciele:
(a) zistiť, či sa obyvatelia Prahy a Košíc líšia v zastúpení krvných skupín AB0 systému.

10 Funkcie

V tabuľkách 10.1, 10.2, 10.3, 10.4, 10.5 a 10.6 sa nachádza tabuľkový prehľad použitých a naprogramovaných funkcií v R a ich stručné definície.

Tabuľka 10.1: Prehľad základných funkcií – charakteristiky polohy a variability a matematické funkcie

funkcia	**popis**
cor()	korelačný koeficient r
length()	rozsah n
mad()	priemer absolútnych odchýlok MAD
max()	maximum $x_{\max}$
mean()	aritmetický priemer $\overline{x}$
median()	medián $\widetilde{x}_{0.5}$
min()	minimum $x_{\min}$
quantile()	vybrané kvantily, napr. $\widetilde{x}_{0.25}$, $\widetilde{x}_{0.5}$ a $\widetilde{x}_{0.75}$
range()	rozpätie D
sd()	smerodajná odchýlka s_x
summary()	päťčíselný súhrn $(x_{\min}, \widetilde{x}_{0.25}, \widetilde{x}_{0.5}, \widetilde{x}_{0.75}, x_{\max})$
var()	rozptyl s_x^2
abs()	absolútna hodnota
atan()	arkus tangens
atanh()	hyperbolický arkus tangens
ceiling()	najbližsie vyššie celé číslo
colMeans()	aritmetický priemer čísel v matici po stĺpcoch
colSum()	suma čísel v matici po stĺpcoch
cos()	kosínus uhla v radiánoch
cumsum()	kumulatívna suma
cut()	rozdelenie spojitej premennej na intervaly
density()	hustota
dim()	rozmery objektu
exp()	exponenciálna funkcia
floor()	celá časť čísla
log()	prirodzený logaritmus
na.omit()	vyradenie chýbajúcich pozorovaní ozn. NA
order()	výpis poradia elementov vektora
round()	zaokrúhľovanie
rowMeans()	aritmetický priemer čísel v matici po riadkoch
rowSum()	suma čísel v matici po riadkoch
scale()	škálovanie dát
sin()	sínus uhla v radiánoch
sort()	zoradenie elementov vektora podľa veľkosti
sqrt()	odmocnina čísla
sum()	suma členov vektora
tanh()	hyperbolický tangens

Tabuľka 10.2: Prehľad funkcií súvisiacich s dátovým manažmentom a funkcia library()

funkcia	popis
addmargins()	pridanie okrajových početností do tabuľky početností
apply()	aplikácia funkcie na maticu
array()	vytvorenie objektu pole
as.data.frame()	zmena nejakého objektu na dátovú tabuľku
as.factor()	zmena nejakého objektu na kategoriálnu premennú
as.matrix()	zmena nejakého objektu na maticu
as.numeric()	zmena nejakého objektu na číselný vektor
as.vector()	zmena nejakého objektu na vektor
attach()	použitie názvov stĺpcov
c()	vytvorenie objektu vektor
cbind()	skladanie vektorov po stĺpcoch
combn()	jednoduchý náhodný výber s vrátením z knižnice utils
data.frame()	vytvorenie dátovej tabuľky
dimnames()	pomenovania dimenzií v objekte
factor()	vytvorenie kategoriálnej premennej
ftable()	vytvorenie kontingenčnej tabuľky
choose()	počet všetkých výberov s vrátením z knižnice utils
levels()	hladiny kategoriálnej premennej
library()	načítanie nainštalovanej knižnice
list()	vytvorenie objektu list
margin.table()	marginálne početnosti v tabuľke početností
matrix()	vytvorenie objektu matica
na.omit()	odstránenie chýbajúcich pozorovani ozn. NA
names()	pomenovanie dimenzií v objekte alebo elementov objektu
numeric()	vytvorenie číselného vektora
print()	výpis objektu (použitie vo funkcii alebo cykle)
prop.table()	vytvorenie tabuľky pravdepodobností
rbind()	skladanie vektorov po riadkoch
read.table()	načítanie (import) dátovej tabuľky do R
rep()	vytvorenie vektora opakovaním jeho elementov
replicate()	opakovanie nejakého príkazu alebo funkcie
return()	výpis objektu (použitie vo funkcii)
sample()	jednoduchý náhodný výber bez vrátenia alebo s vrátením
seq()	vytvorenie sekvencie čísel
split()	rozdelenie dát na skupiny
t()	transpozícia matice
table()	vytvorenie tabuľky početností
tapply()	aplikácia funkcie na pole
write.table()	export tabuľky z R
X[i]	indexovanie vo vektore
X[[i]]	indexovanie v liste
X[i,j]	indexovanie v matici
"retazec"	textový reťazec
==	indexovanie
!is.na()	identifikácia chýbajúcich pozorovaní ako FALSE, ostatných ako TRUE

Tabuľka 10.3: Prehľad štatistických a numericko-matematických funkcií

funkcia	**popis**
cr()	koeficient reliability merania
grafy.dva.vybery()	tri grafy pre exploratórnu analýzu dát
grafy.jeden.vyber()	štyri grafy pre exploratórnu analýzu dát
chisq.stat	χ^2 test dobrej zhody
ISkor()	IS pre klasický Pearsonov korelačný koeficient ρ
ISkor.rozdiel()	IS pre rozdiel dvoch Pearsonových korelačných koeficientov $\rho_1 - \rho_2$
ISkor.uhl()	IS pre lineárno-uhlový Pearsonov korelačný koeficient ρ
ks.l.mc()	MC Lillieforsov test dobrej zhody
ks.mc()	MC Kolmogorov-Smirnovov test dobrej zhody
kvart.sikmost()	kvartilová šikmosť
lambda.multinom()	parameter necentrality λ^2 rozdelenia χ^2_{J-1,λ^2}
lin.uhl.r()	lineárno-uhlový Pearsonov korelačný koeficient r
LRp()	logaritmus funkcie vierohodnosti pre pravdepodobnosť p
min.rozsah.n.kor()	minimálny rozsah n pre rozdiel dvoch korelačných koeficientov $\rho_1 - \rho_2$
min.rozsah.n.p()	minimálny rozsah N pre rozdiel dvoch pravdepodobností $p_1 - p_2$
oktil.sikmost()	oktilová šikmosť b_{1O}
priem.uhol()	priemerný uhol
rho.hat()	odhad korelačného koeficientu $\widehat{\rho}$ (iteračné riešenie rovníc)
rozptyl()	rozptyl s_x^2
SE()	štandardná chyba $s_{\overline{x}}$
sikmost()	šikmosť b_1
sin.cos.uhla()	sínus a kosínus uhla v stupňoch
smerodch()	smerodajná odchýlka s_x
spicatost()	špicatosť b_2
tem()	technická chyba merania
tem.rel()	relatívna technická chyba merania
test.pomeru.sanci()	Waldove testy o pomere šancí
test.relat.rizika()	Waldove testy o relatívnom riziku
test.rozdielu.prav()	Waldove testy o rozdiele pravdepodobností
testy.o.kor.koef()	test pomerom vierohodnosti a χ^2-test o korelačných koeficientoch
testy.o.rozptyloch()	test pomerom vierohodnosti a Bartlettov test o homogenite rozptylov
testy.o.stred.hodnotach()	F-test a test pomerom vierohodnosti o rovnosti stredných hodnôt
uhl.uhl.r()	uhlový Pearsonov korelačný koeficient r
urezanie()	γ-urezaný vektor
winsorizacia()	vektor winsorizovaný pomocou „vnútorných hradieb“
zakl.char()	vybrané základné charakteristiky
aov()	ANOVA model s fixnými efektami
bartlett.test()	viacvýberový test o homogenite rozptylov
binconf()	Waldov 95% empirický IS pre pravdepodobnosť z knižnice Hmisc
cor.test()	jednovýberový t-test nulovosti korelačného koeficientu ρ
ks.test()	Kolmogorov-Smirnovov test dobrej zhody
lillie.test	Lillieforsov test dobrej zhody z knižnice nortest
pairwise.t.test()	viacvýberový test o rovnosti stredných hodnôt
power.t.test()	sila jednovýberového Studentovho t-testu o strednej hodnote
prop.test()	χ^2 test o pravdepodobnosti alebo o vektore pravdepodobností
t.test()	Studentove t-testy
TukeyHSD()	viacvýberový Tukeyho HSD test
var.test()	χ^2 test o podiele rozptylov
optim()	numerické hľadanie maxima alebo minima funkcie
optimize()	numerické hľadanie maxima alebo minima funkcie
uniroot()	numerické hľadanie koreňov rovnice

Tabuľka 10.4: Prehľad funkcií kresliacich rôzne druhy grafov a funkcií s grafmi súvisiacich

funkcia	**popis**
abline()	priamka so zadaným interceptom a sklonom
arrows()	úsečka so šípkami na ich koncoch
axis()	dokreslenie osí x a y do obrázka
barplot()	stĺpcový diagram
box()	ohraničenie grafu
boxplot()	krabicový diagram
contour()	kontúrový graf/vkreslenie kontúr do existujúceho obrázka
curve()	krivka
errbar()	chybová úsečka z knižnice Hmisc
gray()	vytvorenie RGB-vektora z vektora hladín sivej z intervalu $\langle 0, 1\rangle$
grey()	vytvorenie RGB-vektora z vektora hladín sivej z intervalu $\langle 0, 1\rangle$
heat.colors()	teplé farby
histback()	dva histogramy dotýkajúce sa bázami z knižnice Hmisc
histogram()	histogram
identify()	výpis pozície (x,y) v grafe na mieste kurzora
image()	sieť farebných štvoruholníkov
kde2d()	dvojrozmerný graf hustoty
legend()	výpis legendy grafu
lines()	nakreslenie krivky medzi bodmi (x,y)
locator()	výpis pozície (x,y) v grafe na mieste kurzora
matplot()	komplikovanejší bodový graf
mtext()	výpis textu na špecifikované miesto mimo grafu
par()	parametre obrázka
persp()	perspektívny 3D graf
pie()	kruhový diagram
plot()	bodový (rozptylový) graf
plot(ecdf())	empirická distribučná funkcia
points()	body v súradniciach (x,y)
polygon()	vyfarbený mnohouholník medzi dvoma krivkami
qqline()	kvantilová priamka
qqnorm()	kvantilový diagram
qqPlot()	kvantilový diagram s Atkinsonovou obálkou z knižnice car
rainbow()	farby dúhy
rect()	vyfarbený štvoruholník (a jeho vnútro) definovaný svojimi vrcholmi
rug()	„koberec“ pod histogramom
scatterplot3d()	3D bodový graf z kniznice scatterplot3d
segments()	úsečka medzi bodmi (x0,y0) a (x1,y1)
sm.density()	hustota z knižnice sm
spine()	spinogram z knižnice vcd
terrain.colors()	topografické farby
text()	text v polohe (x,y)
title()	popis grafu
topo.colors()	topografické farby
windows()	otvorenie okna danej šírky a výšky v palcoch

Tabuľka 10.5: Prehľad funkcií súvisiacich s grafmi – popis obrázka a jeho parametre

argument	**popis**	**argument**	**popis**
xlab	popis osi x	col	farba
ylab	popis osi y	pch	typ bodu
sub	popis grafu	cex	veľkosť bodu
main	nadpis grafu	lty	typ čiary
axes	zobrazenie osí	lwd	šírka čiary
fig	súradnice plochy obrázka	mfcol	počet stĺpcov v sérii
new	pridanie obrázka do série	mfrow	počet riadkov v sérii
font	typ písma	family	písmová rodina
xlim	hranice grafu v smere osi x	ylim	hranice grafu v smere osi y
asp	pomer škály osí y a x	mar	okraje grafu

Tabuľka 10.6: Prehľad funkcií súvisiacich s rozdeleniami pravdepodobnosti

funkcia	**popis**	**funkcia**	**popis**
binomické rozdelenie		**Poissonovo rozdelenie**	
dbinom()	pravdepodobnostná funkcia	dpois()	pravdepodobnostná funkcia
pbinom()	distribučná funkcia	ppois()	distribučná funkcia
qbinom()	kvantil	qpois()	kvantil
rbinom()	pseudonáhodné čísla	rpois()	pseudonáhodné čísla
multinomické rozdelenie		**gama rozdelenie**	
dmultinom()	pravdepodobnostná funkcia	dgamma()	hustota
pmultinom()	distribučná funkcia	pgamma()	distribučná funkcia
qmultinom()	kvantil	qgamma()	kvantil
rmultinom()	pseudonáhodné čísla	rgamma()	pseudonáhodné čísla
normálne rozdelenie		**Studentovo rozdelenie**	
dnorm()	hustota	dt()	hustota
pnorm()	distribučná funkcia	pt()	distribučná funkcia
qnorm()	kvantil	qt()	kvantil
rnorm()	pseudonáhodné čísla	rt()	distribučná funkcia
χ^2 **rozdelenie**		**Fisherovo rozdelenie**	
dchisq()	hustota	df()	hustota
pchisq()	distribučná funkcia	pf()	distribučná funkcia
qchisq()	kvantil	qf()	kvantil
rchisq()	pseudonáhodné čísla	rf()	pseudonáhodné čísla
mnohorozmerné normálne rozdelenie z knižnice mvtnorm		**mnohorozmerné normálne rozdelenie** z knižnice MASS	
rmvnorm()	pseudonáhodné čísla	mvrnorm()	pseudonáhodné čísla

Literatúra

Agresti, A., 2002: *Categorical Data Analysis.* Hoboken: John Wiley & Sons

Agresti, A., Coull, B.A., 1998: Approximate Is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician* 52(2): 119–126

Alánová, N., 2008: *Sociální a morfologické vlastnosti matek a novorozenců a jejich souvislosti s poměrem pohlaví.* Bakalárska práca. Brno: Masarykova univerzita

Almagor, M., Schwed, P., Yaffe, H., 1998: Male to female ratio in newborns of grand grand multiparous women. *Human Reproduction* 13(8): 2323–2324

Alt, K.W., Benz, M., Müller, W., Berner, M.E., Schultz, M., Schmidt-Schultz, T.H., Knipper, C., Gebel, H.-G.K., Nissen, H.J., Vach, W., 2013: Earliest Evidence for Social Endogamy in the 9,000-Year-Old-Population of Basta, Jordan. *PLoS ONE* 8(6): e65649

Anděl, J., 2011: *Základy matematické statistiky.* Praha: MatfyzPress

Angel, J.L., 1982: A New Measure of Growth Efficiency: Skull Base Height. *American Journal of Physical Anthropology* 58: 297–305

Ashley Montagu, M.F., 1944: Aleš Hrdlička, 1869–1943. *American Anthropologist* 46: 113–117

Atkinson, A.C., 1981: Two Graphical Displays for Outlying and Influential Observations in Regression. *Biometrika* 68: 13–20

Azzalini, A., 1996: *Statistical inference based on likelihood.* Boca Raton: Chapman Hall/CRC Press

Bainbridge, D., Martin, J., Arango, M., Cheng, D., 2012: Perioperative and anaesthetic-related mortality in developed and developing countries: a systematic review and meta-analysis. *The Lancet* 380(9847): 1075–1081

Bartlett, M. S. 1937: Properties of sufficiency and statistical tests. *Proceedings of the Royal Statistical Society, Series A* 160: 268−282

Becker, R.A., Chambers, J.M., Wilks, A.R., 1988: *The New S Language.* Boca Raton: Chapman and Hall/CRC

Benjamini, Y., Hochberg, Y., 1995: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57: 289−300

Benjamini, Y., Yekutieli, D., 2001: The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 29: 1165−1188

Bickel, P.J., Doksum, K.A., 2006: *Mathematical Statistics: Basic Ideas and Selected Topics (Vol. 1).* Upper Saddle River: Pearson

Bland, M., 2009: *An introduction to medical statistics.* Oxford: Oxford University Press

Bogaert, A.F., Blanchard, R., 1996: Physical development and sexual orientation in men: Height, weight and age of puberty differences. *Personality and individual differences* 21: 77–84

Bogaert, A.F., Skorska, M., 2011: Sexual orientation, fraternal birth order, and the maternal immune hypothesis: a review. *Frontiers in Neuroendocrinology* 32(2): 247–254

Bogin, B., 1999: *Patterns of Human Growth.* New York: Cambridge University Press

Bonferroni, C.E., 1936: *Teoria statistica delle classi e calcolo delle probabilita.* Firenze: Libreria Internazionale Seeber

Bouchalová, M., 1987: *Vývoj během dětství a jeho ovlivnění.* Praha: Avicenum

Bowman, A.W., Azzalini, A., 1997: *Applied Smoothing Techniques for Data Analysis.* Oxford: Oxford University Press

Brazzale, A.R., Davison, A.C., Reid, N., 2007: *Applied Asymptotics: Case Studies in Small-Sample Statistics.* New York: Cambridge University Press

Brent, R.P., 1973: Chapter 4: An Algorithm with Guaranteed Convergence for Finding a Zero of a Function. *Algorithms for Minimization without Derivatives.* Englewood Cliffs, NJ: Prentice-Hall

Broca, P., 1862: Sur les proportions relatives du bras, de l'avant bras et de la clavicule chez les nègres et les Européens. *Bulletin Société d'Anthropologie Paris* 3(2): 162–172

Brůžek, J., 2008: Současná česká paleodemografie: falešné naděje přílišného optimismu a nový reálný cíl. *Archeologické rozhledy* 60: 329–344

Budíková, M., Králová, M., Maroš, B., 2010: *Průvodce základními statistickými metodami.* Praha: Grada

Buk, Z., Kordík, P., Brůžek, J., Schmitt, A., Šnorek, M., 2012: The age at death assessment in a multi-ethnic sample of pelvic bones using nature-inspired data mining methods. *Forensic Science International* 220(1–3): 294–294

Býmová, I., 1990: *Dlaňové dermatoglyfy ve vztahu k barvě vlasů.* Magisterská diplomová práca. Brno: Masarykova univerzita

Cardoso, H.F.V., 2008: Secular changes in body height and weight of portuguese boys over one century. *American Journal of Human Biology* 20: 270–277

Casella, G., Berger, R.L., 2002: *Statistical Inference.* Pacific Grove: Duxbury Press

Cox, D.R., 2006: *Principles of statistical inference.* Cambridge: Cambridge University Press

Cox, D.R., Donnelly, C.A., 2011: *Principles of Applied Statistics.* New York: Cambridge University Press

Cox, D.R., Hinkley, D.V., 1974: *Theoretical Statistics.* London: Chapman and Hall

Cressie, N., Read, T.R.C., 1984: Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B* 46: 440–464

Dadejová, V., Králík, M., Urbanová, P., 2011: Věkové a mezipohlavní rozdíly v rozměrech zubního oblouku dolní čelisti nedospělých jedinců: Brněnská sbírka sádrových ortodontických modelů chrupu. *Anthropologia Integra* 2(1): 13–29

Dalgaard, P., 2008: *Introductory Statistics with R.* New York: Springer

Dallal, G.E., Wilkinson, L., 1986: An analytic approximation to the distribution of Lilliefors' test for normality. *The American Statistician* 40: 294–296

DasGupta, A., 2008: *Asymptotic Theory of Statistics and Probability.* New York: Springer

Dodo, Y., 1974: Non-metrical cranial traits in the Hokkaido Ainu and the Northern Japanese of recent times. *Journal of Anthropological Society of Nippon* 82: 31–51

Donner, A., Rosner, B., 1980: On Inference Concerning a Common Correlation Coefficient. *Journal of the Royal Statistical Society, Series C* 29(1): 69–76

Dudoit, S., van der Laan, M.J., 2008: *Multiple Testing Procedures with Applications to Genomics.* New York: Springer

Eveleth, P.B., Tanner, J.M., 1990: *Worldwide variation of human growth.* Cambridge: Cambridge University Press

Fetter, V. a kol., 1967: *Antropologie.* Praha: Academia

Fisher, R.A., 1935: The Design of Experiments. London: Macmillan

Fisher, R.A., 1936, 1971: The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7: 179–188
Flack, V.F., Flores, R.A., 1989: Using Simulated Envelopes in the Evaluation of Normal Probability Plots of Regression Residuals. *Technometrics* 31(2): 219–225
Garn, S.M., Rohmann, C.G.,1963: On the Prevalence of Skewness in Incremental Data. *American Journal of Physical Anthropology* 21: 235–236
Geissler, A., 1889: Beiträge ge zur Frage des Geschlechtsverhältnisses der Geborenen. *Z. Köngl. Sächs. Statist. Bur.* 35: 1–24
Gentle, J.E., 2007: *Matrix Algebra: Theory, Computations, and Applications in Statistics*. New York: Springer
Gentle, J.E., 2009: *Computational Statistics*. New York: Springer
Gerring, J., 2007: *Case Study Research. Principles and Practices*. Cambridge: Cambridge University Press
Givens, G.H., Hoeting, J.A, 2005: *Computational Statistics*. Hoboken: Wiley & Sons
Goto, R., Mascie-Taylor, C.G.N., 2007: Precision of Measurement as a Component of Human Variation. *Journal of Physiological Anthropology* 26: 253–256
Greenwood, M., Yule, G.U., 1920: An inquiry into the nature of frequency distributions representative of multiple happening with particular reference to the occurrence of multiple attacks of disease or repeated accidents. *Journal of the Royal Statistical Society, Series A* 83: 255–279
Hackshaw, A., 2009: *A concise guide to clinical trials*. Chichester: Wiley-Blackwell, BMJ Books
Hackshaw, A., 2011: *How to write a grant application*. Chichester: Wiley-Blackwell, BMJ Books
Halligan, S., 2002: Reproducibility, repeatability, correlation and measurement error. *The British Journal of Radiology* 75: 193–195
Hauser G., De Stefano G. F. a kol., 1989: *Epigenetic Variants of the Human Skull*. Stuttgart: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller)
Hochberg, Y., 1988: A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75: 800–803
Hollier, L.P., Keelan, J.A., Jamnadass, E.S.L., Maybery, M.T., Hickey, M., Whitehouse, A.J.O., 2015: Adult digit ratio (2D:4D) is not related to umbilical cord androgen or estrogen concentrations, their ratios or net bioactivity. *Early Human Development* 91(2): 111–117
Holm, S., 1979: A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6: 65–70
Horníčková, L., 1992: *Barva očí brněnských vysokoškoláků*. Magisterská diplomová práca. Brno: Masarykova univerzita
Horová, I., Zelinka, J., 2008: *Numerické metódy*. Brno: Masarykova univerzita
Hotelling, H., 1979: New light on the correlation coefficient and its transforms. *Journal of the Royal Statistical Society, Series B* 15: 193–232
Hrdlička, A., 1920: *Anthropometry*. Philadelphia: Wistar Institute of Anatomy and Biology
Hrdlička, A., 1939, 1952: *Pactical Anthropometry*. Philadelphia: Wistar Institute of Anatomy and Biology
Hulanicka, B., 1973: *Anthroposcopic features as a measure of similarity. Materialy i prace antropologiczne (Wrocław)* 86: 115–155
Hulley, S.B., Cummings, S.R., Browner, W.S., Grady, D.G., Newman, T.B., 2007: *Designing clinical research*. Philadelphia: Wolters Kluwer
Chambers, J.M., 2008: *Software for Data Analysis: Programming with R*. New York: Springer
Cheng, D.C., Bainbridge, D., Martin, J.E., Novick, R.J., 2005: Does Off-pump Coronary Artery Bypass Reduce Mortality, Morbidity, and Resource Utilization When Compared with Conventional Coronary Artery Bypass? A Meta-analysis of Randomized Trials. *Anesthesiology* 102(1): 188–203
Cheng, H., Shang, X., He, Y., Zhang, T., Zhang, Y.P., Zhou, R., 2007: Insight into human sex ratio imbalance: the more boys born, the more infertile men. *Reproductive Biomedicine Online* 15(5): 487–494
Chernoff, H. , Lehmann, E.L., 1954: The use of maximum likelihood estimates in χ^2 tests for goodness of fit. *Annals of Mathematical Statistics* 25: 579–586
Christensen, R., 1997: *Log-Linear Models and Logistic Regression*. New York: Springer
Christensen, R., 2002: *Plane Answers to Complex Questions*. New York: Springer
James, W.H., 2006: Possible constraints on adaptive variation in sex ratio at birth in humans and other primates. *Journal of Theoretical Biology* 238: 383–394
James, W.H., 2008: The variations of human sex ratio at birth with time of conception within the cycle, coital rate around the time of conception, duration of time taken to achieve conception, and duration of gestation: A synthesis. *Journal of Theoretical Biology* 255: 199–204
Jammalamadaka, S.R., Lund, U.J., 2006: The effect of wind direction on ozone levels: a case study. *Environ Ecol Stat* 13: 287–298
Jit, I., Singh, S., 1966: The sexing of the adult clavicles. *Indian Journal of Medical Research* 54: 551–571
Juntunen, K.S.T., Kvist, A.P., Kauppila, A.J.I., 1997: A shift from a male to a female majority in newborns with the increasing age of grand grand multiparous women. *Human Reproduction* 12: 2321–2323
Jurda, M., 2008: *Tafonomické změny lidské lebky z pohledu geometrické morfometrie*. Magisterská diplomová práca. Brno: Masarykova univerzita
Kabacoff, R.I., 2011: *R in Action: Data analysis and graphics with R*. Shelter Island: Manning Publications
Katina, S., Bodoríková, S., Dornhoferová, 2011: Reliability of measurements in geometric and traditional morphometrics of human skull. *Česká antropologie* 61(2): 16–25
Katz, D., Baptista, J., Azen, S.P., Pike, M.C., 1978: Obtaining confidence intervals for the risk ratio in in cohort studies. *Biometrics* 34: 469–474
Kaur, J., Choudhry, R., Raheja, S., Dhissa, N.C., 2012: Non metric traits of the skull and their role in anthropological studies. *Journal of Morphological Science* 29(4): 189–194
Kinsey, A.C., Pomeroy, W.B., Martin, C.E., 1948: *Sexual Behavior in the Human Male*. Philadelphia and London: W. B. Saunders Company
Knussmann, R. (ed.), 1988: *Anthropologie, Handbuch der vergleichenden Biologie des Menschen* (4. Auflage des Lehrbuchs der Anthropologie begrndet von Rudolf Martin), Band I und II. Stuttgart: Gustav Fischer Verlag
Kolmogorov, A.N., 1933: Sulla determinazione empirica di una legge di distribuzione. *Giornale dell' Istituto Italiano degli Attuari* 4: 83–91
Komenda, S., 2000: *Vypočitatelná náhoda. Elementy počtu pravděpodobnosti a matematické statistiky*. Olomouc: Univerzita Palackého v Olomouci
Králík, M., Urbanová, P., Wagenknechtová, M., 2014: Sex assessment using clavicle measurements: Inter- and intra-population comparisons. *Forensic Science International* 234: 181.e1–181.e15
Kukla, L., 2008: Evropská dlouhodobá studie těhotenství a dětství – ELSPAC. *Vox Pediatriae* 8(5): 24–26
Kumar, V. S., Jeyaseelan, L., Sebastian, T., Regi, A., Mathew, J., Jose, R., 2013: New birth weight reference standards customised to birth order and sex of babies from South India. *BMC Pregnancy and Childbirth* 13: 38–38
Larsen, C.S., 1997: *Bioarchaeology: Interpreting behavior from the human skeleton*. Cambridge: Cambridge University Press

Lehmann, E.L., 1999: *Elements of Large-Sample Theory*. New York: Springer
Lehmann, E.L., Casella, G., 1998: *Theory of Point Estimation*. New York: Springer
Leroux, B.G., Puterman, M.L., 1992: Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* 48: 545–558
LeVay, S., 1991: A difference in hypothalamic structure between homosexual and heterosexual men. *Science* 253: 1034–1037
Lieberman, D.E., 2011: *Complexity, Modularity, and Integration in the Human Head. The Evolution of the Human Head*. Cambridge, Mass: The Belknap Press of Harvard University Press
Lindsey, J.K., Altham, P.M.E., 1998: Analysis of the human sex ratio by using overdispersion models. *Appl. Statist.* 47(1): 149–157
Lorencová, A., Beneš, J. 1976: Body composition of workers in three industrial centres in Moravia (ČSSR). In: M. Dokládal (ed.) *Human growth and physical development*, Brno: J. E. Purkyně University Brno, 141–148
Maccoby, E.E., Doering, C.H., Jacklin, C.N., Kraemer, H., 1979: Concentrations of sex hormones in umbilical–cord blood: their relation to sex and birth order of infants. *Child Development* 50(3): 632–642
Macfarlane, J., Prewett, J., Rose, D., Gard, P., Cunningham, R., Saikku, P., Euden, S., Myint, S., 1997: Prospective casecontrol study of role of infection in patients who reconsult after initial antibiotic treatment for lower respiratory tract infection in primary care. *British Medical Journal* 315(7117): 1211–1214
MacIntyre, C.R., Heywood, A.E., Kovoor, P., Ridda, I. Seale, H., Tan, T., Gao, Z., Katelaris, A.L., Siu, H.W.D., Lo, V., Lindley, R., Dwyer, D.E., 2013: Ischaemic heart disease, influenza and influenza vaccination: a prospective case control study. *Heart* 99(24): 1843–1848
Mardia, K.V., 1976: Linear-circular correlation and rhythmometry. *Biometrika* 63: 403–405
Mardia, K.V., Jupp, P.E., 2000: *Directional statistics*. New York: Wiley
Marini, E., Racugno, W., Borgognini Tarli, S. M., 1999: Univariate estimates of sexual dimorphism: the effects of intrasexual variability. *Americal Journal of Physical Anthropology* 109: 501–508
Martin, R., 1914, 1928: *Lehrbuch der Anthropologie* 1. a 2. Aufl. Band II. Kraniologie, Osteologie. Stuttgart: Gustav Fischer Verlag
Martin, R., Saller, K., 1957–1966: *Lehrbuch der Anthropologie* 3. Aufl. Stuttgart: Gustav Fischer Verlag
Mastrangelo, P., De Luca, S., Alemn, I., Botella, M.C., 2011: Sex assessment from the carpals bones: Discriminant function analysis in a 20th century Spanish sample. *Forensic Science International* 206: 216.e1–216.e10
Matalová, A. (ed.), 2008: Gregor Mendel, *Pokusy s hybridy rostlin*. Brno: Jiří Krejčí, Nakl. K-public
Matloff, N., 2011: *The Art of R Programming: A Tour of Statistical Software Design*. San Francisco: William Pollock
Mays, S., Steele, J., Ford, M., 1999: Directional asymmetry in the human clavicle. *International Journal of Osteoarchaeology* 9: 18–28
McPaul, M.J., Toto, R.D., 2011: *Clinical research from proposal to implementation*. Philadelphia: Wolters Kluwer
Meloun, M., Militký, J., 2004: *Statistické spracování experimentálních dat*. Praha: East Publishing
Mielke, J.H., Konigsberg, L.W., Relethford, J.H., 2011: *Human Biological Variation* (2 ed.). New York – Oxford: Oxford University Press
Mikešová, T., 2008: *Poporodní změny na kostře: kritická historická studie*. Magisterská diplomová práca. Brno: Masarykova univerzita
Mood, A.M., Graybill, F.A., Boes, D.C., 1987: *Introduction to the theory of statistics*. London: McGraw-Hill
Mouri, T., 1976: A Study of non-metrical cranial variants of the modern Japanese in the Kinki district. *Journal of Anthropological Society of Nippon* 84: 191–203
Mueller, W.H., Martorell, R., 1988: Reliability and accuracy of measurement. In: Lohman, T.G., Roche, A.F., Martorell, R. (eds.) *Anthropometric Standardization Reference Manual*. Champaign, Illinois: Human Kinetic Books, 83–86
Murrell, P., 2011: *R Graphics*. Boca Raton: Chapman Hall/CRC Press
Nedoma, K., Freund, K., 1961: Somatosexuální nálezy u homosexuálních mužů. *Československá psychiatrie* 57: 100–103
Newcombe, R.G., 1998: Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine* 17(8): 857–872
Okajima, M., Iwayanagi, C., 1986: Absence of palmar digital triradius d observed in 2,681 Japanese families and clinical significance. *Japanese Journal of Human Genetics* 31(4): 331–336
Parsons, F.G., 1916: On the proportions and characteristics of the modern English clavicle. *Journal of Anatomy* 51: 71–93
Pařízková, J., 1973: *Body composition and lipid metabolism in different regimes of physical activity*. Praha: Avicenum
Paul, S.R., 1989: Test for equality of several correlation coefficients. *The Canadian Journal of Statistics* 17(2): 217–227
Pawitan, Y., 2001: *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford: Oxford University Press
Peacock, J.L., Peacock, P.J., 2011: *Oxford handbook of medical statistics*. Oxford: Oxford University Press
Pearson, K., 1899: IV. *Mathematical contribution to the theory of evolution. – V. On the reconstruction of the stature of prehistoric races*. Philosophical Transactions of the Royal Society of London, Series A 192: 169–244
Pearson, K., 1933: On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika* 25: 379–410
Preedy, V.R. (ed), 2012: *Handbook of Anthropometry: Physical Measures of Human Form in Health and Disease*. New York: Springer
R Development Core Team, 2013: *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing
Rasch, D., Pilz, J., Verdooren, R., Gebhardt, A., 2011: *Optimal Experimental Design with R*. Boca Raton: Chapman Hall/CRC Press
Rizzo, M.L., 2007: *Statistical Computing with R*. Boca Raton: Chapman Hall/CRC Press
Robert, C.P.R., Casella, G., 2010: *Introducing Monte Carlo Methods with R*. New York: Springer
Rousseeuw, P.J., van Zomeren, B.C., 1990: Unmasking multivariate outliers and leverage points. *J. Amer. Statist. Assoc.* 85: 633–651
Rousso, D., Panidis, D., Gkoutzioulis, F., Kourtis, A., Mavromatidis, G., Kalahanis, I., 2002: Effect of the interval between pregnancies on the health of mother and child. *European Journal of Obstetrics and Gynecology and Reproductive Biology* 105(1):4–6
Russell Bernard, H., 2006: *Research Methods in Anthropology. Qualitative and Quantitative Approaches*. Fourth Edition. New York: Rowman and Littlefield Publishers
Rücker, G., Schwarzer, G., Carpenter, J., Olkin, I., 2009: Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Statistics in Medicine* 28(5): 721–738
Seidman, D.S., Ever-Hadani, P., Stevenson, D.K., Slater, P.E., Harlap, S., Gale, R., 1988: Birth Order and Birth Weight Reexamined. *Obstetrics and Gynecology* 72(2): 158–162
Shaffer, J.P., 1995: Multiple hypothesis testing. *Annual Review of Psychology* 46: 561–576
Schaie, K.W., 2000: The impact of longitudinal studies on understanding development from young adulthood to old age. *Inter-*

national Journal of Behavioral Development 24: 257–266
Scheffe, 1953: *The Analysis of Variance*. Hoboken: John Wiley Sons
Schmidt, E., 1888: Catalog der im anatomischen Institut der Universitt Leipzig aufgestellten craniologischen Sammlung des Herrn Dr. Emil Schmidt. *Archiv für Anthropologie* 17. Braunschweig: Vieweg und sohn
Singh, D., Jit, I., 1996: Identification of sex from the volume of the clavicle. *Journal of the Anatomical Society of India* 45(2): 119–124
Singh, S., Gangrade, K.C., 1968: The sexing of adult clavicles – demarking points for Varanasi zone. *Journal of the Anatomical Society of India* 17: 89–100
Smirnov, N.V., 1933: Estimate of deviation between empirical distribution functions in two independent samples. *Bulletin Moscow University* 2: 3–16
Sovinová, H., Sadílek, P., Csémy, L., 2012: Vývoj prevalence kuřáctví v dospělé populaci ČR. Názory a postoje občanú ČR k problematice kouření (období 1997–2011). Výzkumná zpráva. Státní zdravotní ústav.
Spector, P., 2008: *Data Manipulation with R*. New York: Springer
Stephens, M.A., 1974: EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association* 69: 730–737
Stewart, T. D., 1970: Identification of the scars of parturition in the skeletal remains off females. In: Stewart, T.D. (ed.) *Personal identification in mass disasters. Report of a seminar held in Washington, D.C., 9–11 December 1968, by arrangement between the Support Services of the Department of the Army and the Smithsonian Institution*. Washington, D.C.: National Museum of Natural History, Smithsonian Institution, 127–135
Suess, E.A., Trumbo, B.E., 2010: *Introduction to Probability Simulation and Gibbs Sampling with R*. New York: Springer
Suchý, J., 1967: Nejčastěji používané metody zpracování antropologického materiálu. In: Fetter, V. a kol., 1967: *Antropologie*, 173–196. Praha: Academia
Swaab, D.F., Gooren, L.J., Hofman, M.A., 1992: Gender and sexual orientation in relation to hypothalamic structures. *Hormone Research* 38(2): 51–61
Swamy, G.K., Edwards, S., Gelfand, A., James, S.A., Miranda, M.L., 2012: Maternal age, birth order, and race: differential effects on birthweight. *Journal of Epidemiology and Community Health* 66(2): 136–142
Szilvássy, J., 1988: Altersdiagnose am Skelett. In: Knussmann, R. (ed.) *Anthropologie. Handbuch der vergleichenden Biologie des Menschen (4. Auflage des Lehrbuchs der Anthropologie begrndet von Rudolf Martin)*, Band I und II. Jena, New York, Stuttgart: Gustav Fischer
Šefčáková, A., Katina, S., 2008: Geometrical analysis of adult skulls from Předmostí. In: Velemínská, J., Brůzek, J. (eds.) *Early modern humans from Předmostí nr. Přerov: Old documentation and new reading.* Praha: Academia, 87–101
Šefčáková, A., Katina, S., Mizera, I., Halouzka, R., Barta, P., Thurzo, M., 2011: A late upper palaeolithic skull from Moca (The Slovak Republic) in the context of central Europe. *Acta Musei Nationalis Pragae: Series B: Historia naturalis* 67(1–2): 3–22
Šidák, Z., 1967: Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association* 62(318): 626–633
Štouračová, B., 1992: *Variabilita ušního boltce u brněnských vysokoškoláků*. Magisterská diplomová práca. Brno: Masarykova univerzita
Šťastná, A., 2010: Změny reprodukčních vzorců a individuální souvislosti rodičovství. *Demografie* 52(4): 77–88
Titlbachová, S., 1967: Somatoskopie. In: Fetter, V. a kol., *Antropologie*. Praha: Academia
Tukey, J.W., 1949: Comparing Individual Means in the Analysis of Variance. *Biometrics* 5(2): 99–114
Tukey, J.W., 1953: The problem of multiple comparisons. Unpublished manuscript. In *The Collected Works of John W. Tukey VIII. Multiple Comparisons: 1948–1983*. New York: Chapman & Hall
Tukey, J.W., 1962: The Future of Data Analysis. *The Annals of Mathematical Statistics* 31: 385–391
Tukey, J.W., 1991: The philosophy of multiple comparisons. *Statist. Sci.* 6: 100–116
Ugarte, M.D., Militino, A.F., Arnholt, A.T., 2008: *Probability and Statistics with R*. Boca Raton: Chapman Hall/CRC Press
Ulijaszek, S.J., Kerr, D.A., 1999: Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition* 82: 165–177
Ulijaszek, S.J., Lourie, J.A., 1994: Intra- and inter-observer error in anthropometric measurement. In: Ulijaszek, S.J., Mascie-Taylor, C.G.N. (eds.) *Anthropometry: the individual and the population*. Cambridge: Cambridge University Press
Venables, W.N., Ripley, B.D., 2002: Modern Applied Statistics with S. New York: Springer
Venables, W. N., Smith, D.M., the R Core Team, 2013: *An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics* Version 3.0.1 (2013-05-16). Vienna: R Foundation for Statistical Computing
Verzani, J., 2005: Using R for Introductory Statistics. Boca Raton: Chapman & Hall/CRC Press
von Bortkiewicz, L., 1898: Das Gesetz der kleinen Zahlen [The law of small numbers]. Leipzig: B.G. Teubner
Vondrušková, L., 1983: *AB0/H/ systém*. Magisterská diplomová práca. Brno: Masarykova univerzita
Waaler, H.T., 1984: Height, Weight and Mortality: The Norwegian Experience. *Acta Medica Scandinavica, Supplementum* 679: 1–56
Walker, J., Almond, P., 2010: *Interpreting statistical findings*. Maidenhead: Open University Press
Wasserman, L., 2006: *All of Nonparametris Statistics*. New York: Springer
Weill, C.L., ed., 2009: *Natures Choice. What Science Reveals About the Biological Origins of Sexual Orientation*. New York: Taylor and Francis
Welch, B.L., 1947: The generalization of Student's problem when several different population variances are involved. *Biometrika* 34(1–2): 28–35
White, K.R., 1980: Socio-economic status and academic achievements. *Evaluation in Education* 4: 79–81
Wilks, S.S., 1948: Order statistics. *Bulletin of the American Mathematical Society* 54(1): 6–50
Wilson, E.B., 1927: Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 22: 209–212
Wood, W., Milner, R., Harpending, C., Weiss, M., 1992: The Osteological Paradox: Problems of Inferring Prehistoric Health from Skeletal Samples. *Current Anthropology* 33(4): 343–370
Woolf, B., 1955: On estimating the relation between blood groups and disease. *Annals of Human Genetics* 19(4): 251–253
Yekutieli, H., Benjamini, Y., 1999: Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference* 82: 171–196
Zvára, K., 1999: Statistika v antropologii. In: Stloukal, M. a kol. 1999: *Antropologie. Příručka pro studium kostry*, 433–479. Praha: Národní muzeum
Zvára, K., 2001: *Biostatistika*. Praha: Nakladatelství Karolinum
Živný, M., 2010: *Antropologické zpracování lidských kosterních pozůstatků ze hřbitova u kostela sv. Jakuba v Brně: Výsledky paleodemografické a osteometrické analýzy*. Brno: Akademické nakladatelství CERM

Register

dáta
- databáza, 25
  - bezpečnosť, 26
  - databázový management, 26
  - databázový programátor, 26
  - databázový projekt, 26
  - kontrola kvality, 26
  - štatistický programátor, 26
  - uzamknutie, 27
  - zmrazenie, 27
  - zoznam formulárov, 26
- dokumentácia, 21
- hlavné atribúty, 18
  - objektivita, 18
  - platnosť, 18
  - reliabilita, 18
  - reprezentatívnosť, 18
  - spoľahlivosť, 18
  - validita, 18
- chyba merania, 21
- maskovanie, 25
- metódy zberu, 18
  - dotazník, 19
  - observačné, 18
  - rozhovor, 19
- odľahlé pozorovanie, 85
- otázky, 21
- presnosť merania, 21
- realizácie, 33, 80
  - abundančná matica, 24
  - dátová tabuľka, 24
  - incidenčná matica, 24
  - kontingenčná tabuľka, 50–52, 90
  - matica, 24
  - pole, 24
  - skalár, 24
  - tabuľka početností, 24
  - usporiadané, 80
  - vektor, 24
- stratifikácia, 25
- stratifikačné skupiny, 25
- škála, 19
- štatistické požiadavky, 19
  - jednoznačnosť, 19
  - nezávislosť, 19
  - úplnosť, 19
- triedenie
  - dĺžky triednych intervalov, 24
  - ekvidištantné intervaly, 24
  - hranice triednych intervalov, 24
  - jednostupňové, 24
  - počet triednych intervalov, 24
  - stredy triednych intervalov, 24
  - triedy, 24
  - viacstupňové, 24
- z-skóre, 86
- zaslepenie, 25
  - dvojité, 25
- znáhodnená alokácia, 25
  - alokačný pomer, 25
  - permutačný blokový princíp, 25
- znáhodnenie, 25

dátový súbor
- anova-head.txt, 39, 141, 234, 252, 253, 284
- anova-means-skull.txt, 283
- anova-newborns.txt, 36, 249, 285
- lin-uhl-fm.txt, 171, 277
- more-samples-correlations-skull.txt, 263, 286
- more-samples-probabilities-pubis .txt, 270
- more-samples-probabilities-pubis.txt, 39, 287
- more-samples-variances-clavicle.txt, 86, 104, 259, 285
- multinom-blood-groups.txt, 49, 51, 104, 268, 290
- multinom-earlobe.txt, 49, 51, 104, 290
- multinom-iris-color.txt, 49, 51, 104, 287
- multinom-palmar-lines.txt, 49, 51, 104, 271, 288
- one-sample-correlation-skull-mf.txt, 36, 43, 45, 72, 89, 104, 141, 169, 276
- one-sample-mean-skull-mf.txt, 36, 41, 60, 72, 86, 89, 97–99, 150–152, 273
- one-sample-probability-sexratio.txt, 73, 179, 181, 278
- one-sample-variance-skull-mf.txt, 163, 164, 275
- paired-means-clavicle.txt, 153, 154, 156, 158, 273
- paired-means-clavicle2.txt, 36, 41, 60, 72, 89, 104, 141, 275
- two-samples-correlations-trunk.txt, 39, 100, 104, 207, 281
- two-samples-means-birth.txt, 36, 97, 102, 197, 280
- two-samples-means-skull.txt, 39, 104, 199, 280
- two-samples-probabilities-sexratio.txt, 38, 73, 224, 282
- two-samples-variances-skull.txt, 202, 281
- uhl-uhl-fm.txt, 173, 277

efekt, 39
- korelačný koeficient, 165
- podiel rozptylov, 199
- pomer šancí, 209
- pomer rizík, 209
- pravdepodobnosť, 175
- relatívne riziko, 209
- rozdiel korelačných koeficientov, 204
- rozdiel rizík, 209
- rozdiel stredných hodnôt, 191
- rozptyl, 159
- stredná hodnota, 122, 144, 152

funkcia vierohodnosti, 62
- binomického rozdelenia, 63, 221
- Fisherova miera informácie, 64
  - očakávaná, 64
  - pozorovaná, 64
- jadro, 63
- kvadratická aproximácia, 64, 65, 67, 70
- logaritmus, 63
- maximalizácia, 69
- multinomického rozdelenia, 63
- normálneho rozdelenia, 67, 149, 162, 168, 195, 201, 202, 205, 206, 233, 256–258, 261, 262, 266, 269, 270
- Poissonovho rozdelenia, 64
- pomer vierohodnosti
  - jednoduchý, 120
  - zovšeobecnený, 120
- profilová, 68
  - zakrivenie, 68
- relatívna, 120
- skóre funkcia, 64, 66, 70
- štandardizovaná, 120
- Taylorov rozvoj, 70

graf
- bodový, 91
- empirická distribučná funkcia, 99
- exploratórna analýza dát, 33, 89
- histogram, 89, 96
  - Freedman-Diaconisova formula, 96
  - kumulatívnych početností, 99
  - súčtový, 99
  - Scottova formula, 96
  - Sturgesova formula, 96

    - v absolútnej škále, 96
    - v relatívnej škále, 96
  - koláčový, 94
  - krabicový diagram, 99
    - so zárezom, 99
  - kruhový diagram, 94
  - kvantilový diagram, 101
    - Atkinsonova obálka, 101
  - odtiene
    - farby sivej, 95
    - farieb dúhy, 95
    - teplých farieb, 95
    - topografických farieb, 95
  - *qq*-diagram, 101
  - rozptylový, 91
  - spinogram, 90
  - stĺpcový diagram, 89
  - výsekový diagram, 94
  - veková pyramída, 89

- charakteristika polohy
  - aritmetický priemer, 80
  - decil, 80, 106
  - dolný kvartil, 78, 80, 106
  - druhý kvartil, 80, 106
  - horný kvartil, 78, 80, 106
  - maximum, 80
  - medián, 77, 80, 106
  - minimum, 80
  - modus, 80
  - päťčíselný súhrn, 80
  - percentil, 78, 80, 106
  - poradie, 81
  - prvý kvartil, 80, 106
  - stredná hodnota, 77
  - stredná hodnota mediánu, 79
  - strednoporadie, 81
  - tretí kvartil, 80, 106
  - urezaný aritmetický priemer, 81
  - winsorizovaný aritmetický priemer, 82
- charakteristika variability
  - decilové rozpätie, 84
  - koeficient šikmosti, 83
  - koeficient špicatosti, 83
  - koeficient variácie, 77, 82, 113
  - kvartilový koeficient šikmosti, 84
  - kvartilový koeficient variácie, 84
  - medzikvartilové rozpätie, 83
  - oktilový koeficient šikmosti, 85
  - percentilové rozpätie, 84
  - priemer absolútnych odchýlok, 83
  - rozpätie, 79, 83
  - rozptyl, 77, 82
  - rozptyl aritmetického priemeru, 83
  - rozptyl mediánu, 79
  - smerodajná odchýlka, 82
  - stredná chyba priemeru, 83
  - suma štvorcov, 83
  - štandardná chyba, 83
  - suma absolútnych odchýlok, 83
  - urezaný rozptyl, 84
  - vnútorné hradby, 84
  - vonkajšie hradby, 84
  - winsorizovaný rozptyl, 84
- chyba
  - druhého druhu, 105
  - interindividuálna, 5, 152
  - intraindividuálna, 5, 152
  - kalibrácie meracieho prístroja, 21
  - koeficient reliability, 23
  - náhodná, 22
    - registrácie, 22
  - priemerná kvadratická, 113
  - prvého druhu, 105
  - relatívna technická chyba merania, 22
  - systematická
    - kombinovaná, 21
    - merania, 21
    - prístroja, 21
    - spôsobená externými faktormi, 21
  - technická chyba merania, 22

- interval
  - dĺžka, 115
  - dvojstranný, 113
  - elipsa spoľahlivosti, 121
  - elipsoid spoľahlivosti, 121
  - empirický, 114
  - jednostranný, 113, 117
  - koeficient spoľahlivosti, 105, 113
  - konzervatívny, 118
  - liberálny, 118
  - obojstranný, 117
  - pravdepodobnosť pokrytia
    - aktuálna, 118
    - nominálna, 118
    - skutočná, 118
  - spoľahlivosti, 113
    - skóre, 177
    - vierohodnostný, 121, 127, 128, 150, 162, 169, 178, 187, 202, 207, 222
    - Waldov, 125, 144, 160, 166, 175, 186, 190, 200, 204, 212, 213, 216, 217, 219, 220
  - šírka, 115

- knižnica
  - Hmisc, 97
  - MASS, 45
  - PASWR, 132
  - car, 101
  - kolmim, 79
  - mvtnorm, 45
  - nortest, 140
  - scatterplot3d, 91
  - sm, 98
  - utils, 35
  - vcd, 90
- konvergencia
  - podľa pravdepodobnosti, 109
  - skoro všade, 109
  - v distribúcii, 110
  - v kvadratickom strede, 109
- kritická hodnota, 107
  - Fisherovho rozdelenia, 107
  - χ^2 rozdelenia, 107
  - normálneho rozdelenia, 107
  - Studentovho rozdelenia, 107
- kvantil, 106
  - Fisherovho rozdelenia, 106
  - χ^2 rozdelenia, 106
  - normálneho rozdelenia, 106
  - Studentovho rozdelenia, 106

- metóda
  - BFGS, 71
  - Brent-Dekkerova metóda, 130
  - Brentova, 130
  - dotyčníc, 69
  - Fisherova skóringová, 71
  - Nelder-Meadova, 71
  - Newton-Raphsonova, 69
  - Newtonova, 69
  - quasi Newtonova, 71
  - sečníc, 130
  - simplexov, 71
  - sukcesívnej parabolickej interpolácie, 70
  - zlatého rezu, 70
- model, 34
  - aplikácia na dáta, 33
  - distribučná funkcia, 33
  - funkcia hustoty, 33
  - množina modelov, 34
  - na dáta, 33
  - neparametrický, 40
  - parametrický, 33, 40
  - pre štatistiku, 33

pre chyby, 33
pre testovaciu štatistiku, 33
rozdelenia pravdepodobnosti, 2, 15, 33, 39
štatistická inferencia, 34
štatistický, 2, 11–13, 15, 24, 27, 28, 33

parameter, 39
odhad
asymptoticky nevychýlený, 110
asymptoticky normálny, 111
intervalový, 113
konzistentný, 110
konzistentný a asymptoticky efektívny, 111
lepší, 110
maximálne vierohodný, 63, 64, 69
nevychýlený, 110
vlastnosti odhadu, 112
vychýlený, 110
premenná, 1, 23
mätúca, 11
nezávislá, 2
pozri znak, 23
závislá, 1, 2

rozdelenie
asymptotické, 116
diskrétne, 35
binomické, 47, 208
multinomické, 49, 54
negatívne binomické, 54, 56
Poissonovo, 52, 54, 56
súčinové multinomické, 50
ZIP, 135
funkcie parametra, 209
marginálne, 42, 49, 56
parametra
normálne, 121, 128, 166
pivotovej štatistiky, 114
podmienené, 54
poriadkovej štatistiky, 78
pravdepodobnosti, 2, 15, 33
spojité, 36
dvojrozmerné normálne, 42, 204
exponenciálne, 59
Fisherovo, 106, 107, 109
gama, 56
χ^2, 106
normálne, 41, 75, 76, 84–86, 96, 100, 101, 119, 131, 138
rovnomerné, 59
štandardizované dvojrozmerné normálne, 44
štandardizované normálne, 41, 57, 107
Studentovo, 106–108
stupne voľnosti, 107
symetrické, 80, 99
štatistiky
normálne, 119
testovacej štatistiky
Fisherovo, 118, 199
χ^2, 118, 136, 159
necentrálne Studentovo, 144, 191
normálne, 118, 119, 204, 215–220
Studentovo, 118, 144, 169, 191
zošikmené, 75, 86, 96
ľavostranne, 47, 80, 99
negatívne, 80
pozitívne, 80
pravostranne, 47, 80, 99

štúdia
účely, 1
analytická, 2
biologické súvislosti, 1
ciele, 1, 3
hlavné, 4
vedľajšie, 5
vedecká otázka, 5
deskriptívna, 2
dizajn, 1
experimentálna, 6
hypotézy, 1
interná, 3
klinická, 2
klinické vyšetrenia, 2
kontrola kvality, 5
kostrových pozostatkov, 14
metóda, 3
meta-analýza, 7
metodika, 3
observačná, 6
audit pacientov, 7
kohortová, 9
kvalitatívna, 7
plán, 11
prípadov a kontrol, 9
prípadová, 7
prierezová, 8
sekvenčná, 11
semilongitudinálna, 11
opis štatistickej analýzy, 2
opisná, 2
plán, 1
subjekty, 1
systematický prehľad, 7, 13
technika, 3
translačná, 7
trvanie, 2
validita, 3
externá, 3
konštrukčná, 3
štatistických záverov, 3
validizácia
externá, 5
interná , 5
zákroky, 2
štatistika, 11
abundancia, 11
frekvencia, 11
incidencia, 11
kumulatívna incidencia, 11
kumulatívna prevalencia, 11
pivot, 114
pivotová, 114
počet, 11
početnosť, 11
poriadková, 77
postačujúca, 60
prevalencia, 11
testovacia, 61, 105
pomerom vierohodnosti, 120, 126, 150, 152, 162, 169, 178, 187, 194–196, 202, 206, 222, 233, 258, 262, 266, 269, 271
skóre, 120, 177, 187, 267, 270, 271
typy, 121
Waldova, 120, 122, 144, 159, 166, 169, 175, 185, 189, 191, 199, 204, 211, 212, 215–220, 266, 270, 271

test
štatistický, 105
dobrej zhody
χ^2, 131
Kolmogorov-Smirnovov, 137
Lillieforsov, 139
dvojvýberový
alternatívny Z-test o logaritme pomeru šancí, 219
alternatívny Z-test o logaritme relatívneho rizika, 216
alternatívny Z-test o rozdiele pravdepodobností, 212
alternatívny Z-test o pomere šancí, 220
alternatívny Z-test o relatívnom riziku, 217
F-test o podiele rozptylov, 199
t-test o rozdiele stredných hodnôt, 191
Welchov t-test o rozdiele stredných hodnôt, 191
Z-test o logaritme pomeru šancí, 218

        - Z-test o logaritme relatívneho rizika, 215
        - Z-test o pomere šancí, 220
        - Z-test o relatívnom riziku, 217
        - Z-test o rozdiele korelačných koeficientov, 204
        - Z-test o rozdiele pravdepodobností, 211
        - Z-test o rozdiele stredných hodnôt, 189
    - hladina významnosti, 106, 115, 116
        - pozorovaná, 117
    - hypotéza
        - alternatívna, 105
        - jednostranná, 115–117
        - nezamietanie, 116, 117
        - nulová, 105
        - obojstranná, 115–118
        - zamietanie, 116, 117
    - jednovýberový
        - χ^2-test o rozptyle, 159
        - Studentov t-test o strednej hodnote, 144
        - Z-test o λ, 185
        - Z-test o korelačnom koeficiente, 166
        - Z-test o pravdepodobnosti, 175
        - Z-test o strednej hodnote, 122
    - konzervatívny, 118
    - kritická hodnota, 116, 120
    - liberálny, 118
    - Neyman-Pearsonov prístup, 106
    - oblasť
        - kritická, 105
        - nezamietania, 105, 115
        - zamietania, 105
    - obor
        - kritický, 116, 124, 144, 160, 166, 175, 185, 189, 191, 200, 204, 211, 212
        - nezamietania, 105, 115, 116
        - zamietania, 116
    - p-hodnota, 117, 118, 144, 160, 166, 175, 186, 190, 192, 200, 204, 211, 212, 216, 217, 219, 220
    - párový
        - Studentov t-test o strednej hodnote, 152
    - sila, 105
    - silofunkcia, 106, 124, 144, 160, 166, 175, 185, 189, 191, 200, 204, 211, 212
    - testovacie kritérium, 116
    - výsledok
        - hraničná štatistická signifikancia, 117
        - štatisticky nesignifikantný, 117
        - štatisticky signifikantný, 117
    - viacvýberový
        - Bartlettov test o homogenite rozptylov, 258
        - F-test o rovnosti stredných hodnôt, 230
        - χ^2-test o homogenite korelačných koeficientov, 261
        - χ^2-test o homogenite pravdepodobností, 270
        - χ^2-test o homogenite vektorov pravdepodobností, 271
        - χ^2-test v multinomickom rozdelení, 266

- výber
    - alokačný čas, 28
    - chyba
        - registrácie, 17
        - reprezentatívnosti, 17
    - dobrovoľníci, 17
    - informovaný súhlas, 28, 29
    - infraštruktúra, 28
    - kontrola, 29
    - kritériá
        - prijímacie, 29
        - vylučovacie, 29
    - nábor, 30
    - náhodný, 16
        - homogénny, 85
        - jednoduchý, 16
        - mechanický, 16
        - oblastný, 16
        - skupinový, 16
        - stratifikovaný, 16
        - systematický, 16
        - usporiadaný, 77
    - nepravdepodobnostný, 16
    - počet návštev, 28
    - pomer vyšetrených ku zahrnutým, 28
    - pravdepodobnostný, 16
    - propagácia, 28
    - protokol, 29
    - rozsah, 33
        - minimálny, 27, 124, 167, 176, 205
    - selektívny, 17
    - školenie subjektov, 28
    - štatistické jednotky, 15
    - štatistický súbor, 15
    - typu snehovej gule, 17
    - typu vhodnosti a príležitosti, 17
    - typy, 15
    - výberové jednotky, 15
    - vedecká etika, 29
    - vyšetrenie, 30
    - základný súbor, 15
    - zámerný, 17
    - zápis subjektov do štúdie, 29

- znak, 23
    - kardinálny, 23
    - kvalitatívny, 24
        - alternatívny, 24
        - binárny, 24
        - neusporiadaná kategorizácia, 24
        - nominálny, 24
        - ordinálny, 24
        - usporiadaná kategorizácia, 24
    - kvantitatívny, 23
        - diskrétny, 23
        - frekvencia, 24
        - intervalový, 23
        - spojitý, 23
        - výskyt, 24
    - metrický, 23
    - pozri premenná, 23
    - prognostický, 25
    - realizácie, 24
    - škála merania, 23

Aplikovaná štatistická inferencia I

Biologická antropológia očami matematickej štatistiky

Dnešný antropologický výskum sa bez uplatnenia štatistického uvažovanie nezaobíde. I keď sa väčšina biológov, lekárov ani antropológov štatistikmi nikdy nestanú, výsledky výskumov i možnosti ich publikácie na správnom usporiadaní a štatistickom spracovaní dát silno závisia. Kľúčovým pre obojstranne úspešnú spoluprácu preto je na strane týchto odborov porozumieť základom štatistických postupov, vedieť sa vyjadrovať aspoň základnou štatistickou terminológiou a vedieť štatistikovi vysvetliť svoj výskumný zámer; na strane štatistika potom pochopiť podstatu zamerania jednotlivých odborov, charakter ich metodológie a typické problémy, s ktorými sa po štatistickej stránke stretávajú. Jedným z cieľov tejto knihy je podporiť kladný prístup mladých ľudí k využívaniu štatistických prostriedkov v antropológii a ďalších odboroch a umožniť ich aplikáciu v prostredí programu R, univerzálneho výpočtového prostredia, voľne dostupného všetkým. Táto kniha nie je „kuchárskou knihou" (pretože matematická štatistika nie je knihou receptov alebo návodov), nie je ani „teoretickou matematickou štatistikou" (pretože teória je vysvetlená prostredníctvom príkladov pochádzajúcich z reality praxe biologickej antropológie) a nie je ani „učebnicou prostredia R" (pretože program R je v knihe použitý len ako nástroj, prostredníctvom ktorého sú riešené reálne situácie). Čím teda kniha je? Kniha je unikátnou kombináciou sily teoretického základu matematickej štatistiky implementovaného v prostredí R s cieľom pochopiť a riešiť praktické situácie z biologickej antropológie (ako aj biológie a medicíny). Internetová stránka knihy: http://www.math.muni.cz/pro-studenty/ucebni-text-el-podoba/352-aplikova-statisticka-inferencia.html

Katina, Stanislav (*1976), bioštatistik; v súčasnosti pôsobí ako docent a vedúci Oddelenia aplikovanej matematiky na Ústave matematiky a štatistiky Masarykovej univerzity a tiež ako Honorary Research Fellow na School of Mathematics and Statistics (The University of Glasgow). Dlhodobo pôsobil v zahraničí na univerzitách v USA, Kanade, Nemecku, Rakúsku, Francúzsku a Veľkej Británii. Je členom výkonného výboru International Society for Clinical Biostatistics a členom dvoch subkomisií National Groups a Statistics in Regulatory Affairs v tejto spoločnosti. Vo vedeckej činnosti sa venuje biomedicínskym štatistickým analýzam v kardiológii, neurológii, imunológii a antropológii, štatistickej analýze tvaru a obrazu, časopriestorovému štatistickému modelovaniu elektroencefalografických signálov, elektrických signálov srdca a biogeografických javov, dizajnu klinických, medicínskych a antropologických štúdií, štatistickej vizualizácii a implementácii štatistických a numericko-matematických metód do jazyka R. Kontakt: Oddělení aplikované matematiky, Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Kotlářská 267/2, 611 37, Brno, Česká republika, e-mail: katina@math.muni.cz, tel.: 00420-549-49-7265. Internetová stránka: http://www.muni.cz/sci/people/Stanislav.Katina

Králík, Miroslav (*1973), antropológ; v súčasnosti pôsobí ako docent Ústavu antropológie Prírodovedeckej fakulty Masarykovej univerzity. Zaoberá sa teoretickými aspektmi sexuálneho dimorfizmu, jeho variabilitou, prenatálnou a postnatálnou ontogenézou, významom sexuálneho dimorfizmus z hľadiska ľudského konania a sexuality, a tiež praktickými aplikáciami sexuálneho dimorfizmu vo forenznej oblasti a archeológii. Po metodologickej stránke sa venuje aplikácii a rozvoju moderných morfometrických prístupov v antropológii. Kontakt: Laboratoř morfologie a forenzní antropologie, Ústav antropologie, Přírodovědecká fakulta, Masarykova univerzita, Kotlářská 267/2, 611 37, Brno, Česká republika, e-mail: mirekkralik@seznam.cz, tel.: 00420-549-49-4966. Internetová stránka: http://www.muni.cz/sci/people/Miroslav.Kralik

Hupková, Adela (*1985, rod. Koprdová), antropologička so zameraním na kostrovú a dentálnu antropológiu; v súčasnosti pôsobí ako asistentka Ústavu antropológie Prírodovedeckej fakulty Masarykovej univerzity. Zaoberá sa ontogenézou človeka, vplyvom vývinového stresu na rozvoj bilaterálne symetrických znakov, indikátormi stresu na zuboch a kostrových pozostatkoch človeka. Po metodologickej stránke sa venuje aplikácii histologických analýz mikroštruktúry tvrdých zubných tkanív v antropológii a bioarcheológii. Kontakt: Laboratoř morfologie a forenzní antropologie, Ústav antropologie, Přírodovědecká fakulta, Masarykova univerzita, Kotlářská 267/2, 611 37, Brno, Česká republika, e-mail: hupkova@sci.muni.cz, tel.: 00420-549-49-4073. Internetová stránka: http://www.muni.cz/sci/people/Adela.Hupkova

ISBN 978-80-210-7752-2

9 788021 077522