

Obsah

ÚVOD	7
1. ANALÝZA ROZPTYLU	9
1.1. Základné predpoklady použitia a princíp analýzy rozptylu	10
1.1.1. Jednorozmerná ANOVA – jednoduché triedenie	10
1.1.2. Jednorozmerná ANOVA – dvojné triedenie bez interakcie	14
1.1.3. Jednorozmerná ANOVA – dvojné triedenie s interakciou	15
1.1.4. Viacrozmerná analýza rozptylu (MANOVA)	17
1.2. Analýza rozptylu v programe SPSS	19
1.3. Analýza rozptylu v programe R	27
2. DISKRIMINAČNÁ ANALÝZA	35
2.1. Základné predpoklady použitia a princíp diskriminačnej analýzy	35
2.1.1. Kanonická diskriminačná analýza	37
2.1.2. Lineárna diskriminačná analýza	44
2.1.3. Kvadratická diskriminačná analýza	44
2.2. Diskriminačná analýza v programe SPSS	45
2.3. Diskriminačná analýza v programe R	50
3. LOGISTICKÁ REGRESIA	56
3.1. Základné predpoklady použitia a princíp logistickej regresie	56
3.1.1. Odhad parametrov metódou maximálnej vierohodnosti	57
3.1.2. Hodnotenie kvality modelu a jeho interpretácia	60
3.2. Logistická regresia v programe SPSS	63
3.3. Logistická regresia v programe R	72
4. ANALÝZA HLAVNÝCH KOMPONENTOV	88
4.1. Základné predpoklady použitia a princíp analýzy hlavných komponentov	88
4.2. Analýza hlavných komponentov v programe SPSS	95
4.3. Analýza hlavných komponentov v programe R	103
5. FAKTOROVÁ ANALÝZA	112
5.1. Základné predpoklady použitia a princíp faktorovej analýzy	112
5.2. Faktorová analýza v programe SPSS	125
5.3. Faktorová analýza v programe R	130

6. ZHLUKOVÁ ANALÝZA	137
6.1. Základné predpoklady použitia a princíp zhlukovej analýzy	138
6.1.1. Miery podobnosti	138
6.1.2. Zhlukovacie metódy	141
6.2. Zhluková analýza v programe SPSS	147
6.2.1. Hierarchické zhlukovanie v programe SPSS	147
6.2.2. Nehierarchické zhlukovanie v programe SPSS	149
6.3. Zhluková analýza v programe R	154
7. KOREŠPONDENČNÁ ANALÝZA	163
7.1. Základné predpoklady použitia a princíp korešpondenčnej analýzy	163
7.2. Korešpondenčná analýza v programe SPSS	166
7.3. Korešpondenčná analýza v programe R	170

Contents

INTRODUCTION	7
1. ANALYSIS OF VARIANCE	9
1.1. Basic Assumptions and Principles of ANOVA	10
1.1.1. One-way ANOVA	10
1.1.2. Two-way ANOVA without Interaction Effects	14
1.1.3. Two-way ANOVA with Interaction Effects	15
1.1.4. Multivariate Analysis of Variance (MANOVA)	17
1.2. ANOVA in SPSS	19
1.3. ANOVA in R	27
2. DISCRIMINANT ANALYSIS	35
2.1. Basic Assumptions and Principles of Discriminant Analysis	35
2.1.1. Canonical Discriminant Analysis	37
2.1.2. Linear Discriminant Analysis	44
2.1.3. Quadratic Discriminant Analysis	44
2.2. Discriminant Analysis in SPSS	45
2.3. Discriminant Analysis in R	50
3. LOGISTIC REGRESSION	56
3.1. Basic Assumptions and Principles of Logistic Regression	56
3.1.1. Maximum Likelihood Parameter Estimation	57
3.1.2. Quality and Interpretation of Models	60
3.2. Logistic Regression in SPSS	63
3.3. Logistic Regression in R	72
4. PRINCIPAL COMPONENT ANALYSIS	88
4.1. Basic Assumptions and Principles of Principal Component Analysis	88
4.2. Principal Component Analysis in SPSS	95
4.3. Principal Component Analysis in R	103
5. FACTOR ANALYSIS	112
5.1. Basic Assumptions and Principles of Factor Analysis	112
5.2. Factor Analysis in SPSS	125
5.3. Factor Analysis in R	130

6. CLUSTER ANALYSIS	137
6.1. Basic Assumptions and Principles of Cluster Analysis	138
6.1.1. Similarity Measures	138
6.1.2. Clustering Methods	141
6.2. Cluster Analysis in SPSS	147
6.2.1. Hierarchical Clustering in SPSS	147
6.2.2. Non-hierarchical Clustering in SPSS	149
6.3. Cluster Analysis in R	154
7. CORRESPONDENCE ANALYSIS	163
7.1. Basic Assumptions and Principles of Correspondence Analysis	163
7.2. Correspondence Analysis in SPSS	166
7.3. Correspondence Analysis in R	170

ÚVOD

V ekonomickej praxi sa veľmi často stretávame so situáciami, keď musíme na základe určitých vstupných informácií uskutočniť rozhodnutie, resp. vybrať najvýhodnejšiu možnosť spomedzi viacerých variantov. V triviálnych prípadoch je možné sa rozhodovať intuitívne na základe predchádzajúcich skúseností, resp. okamžitého posúdenia dôležitosti faktorov vplývajúcich na daný jav. Tento spôsob rozhodovania využívajú v ekonomickej praxi najmä malí podnikatelia, ktorí týmto spôsobom riešia väčšinu svojich operatívnych úloh. Zložitosť ekonomických problémov spočíva najmä v množstve faktorov, ktoré viac alebo menej intenzívne daný jav ovplyvňujú. Nie vždy je možné a vhodné upriamiť pozornosť na jedinú vlastnosť skúmaného ekonomického javu (charakterizovanú jedným štatistickým znakom), toto zjednodušenie je často veľmi vzdialené od reálneho ekonomického sveta. Práve zohľadnenie viacerých (prípadne všetkých) relevantných aspektov, sledovaním viacerých štatistických znakov, vyúsťuje k omnoho väčšiemu priblíženiu analyzovaných javov k ekonomickej realite. Táto skutočnosť vedie ekonómov k použitiu viacerozmerných štatistických metód ako jedného z dôležitých nástrojov na podporu rozhodovania v manažérskej praxi. Cieľom predkladanej učebnice je prezentovať vybrané viacerozmerné štatistické metódy ako základný nástroj na podporu rozhodovania kvantitatívne orientovaného ekonóma a poukázať na ich relatívne ľahkú aplikáciu pri využití vhodného štatistického softvéru.

Na základe uvedeného predpokladáme, že čitateľmi tejto učebnice budú najmä študenti a absolventi študijných odborov ekonomického zamerania, no mala by byť užitočnou pomôckou pre každého, kto má záujem osvojiť si základy viacerozmerných štatistických metód pre potreby ich praktických aplikácií. Štúdiom tejto knihy čitateľ získa základný prehľad o vybraných klasických viacerozmerných štatistických metódach. Tento prehľad by mal byť dostatočný na to, aby získal primeranú predstavu o možnostiach využitia jednotlivých štatistických metód, o ich predpokladoch a obmedzeniach a mohol v prípade veľmi jednoduchých situácií analogických s príkladmi prezentovanými v tejto knihe vytvoriť vhodné modely s využitím štatistického softvéru a korektne ich interpretovať. Predkladaná publikácia by mala čitateľovi zároveň umožniť aktívnu spoluprácu s profesionálnym štatistikom pri riešení zložitejších praktických problémov. Predpokladáme, že čitateľ má znalosti z deskriptívnej štatistiky, zo základov výberového skúmania a regresnej a korelačnej analýzy približne v rozsahu základného kurzu štatistiky na školách ekonomického zamerania. Predpokladáme tiež znalosti v rozsahu základného kurzu matematiky na tomto type škôl.

V súčasnosti nie je možné predstaviť si úspešnú aplikáciu štatistiky bez použitia kvalitného štatistického programu. Na druhej strane často dochádza k nesprávnemu používaniu štatistických metód a chybnjej interpretácii výsledkov, ktorá je spôsobená nedostatočným osvojením si teoretických predpokladov a základných princípov štatistických