

# Obsah

<b>1. Úvod</b>	<b>11</b>
<b>2. Obecné úvahy</b>	<b>17</b>
2.1. Rozdíly biologie a matematiky	17
2.2. Přístupy k řešení problémů	18
2.3. Populace a výběr – základ statistické indukce	19
<b>3. Typy sledovaných veličin</b>	<b>23</b>
3.1. Co můžeme sledovat	23
3.2. Typy náhodných veličin	24
3.2.1. Alternativní veličiny	25
3.2.2. Nominální veličiny	25
3.2.3. Ordinální veličiny	26
3.2.4. Kvantitativní veličiny	28
3.2.5. Celočíselné veličiny	30
<b>4. Základní statistické charakteristiky</b>	<b>33</b>
4.1. Míry pro kvalitativní veličiny	34
4.1.1. Pravděpodobnost	35
4.1.2. Relativní četnost	36
4.2. Míry polohy	37
4.2.1. Průměr (aritmetický)	37
4.2.2. Geometrický průměr	38
4.2.3. Medián	39
4.2.4. Modus	40
4.2.5. Useknutý průměr	41
4.2.6. Kvantil	41
4.3. Míry měřítka (variability)	41
4.3.1. Rozptyl	42
4.3.2. Rozpětí	43
4.3.3. Mezikvartilové rozpětí	44
4.3.4. Odhad MAD	44
4.3.5. Variační koeficient	45
4.4. Ostatní charakteristiky	45
4.4.1. Šikmost – skewness	45
4.4.2. Špičatost – kurtosis	46
4.5. Praktické příklady jednotlivých charakteristik	47
<b>5. Modely náhodné veličiny – rozložení pravděpodobnosti</b>	<b>51</b>
5.1. Nominální veličiny	51
5.2. Diskrétní (celočíselné) kvantitativní veličiny	51
5.2.1. Binomické rozložení	51
5.2.2. Multinomické rozložení	52

5.2.3.	Poissonovo rozložení . . . . .	53
5.2.4.	Negativně binomické (Pascalovo) rozložení . . . . .	54
5.2.5.	Nakažlivá rozložení . . . . .	54
5.3.	Spojité kvantitativní veličiny . . . . .	55
5.3.1.	Normální (Gaussovo) rozložení . . . . .	55
5.3.2.	Logaritmicko-normální rozložení . . . . .	57
5.3.3.	Exponenciální rozložení . . . . .	58
5.3.4.	Weibullovo rozložení . . . . .	58
5.3.5.	Rovnoměrné rozložení . . . . .	59
5.3.6.	Logistické rozložení . . . . .	59
5.4.	Výběrová rozložení – rozložení testovacích statistik . . . . .	60
5.4.1.	$\chi^2$ -rozložení . . . . .	61
5.4.2.	Studentovo t-rozložení . . . . .	61
5.4.3.	Fisherovo F-rozložení . . . . .	62
<b>6.</b>	<b>Statistické odhady a testy – základní principy</b> . . . . .	<b>65</b>
6.1.	Odhady populačních charakteristik . . . . .	65
6.2.	Bodové odhady . . . . .	66
6.3.	Intervalové odhady . . . . .	66
6.3.1.	Intervalové odhady populačních charakteristik – intervaly spolehlivosti . . . . .	68
6.3.2.	Intervalové odhady – predikční intervaly . . . . .	70
6.3.3.	Intervalové odhady – toleranční intervaly . . . . .	71
6.4.	Rozdíl interpretace intervalu spolehlivosti a tolerančního intervalu . . . . .	71
6.5.	Obecné principy při konstrukci odhadů . . . . .	74
6.6.	Statistické testy . . . . .	75
6.7.	Nové možnosti výpočetní techniky . . . . .	78
<b>7.</b>	<b>Ověřování typu rozložení dat – klíč k volbě modelu</b> . . . . .	<b>81</b>
7.1.	Grafické zobrazení výběrového rozložení . . . . .	81
7.2.	Testy k ověření typu rozložení . . . . .	85
7.2.1.	$\chi^2$ – testy dobré shody . . . . .	85
7.2.2.	Kolmogorův-Smirnovův test . . . . .	86
7.2.3.	Test normality Šapirův-Wilkův . . . . .	87
7.2.4.	Další možnosti . . . . .	87
7.3.	Význam znalosti typu rozložení . . . . .	88
<b>8.</b>	<b>Porovnání kvantitativní veličiny jednoho výběru s pevnou hodnotou</b> . . . . .	<b>93</b>
8.1.	Testy charakteristik . . . . .	93
8.1.1.	Jednovýběrový Z-test . . . . .	94
8.1.2.	Jednovýběrový t-test . . . . .	95
8.1.3.	Jednovýběrový znaménkový (mediánový) test . . . . .	97
8.1.4.	Jednovýběrový Wilcoxonův test . . . . .	98
8.2.	Intervalové odhady . . . . .	100
8.2.1.	Intervaly spolehlivosti . . . . .	100
8.2.2.	Predikční intervaly . . . . .	102
8.2.3.	Toleranční intervaly . . . . .	103
8.2.4.	Konstrukce intervalových odhadů metodou bootstrap . . . . .	104
8.2.5.	Praktické ukázky intervalových odhadů . . . . .	104
8.2.6.	Co nejsou intervalové odhady . . . . .	104

<b>9. Porovnání kvantitativní veličiny ve dvou různých výběrech</b>	<b>109</b>
9.1. Dvě skupiny . . . . .	109
9.1.1. Dvouvýběrový t-test . . . . .	109
9.1.2. Porovnání dvou rozptylů . . . . .	111
9.1.3. Dvouvýběrový znaménkový test (mediánový) . . . . .	111
9.1.4. Dvouvýběrový Wilcoxonův test . . . . .	112
9.1.5. Poznámka k testům porovnání dvou skupin . . . . .	113
9.2. Párové porovnání . . . . .	113
9.2.1. Párový t-test . . . . .	114
9.2.2. Párový znaménkový test (mediánový) . . . . .	114
9.2.3. Párový Wilcoxonův test . . . . .	114
9.2.4. Praktické použití párových porovnání . . . . .	115
9.2.5. Několik poznámek k párovému testu a korelaci . . . . .	117
<b>10. Analýza vztahu dvou spojitých veličin</b>	<b>121</b>
10.1. Společné rozložení dvou veličin . . . . .	121
10.2. Kovariance – míra lineárního vztahu dvou veličin . . . . .	124
10.3. Koeficient lineární korelace . . . . .	125
10.4. Robustní varianty korelačních koeficientů . . . . .	129
10.4.1. Spearmanův koeficient monotónní korelace . . . . .	129
10.4.2. Kendallův koeficient monotónní korelace . . . . .	129
10.5. Praktické ukázky různých typů závislostí . . . . .	130
10.6. Lineární regresní model . . . . .	131
10.6.1. Lineární regresní model normálně rozložené náhodné veličiny . . . . .	132
10.6.2. Regresní modely procházející počátkem (bez interceptu) – regrese procházející počátkem . . . . .	136
10.6.3. Vztah regresního lineárního modelu a lineárního korelačního koeficientu . . . . .	138
10.6.4. Oblasti spolehlivosti – intervalové odhady . . . . .	141
10.6.5. Problémy s linearitou a normalitou – transformace modelu . . . . .	143
10.6.6. Ověření předpokladu lineárního regresního modelu . . . . .	144
10.6.7. Odlehlá pozorování v regresi . . . . .	147
10.6.8. Analýza reziduí . . . . .	149
10.7. Vztah více než dvou veličin . . . . .	154
10.7.1. Vícenásobná regrese . . . . .	155
10.7.2. Vícerozměrná regrese . . . . .	156
10.7.3. Korelace více veličin . . . . .	157
10.7.4. Porovnání modelů . . . . .	159
10.7.5. Polynomická regrese . . . . .	161
10.8. Nelineární regrese . . . . .	163
10.9. Robustní regresní metody . . . . .	165
10.10. Metody vyhlazování časových řad . . . . .	168
<b>11. Porovnání kvantitativní veličiny ve více skupinách – analýza rozptylu – ANOVA</b>	<b>171</b>
11.1. Podmínky použitelnosti analýzy rozptylu . . . . .	173
11.1.1. Test shody rozptylů . . . . .	174
11.2. Více skupin – analýza rozptylu jednoduchého třídění – způsob výpočtu . . . . .	176
11.2.1. Kontrasty . . . . .	181
11.3. Metody mnohonásobného srovnávání . . . . .	184
11.4. Neparametrické varianty analýzy rozptylu . . . . .	188
11.5. Vztah mezi regresi a analýzou rozptylu . . . . .	189

11.6. Analýza rozptylu dvojného třídění . . . . .	190
11.7. Opakovaná pozorování . . . . .	196
11.8. Testování modelu a „podmodelu“ . . . . .	198
11.9. Obecnější modely analýzy rozptylu . . . . .	199
11.10 Model se smíšenými efekty . . . . .	200
11.10.1 Párový t-test pomocí modelu se smíšenými efekty . . . . .	202
11.10.2 Dvouvýběrový t-test pomocí modelu se smíšenými efekty . . . . .	203
11.10.3 Obecnější model smíšených efektů . . . . .	203
<b>12. Kvalitativní veličiny a jejich vztah</b> . . . . .	<b>207</b>
12.1. Odhad a testy pravděpodobnosti alternativní veličiny . . . . .	207
12.1.1. Aproximace normálním rozložením . . . . .	207
12.1.2. Fleissova kvadratická aproximace . . . . .	208
12.1.3. Exaktní binomický test . . . . .	208
12.2. Obecná kontingenční tabulka . . . . .	209
12.3. Kontingenční tabulka $2 \times 2$ . . . . .	213
12.3.1. Míry vztahu dvou alternativních veličin . . . . .	218
12.3.2. McNemarova hypotéza symetrie . . . . .	221
12.3.3. Shoda dvou hodnotitelů . . . . .	223
12.4. Typy studií – způsoby konstrukce kontingenčních tabulek . . . . .	224
12.4.1. Průřezová studie . . . . .	225
12.4.2. Kohortová studie . . . . .	225
12.4.3. Studie případů-kontrol . . . . .	225
12.4.4. Typy studií a míry nezávislosti . . . . .	226
12.4.5. Studie typu případů a kontrola . . . . .	226
12.4.6. Průřezová studie . . . . .	227
12.4.7. Kohortová studie . . . . .	227
12.5. Stratifikované kontingenční tabulky . . . . .	228
12.6. Test trendu v kontingenční tabulce . . . . .	232
12.7. Souvislost testů pro kategoriální a spojité veličiny . . . . .	234
12.8. Intenzita incidence . . . . .	235
12.9. Hodnocení kvality screeningových testů . . . . .	237
12.10 ROC křivky . . . . .	240
<b>13. Výběr a jeho reprezentativnost</b> . . . . .	<b>243</b>
13.1. Rušivé faktory . . . . .	244
13.2. Konstrukce výběru pro studie popisující populaci . . . . .	244
13.3. Plány experimentu . . . . .	246
13.3.1. Rozdělení na skupiny (do větví) . . . . .	246
13.3.2. Volba kontrolní skupiny . . . . .	246
13.3.3. Použití placeba . . . . .	247
13.3.4. Párové uspořádání dat . . . . .	248
13.3.5. Křížový pokus . . . . .	248
13.4. Stanovení rozsahu výběru . . . . .	249
13.4.1. Rozsah výběru pro jednovýběrový t-test . . . . .	249
13.4.2. Rozsah výběru pro dvouvýběrový t-test . . . . .	250
13.4.3. Rozsah výběru pro test binomické veličiny . . . . .	251
13.5. Metoda vážení . . . . .	252
13.6. Standardizace . . . . .	253
13.6.1. Přímá standardizace . . . . .	255

13.6.2. Nepřímá standardizace . . . . .	256
13.6.3. Inverzní standardizace . . . . .	256
13.6.4. Intervaly spolehlivosti pro standardizované ukazatele . . . . .	257
<b>14. Další modely pro studium závislosti veličin</b>	<b>261</b>
14.1. Logistická regrese – model závislosti alternativní veličiny . . . . .	261
14.2. Další modely pro alternativní veličinu . . . . .	266
14.2.1. Účinná dávka ED50 či LD50 . . . . .	266
14.3. Poissonovská regrese – model závislosti počtů na spojité či kvalitativní veličině . . . . .	267
<b>15. Analýza cenzorovaných dat</b>	<b>269</b>
15.1. Neúplná informace – cenzorovaná data . . . . .	269
15.2. Analýza přežití . . . . .	271
15.2.1. Odhad doby do události (doby přežití) . . . . .	273
15.2.2. Tabulky přežití . . . . .	274
15.2.3. Neparаметrické metody . . . . .	275
15.2.4. Semiparametrické metody . . . . .	279
15.2.5. Parametrické metody . . . . .	281
15.2.6. Složitější parametrické modely pro analýzu přežití . . . . .	283
15.2.7. Rozdíly mezi neparаметrickým, parametrickým a semiparametrickým pří- stupem . . . . .	284
15.3. Cenzorovaná data – hodnoty pod detekčním limitem . . . . .	284
15.4. Použití analýzy cenzorovaných dat k odfiltrování epidemií . . . . .	286
15.4.1. Nalezení epidemického prahu . . . . .	286
15.4.2. Odhad počtu úmrtí zvýšeného výskytem epidemie . . . . .	288
15.4.3. Složitější modely pro nalezení odhadu očekávaného výskytu – „baseline“. . . . .	291
<b>A. Jemný úvod do programu R</b>	<b>295</b>
<b>B. Využití výpočetní techniky pro statistická hodnocení</b>	<b>313</b>
<b>C. Grafy – užitečný nástroj interpretace a jejich úskalí</b>	<b>315</b>
<b>D. Ukázky chybných použití statistiky</b>	<b>321</b>
D.1. Chyby při používání statistiky a interpretaci výsledků analýz . . . . .	321
D.2. Cestou statistiky i medicíny k stejnému závěru . . . . .	330
<b>E. Data a skripty k jednotlivým kapitolám</b>	<b>331</b>
<b>Literatura</b>	<b>333</b>
<b>Rejstřík</b>	<b>337</b>