

1	Úvodem	13
2	Dříve les než stromy	17
2.1	Nejednoznačnost terminologie	17
2.2	Volba metody analýzy dat	23
2.3	Přehled vybraných vícerozměrných metod	25
2.3.1	Metoda hlavních komponent	26
2.3.2	Faktorová analýza	27
2.3.3	Kanonická korelační analýza	28
2.3.4	Regresní analýza	28
2.3.5	Analýza rozptylu	30
2.3.6	Analýza kovariance	31
2.3.7	Diskriminační analýza	31
2.3.8	Logistická regrese	32
2.3.9	Analýza kategoriálních dat	33
2.3.10	Shluková analýza	34
2.3.11	Vícerozměrné škálování	34
2.3.12	Korespondenční analýza a vizualizace dat	35
2.3.13	Analýza conjoint	35
2.4	Klasifikace vícerozměrných statistických metod	36
	<i>Literatura</i>	40
3	Operace s maticemi	41
3.1	Maticový počet pro potřeby této knihy	41
3.2	Vektory a matice	42
3.3	Některé věty maticového počtu	47
3.4	Rozšíření maticového počtu	51
3.4.1	Ortogonální matice	51
3.4.2	Idempotentní matice	53
3.4.3	Charakteristická čísla a charakteristické vektory	54
3.5	Lineární a kvadratické formy	57
3.6	Pozitivně definitní a pozitivně semidefinitní matice	58
3.7	Odmocninová matice	59
3.8	Singulární rozklad	60
3.9	Pravidla pro derivování funkcí vektorů a matic	61
	<i>Literatura</i>	62

4	Rozdělení náhodných veličin	63
4.1	Náhodné jevy a jejich pravděpodobnosti	63
4.2	Pravidla pro počítání s pravděpodobnostmi	67
4.3	Rozdělení náhodných veličin	73
4.4	Rozdělení náhodného vektoru	86
4.5	Charakteristiky náhodného vektoru	88
4.6	Další modely diskretních a spojitých rozdělení	90
4.6.1	Vícerozměrné rozšíření diskretních rozdělení	91
4.6.2	Vícerozměrné normální rozdělení	98
4.6.3	Výběrová spojitá jednorozměrná rozdělení	102
4.6.4	Výběrová vícerozměrná spojitá rozdělení	105
	<i>Literatura</i>	110
5	Data	111
5.1	Metody pořizování dat	112
5.2	Datová matice	120
5.3	Soubor objektů a jeho rozsah	122
5.4	Proměnné v datové matici	123
	<i>Literatura</i>	128
6	Bayesovské a klasické myšlení	129
6.1	Oprávněná kritika statistických úsudků	129
6.2	Pohled do historie	132
6.3	Bayesovské a četnostní pojetí pravděpodobnosti	136
6.4	Bayesův vzorec pro náhodné jevy a hypotézy	143
6.5	Bodové odhady	147
6.5.1	Klasický přístup k bodovým odhadům	147
6.5.2	Bayesovský přístup k bodovým odhadům	151
6.5.3	Hodnocení kvality bayesovského bodového odhadu	153
6.6	Intervalové odhady	160
6.7	Testování hypotéz	162
6.8	Výpočetní aspekty bayesovské statistiky	168
	<i>Literatura</i>	176
7	Ověřování kvality dat	179
7.1	Důvody a nástroje zkoumání dat	179
7.2	Chybějící údaje	183
7.3	Intervalové a skupinové rozdělení četnosti	185
7.4	Vyhledávání odlehlých pozorování	187
7.4.1	Identifikace jednorozměrných odlehlých pozorování	190
7.4.2	Identifikace vícerozměrných odlehlých pozorování	194
7.5	Předpoklad normálního rozdělení	199
7.6	Grafické posouzení normálního rozdělení	201
7.6.1	Grafické posouzení jednorozměrného normálního rozdělení	201

7.6.2	Grafické posouzení vícerozměrného normálního rozdělení	206
7.7	Ověřování typu rozdělení pomocí statistických testů	209
7.7.1	Chí-kvadrát test dobré shody	209
7.7.2	Výběrová distribuční funkce a Kolmogorovův test	211
7.7.3	Testy vycházející z šikmosti a špičatosti rozdělení	217
7.7.4	Shapiroův a Wilkův test normality	222
7.7.5	D'Agostinův test	225
7.7.6	Test založený na studentizovaném rozpětí	227
7.7.7	Andersonův-Darlingův test	228
7.8	Transformace dat	231
<i>Literatura</i>		233
8	Závislost a vztah dvou proměnných	235
8.1	Dvě kvantitativní proměnné	235
8.2	Asociace a korelace	237
8.2.1	Pearsonův korelační koeficient	239
8.2.2	Regresní funkce s jednou vysvětlující proměnnou	243
8.3	Jednofaktorová analýza rozptylu	255
8.3.1	Vliv třídícího faktoru na úroveň kvantitativní proměnné	257
8.3.2	Mnohonásobná porovnávání (simultánní úsudky)	263
8.4	Dvě kategoriální proměnné	267
8.4.1	Výběrové míry asociace	269
8.4.2	Testy nezávislosti	276
<i>Literatura</i>		281
9	Vektor středních hodnot	283
9.1	Úsudky o středních hodnotách	283
9.2	Úsudky o střední hodnotě normálního rozdělení	284
9.3	Úsudky o vektoru středních hodnot normálního rozdělení	285
9.4	Úsudky o složkách vektoru středních hodnot	289
9.5	Úsudky o vektoru středních hodnot při velkých výběrech	297
9.6	Intervaly a testy pro četnosti při velkých výběrech	298
9.7	Test o lineární formě vektoru středních hodnot	300
<i>Literatura</i>		302
10	Dva vektory středních hodnot	303
10.1	Úsudky při srovnání výběrů ze dvou populací	303
10.2	Dva nezávislé jednorozměrné náhodné výběry	303
10.3	Dva nezávislé vícerozměrné výběry	308
10.4	Simultánní úsudky o shodě složek dvou vektorů	314
10.5	Nestejně kovarianční matice	315
10.6	Ověření shody středních hodnot při dvou závislých výběrech	316
<i>Literatura</i>		320

11	Kovarianční a korelační matice	321
11.1	Úsudky o kovarianční matici	321
11.2	Testy shody rozptylů	325
11.3	Test shody kovariančních matic	328
11.4	Populační korelační koeficienty	333
11.5	Induktivní úsudky o korelačních koeficientech	335
11.6	Důsledky vztahů mezi korelačními koeficienty	341
11.7	Ověření úplné nezávislosti proměnných	342
11.8	Ověření nezávislosti mezi skupinami proměnných	343
11.9	Robustní odhady jednoduchého korelačního koeficientu	344
	<i>Literatura</i>	347
12	Metoda hlavních komponent	349
12.1	Hlavní komponenty	349
12.2	Cíle analýzy hlavních komponent	349
12.3	Hlavní komponenty v populaci	351
12.4	Hlavní komponenty ve výběru	357
12.5	Geometrický význam hlavních komponent	364
12.6	Shrnutí metody hlavních komponent	364
	<i>Literatura</i>	376
13	Faktorová analýza	377
13.1	Explorativní a konfirmativní faktorová analýza	377
13.2	Historie a názory na faktorovou analýzu	378
13.3	Model faktorové analýzy	379
13.4	Nejednoznačnost faktorového modelu	382
13.5	Odhad parametrů faktorového modelu	383
13.6	Řešení faktorových rovnic a počet faktorů	389
13.7	Jednoduchá struktura a rotace faktorů	392
13.8	Odhad faktorového skóre	397
	<i>Literatura</i>	409
14	Kanonická korelační analýza	411
14.1	Úvod do kanonické analýzy	411
14.2	Kanonické veličiny ve statistických metodách	412
14.3	Korelace dvou skupin proměnných	414
14.4	Kanonická korelace v populaci	415
14.5	Kanonická korelace ve výběru	417
14.6	Vztahy mezi původními a kanonickými veličinami	419
	<i>Literatura</i>	427
15	Vícenásobná lineární regrese	429
15.1	Statistické modelování závislosti	429
15.1.1	Měření závislosti a regrese není totéž	430

15.1.2	Heuristický úvod a skrytá korelace	434
15.2	Regrese a korelace	438
15.2.1	Úkoly regresní a korelační analýzy	442
15.2.2	Regresní modely a jejich klasifikace	444
15.2.3	Vyrovňovací kritéria	448
15.3	Lineární regresní model	453
15.3.1	Klasický lineární regresní model	457
15.3.2	Odhad parametrů regresní funkce	460
15.3.3	Metoda maximální věrohodnosti v regresi	463
15.3.4	Bodové odhady v lineárním regresním modelu	465
15.3.5	Odhad lineární funkce regresních koeficientů	469
15.4	Intervaly spolehlivosti a testy hypotéz v KLM	478
15.4.1	Intervaly spolehlivosti v KLM	479
15.4.2	Test obecné lineární hypotézy	485
15.4.3	Testy hypotéz v KLM	488
15.5	Hodnocení kvality regresního modelu	497
15.5.1	Rozklady součtů čtverců	498
15.5.2	Analýza reziduí a vlivná pozorování	504
15.5.3	Projekční matice	505
15.5.4	Různé typy reziduí a jejich vlastností	507
15.5.5	Vlivná pozorování	511
15.5.6	Vysvětlující proměnné a kvalita modelu	513
15.5.7	Chybná specifikace proměnných	514
15.5.8	Výběr vysvětlujících proměnných	516
15.5.9	Kritéria výběru proměnných	518
15.5.10	Metody výběru podmnožiny proměnných	521
15.6	Nesplněné podmínky KLM	525
15.6.1	Zobecněný lineární model	527
15.6.2	Heteroskedasticita	530
15.6.3	Autokorelace	539
15.6.4	Náhodné vysvětlující proměnné	548
15.6.5	Multikolinearita	549
15.6.6	Transformace dat a modelu	558
15.7	Nevýběrová informace ve formě přesných lineárních omezení	560
	<i>Literatura</i>	566
16	Obecný lineární model	569
16.1	Základní pojmy	569
16.2	Analýza rozptylu – jednorozměrné úlohy s více faktory	579
16.2.1	Faktoriální plán	580
16.2.2	Uspořádání do bloků, model bez interakce	587
16.2.3	Hierarchický experiment	590
16.2.4	Opakovaná měření	593
16.3	Vícerozměrná analýza rozptylu	602
16.3.1	Jeden faktor	603
16.3.2	Faktoriální návrh, dva faktory	612

OBSAH

16.3.3	Longitudinální data	617
16.3.4	Analýza rozptylu s využitím kanonických proměnných	620
16.4	Analýza kovariance	624
16.4.1	Jeden faktor, jedna doprovodná proměnná	624
16.4.2	Předpoklady modelu	627
16.4.3	Testované hypotézy	630
16.4.4	Dva faktory, jedna doprovodná proměnná	633
16.5	Vícerozměrná analýza kovariance	637
16.6	Vícerozměrný regresní model	644
16.6.1	Vícerozměrné testy v regresním modelu	646
16.6.2	Simultánní intervaly v regresním modelu	650
16.7	Odhad parametrů a testování hypotéz v GLM	652
16.7.1	Odhadnutelné parametrické funkce	653
16.7.2	Testy hypotéz v GLM	657
16.7.3	Typ I součtu čtverců	658
16.7.4	Typ II součtu čtverců	661
16.7.5	Typ III součtu čtverců	662
16.7.6	Typ IV součtu čtverců	665
16.7.7	Typ V součtu čtverců	668
16.7.8	Typ VI součtu čtverců	668
16.7.9	Porovnání regresních přímek	668
16.7.10	Konstrukce obecného lineárního modelu	672
<i>Literatura</i>		678
17	Diskriminační analýza	681
17.1	Podstata a cíl diskriminační analýzy	681
17.2	Diskriminace dvou vícerozměrných normálních rozdělení	682
17.3	Diskriminace v případě tří a více skupin	691
17.4	Kanonická diskriminační analýza	696
<i>Literatura</i>		710
18	Logistická regrese	711
18.1	Alternativní vysvětlovaná proměnná	711
18.2	Kategoriální vysvětlovaná proměnná	723
<i>Literatura</i>		728
19	Kategoriální proměnné	729
19.1	Vícerozměrná kontingenční tabulka	729
19.2	Loglineární model	731
19.3	Odhad parametrů loglineárního modelu	736
19.4	Volba vhodného loglineárního modelu	738
19.5	Loglineární a logitový model	742
<i>Literatura</i>		744

20	Shluková analýza	745
20.1	Cíle shlukové analýzy	745
20.2	Klasifikace metod shlukové analýzy	747
20.3	Míry podobnosti a vzdálenosti	748
20.3.1	Hodnocení vztahu dvou objektů	749
20.3.2	Hodnocení vztahu dvou proměnných	754
20.3.3	Hodnocení vztahu dvou kategorií nominální proměnné	755
20.3.4	Hodnocení vztahu dvou shluků	755
20.4	Metody k-shlukování	757
20.4.1	Algoritmy k-centroidů	757
20.4.2	Algoritmy k-medoidů	760
20.4.3	Algoritmy fuzzy k-shlukování	761
20.4.4	Algoritmy založené na modelu	764
20.5	Metody hierarchického shlukování	764
20.5.1	Aglomerativní algoritmy	764
20.5.2	Divizivní algoritmy	769
20.5.3	Dvourozměrné shlukování	771
20.6	Speciální metody shlukování	771
20.6.1	Algoritmy pro shlukování velkých souborů	772
20.6.2	Algoritmy pro shlukování s kategoriálními proměnnými	775
20.6.3	Algoritmy založené na hustotě a mřížce	776
20.7	Hodnocení vytvořených shluků	776
20.7.1	Hodnocení výsledků disjunktčního shlukování	777
20.7.2	Hodnocení výsledků fuzzy shlukování	780
	<i>Literatura</i>	781
21	Vícerozměrné škálování	783
21.1	Podstata metody	783
21.2	Vstupní data pro vícerozměrné škálování	783
21.2.1	Vzdálenost objektů	784
21.2.2	Podobnost objektů	784
21.3	Modely MDS	785
21.4	Metrické MDS	785
21.4.1	Klasické MDS	785
21.4.2	Odvození souřadnic z měř nepodobnosti	787
21.5	Nemetrické MDS	790
21.5.1	Ztrátová funkce	791
21.5.2	Výpočetní algoritmus	793
21.6	Modely individuálních diferencí	799
21.6.1	Vážený euklidovský model	799
21.6.2	Zobecněný euklidovský model	802
	<i>Literatura</i>	804

OBSAH

22	Korespondenční analýza	805
22.1	Podstata korespondenční analýzy	805
22.2	Elementární analýza kontingenčních tabulek	806
22.3	Jednoduchá korespondenční analýza	807
22.4	Míry vzdálenosti	810
22.5	Výpočetní algoritmus	811
22.6	Korespondenční mapa	815
22.7	Hodnocení a interpretace modelu	817
22.8	Vícenásobná korespondenční analýza	820
<i>Literatura</i>		825
23	Analýza conjoint	827
23.1	Základní pojmy a principy v analýze conjoint	827
23.2	Tradiční conjoint	829
23.3	Metoda částečného profilu	834
23.4	Hybridní modely	835
23.5	Metoda založená na výběru	837
23.6	Modelování heterogenity pomocí bayesovských modelů	840
23.7	Současné trendy v analýze conjoint	841
23.7.1	Adaptivní metoda založená na výběru	842
23.7.2	Metoda založená na volbě z nabídky (MBC)	845
<i>Literatura</i>		847
24	Vizualizace vícerozměrných dat	849
24.1	Cíle vizualizace	849
24.2	Metody vizualizace	850
24.2.1	Paralelní osy	850
24.2.2	Bertinovy (permutační) matice	852
24.2.3	Metoda projekčního prohledávání	856
24.2.4	Metoda RadViz	857
24.2.5	Kohonenovy mapy	860
24.3	Dynamická grafická analýza a grafický software	863
24.4	Grafická analýza korelační matice	864
<i>Literatura</i>		866
Rejstřík		867
Summary		877