

*Bohdaloč*

2/2005

# FORUM STATISTICUM SLOVACUM



ISSN 1336-7420

52

977 1336 742001

## OBSAH

|  |     |
|--|-----|
| Chajdiak Jozef, Luha Ján, Z histórie seminárov Výpočtová štatistika  | 2   |
| Bačišin Vladimír: Inovácie pohľadom teórií QWERTY a Path Dependency  | 4   |
| Bartošová Jitka: Validity of the logarithm-normal model of HOUSEHOLD Income Distribution in the Czech Republic   | 9   |
| Bohdalová mária: Monte Carlo simulačná štúdia pre metódu ANOVA v systéme SAS®  | 14  |
| Garaj Ivan: Porovnanie Akahirovej aproximácie s exaktným výpočtom jednostranných tolerančných činiteľov normálneho rozdelenia                            | 19  |
| Gavliak Rudolf: Porovnanie modelov CAPM a APT v slovenských podmienkach  | 25  |
| Hroncová Ivana: Využitie štatistického testovania na zhodnotenie vplyvu vybraných nástrojov marketingovej komunikácie na spotrebiteľa                    | 30  |
| Hurbánková Ľubica: Zvyšovanie konkurencieschopnosti slovenského vysokého školstva  | 35  |
| Huťa Anton: Diferenčné rovnice neceločíselného rádu a ich riešenie pomocou softvéru MATHEMATICA  | 42  |
| Chajdiak Jozef: Vývoj reálnych miezd v SR  | 45  |
| Katina Stanislav: Comparison of Procrustes and Bookstein 2D coordinates in shape analysis: simulations, bagplots for landmarks and applications          | 47  |
| Kepič Tomáš, Török Csaba.: Algoritmus výberu segmentov plôch   | 52  |
| Klein Daniel: Porovnanie Slangerových odhadov variančných komponentov so štandardnými odhadmi pomocou simulácií  | 61  |
| Koróny Samuel: Opisná štatistika vybraných absolútnych ukazovateľov stavebných podnikov v SR   | 65  |
| Kotlebová Eva: Empirický bayesovský bodový odhad parametra Poissonovho rozdelenia  | 69  |
| Labudová Viera, Rybánska Zuzana: Použitie jednoduchých metód viacozmerného porovnávania pri analýze životnej úrovne krajov Slovenskej a Českej republiky | 75  |
| Ľapinová Erika: Meranie chudoby v krajinách Európskej únie a na Slovensku  | 81  |
| Linda Bohdan, Kubanová Jana: Demographical and economical indicators of the countries with low scale of economical development                           | 86  |
| Löster Tomáš: Srovnání systémů STATGRAPHICS PLUS, SAS a MS EXCEL při vyhledávání kvantilů a hodnot distribučních funkcí                                  | 90  |
| Luha Ján: Reprezentatívnosť vo výskumoch verejnej mienky   | 95  |
| Megyesiová Silvia: Softvérové riešenie úloh zhlukovej analýzy  | 100 |
| Mikuš Andrej: Porovnanie vývoja HTFK vo vzťahu ku Hrubému domácomu produktu  | 106 |
| Pecáková Iva: Testy nezávislosti v řídkých kontingenčních tabulkách  | 109 |
| Řezanková Hana: Testy pro alternativní proměnné ve statistických programových systémech  | 114 |
| Šimsová Jana: Analýza časových řad míry nezaměstnanosti dvou regionů Ústeckého kraje   | 119 |
| Stankovičová Iveta: Analýza rozptylu a systém SAS®   | 125 |
| Stehlíková Beáta: Ukážka využitia SAS-u pri hodnotení genetických zdrojov  | 130 |
| Stehlíková Beáta: Využitie atribútu farby pri hodnotení genetických zdrojov  | 133 |
| Úradníček Vladimír: Využitie predajno-kúpnej parity pri stanovení spravodlivej ceny opcie na akciu   | 136 |

## ANALÝZA ROZPTYLU A SYSTÉM SAS®

Iveta Stankovičová

### Súhrn

Analýza rozptylu (ANOVA) je štatistická metóda vhodná na vyhodnocovanie experimentálnych údajov. Procedúry ANOVA a GLM sú len dve procedúry, ktoré je možné použiť na analýzu rozptylu v systéme SAS/STAT®. Procedúra ANOVA je určená na analýzu vyvážených údajov, kým procedúra GLM je vhodná pre vyvážené aj nevyvážené dátu. Pre špeciálne prípady analýzy rozptylu sú určené procedúry MIXED, NESTED a NPAR1WAY.

### Kľúčové slová

Analýza rozptylu (ANOVA), všeobecný lineárny model (GLM), efekty, viacnásobné porovnávanie, systém SAS®.

### Summary

The analysis of variance (ANOVA) is statistical method for analyzing data from a wide variety of experimental designs. The ANOVA and GLM procedures are two of several procedures available in SAS/STAT® software for analysis of variance. The ANOVA procedure is designed to handle balanced data, whereas the GLM procedure can analyze both balanced and unbalanced data. The following procedures can be used for special analysis of variance: MIXED, NESTED and NPAR1WAY.

### Key words

Analysis of variance (ANOVA), General Linear Model (GLM), Effects, Multiple Comparisons, SAS® System.

### Úvod

Analýza rozptylu je štatistická metóda vhodná na skúmanie vzťahov medzi premennými a používa sa hlavne na vyhodnocovanie experimentálnych údajov. ANOVA je špeciálnym prípadom viacnásobnej regresie a teda aj špeciálnym prípadom všeobecného lineárneho modelu (GLM). Výpočtovo môže byť ANOVA interpretovaná ako regresia s umelými (dummy) vysvetľujúcimi premennými. Maticové riešenie umožňuje odvodiť pomerne jednoduché vzorce pre výpočet parametrov a testovanie hypotéz.

Vysvetľovanou (závislou) premennou v ANOVe je spojité kvantitatívna premenná Y a ako faktor (vysvetľujúca, nezávislá premenná) vystupuje kategoriálna premenná A s malým počtom obmien. Podľa týchto obmien je možné hodnoty vysvetľovanej premennej triediť do skupín a tak skúmať, či faktorová premenná vplýva na vysvetľovanú premennú.

Analýzu rozptylu klasifikujeme z viacerých hľadišť. Ak skúmame vplyv len jedného faktora na vysvetľovanú premennú, tak ide o jednofaktorovú ANOVu. Jedná sa teda len o jednoduché triedenie. Pri viacerých faktoroch hovoríme o viacfaktorovej ANOVe (dvojité, trojité, ... triedenie). Podľa počtu vysvetľovaných premenných rozoznávame jednorozmernú (ANOVA) a viacrozmernú (MANOVA) analýzu rozptylu. ANOVu považujeme za vyváženú (balanced), ak v skupinách podľa hodnôt faktorovej premennej je rovnaký počet štatistických jednotiek. V opačnom prípade ide o nevyvážený (unbalanced) ANOVA model. Podľa toho, či analýzu rozptylu počítame z nameraných hodnôt závislej premennej alebo len z ich poradových (rank) hodnôt, rozoznávame parametrickú a neparametrickú ANOVu. Neparametrická ANOVA je vhodná hlavne pre malé súbory, v ktorých sa nedá overiť podmienka o normalite rozdelenia dát závislej premennej v podsúboroch podľa hodnôt faktorovej premennej.

V parametrických modeloch ANOVA uvažujeme, že hodnoty vysvetľovanej premennej Y je možné vyjadriť ako súčet neznámej strednej hodnoty  $\mu$ , hodnoty  $A_i$ , ktorá vyjadruje tzv. efekt i-tej úrovne faktora a neznámej zložky  $\epsilon$ , ktorá obsahuje ostatné vplyvy:

$$Y = \mu + A_i + \varepsilon \quad i=1,2, \dots, k \quad (1)$$

Tento model je špeciálnym prípadom všeobecného lineárneho modelu (GLM) a môžeme ho vyjadriť v maticovom tvare:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon \quad (2)$$

kde,  $\mathbf{y}$  je vektor pozorovaní členený na  $k$  podvektorov, ktoré zodpovedajú úrovniam faktorovej premennej. Matica  $\mathbf{X}=[\mathbf{1}, \mathbf{X}_A]$  je tvorená stĺpcom jednotiek, ktoré patria k parametru  $\mu$  a d'alších  $k$  stĺpcov, ktoré patria k parametrom  $A_1$  až  $A_k$ , vektor  $\boldsymbol{\beta}$  je vektor neznámych parametrov a  $\varepsilon$  je vektor neznámych zložiek.

Ak chceme odhadnúť parametre tohto modelu (2), čiže efekty jednotlivých úrovní faktora, tak sa nedá priamo použiť metóda najmenších štvorcov. Matica  $\mathbf{X}$  nemá plnú hodnosť, t.j.  $h(\mathbf{X}) < k+1$  a preto je matica  $\mathbf{X}^T \mathbf{X}$  singulárna. Uvažuje sa preto naviac lineárne obmedzenie parametrov

$$\sum_{i=1}^k n_i A_i = 0 \quad (3)$$

a redukuje sa počet odhadovaných parametrov, alebo sa namiesto matice  $\mathbf{X}$  typu  $n \times (k+1)$  uvažuje s maticou kontrastov  $\mathbf{X}=[\mathbf{I}_n \ \mathbf{X}_A \mathbf{C}]$  typu  $n \times k$ , ktorá má už plnú hodnosť a tak získame rôzne typy kontrastov. Pre prvky matice  $\mathbf{C}$ , t.j. pre koeficienty kontrastov platí

$$\sum_{i=1}^k c_{il} = 0, \quad \sum_{i=1}^k c_{il}^2 > 0 \quad \text{kde } l=1,2,\dots,k-1 \quad (4)$$

Namiesto parametrov  $A_i$  potom uvažujeme parametre  $A_i^*$ , ktoré sú navzájom vo vzťahu

$$A_i = \sum_{l=1}^{k-1} c_{il} A_l^* = 0, \quad \text{kde } i=1,2,\dots,k \quad (5)$$

Volba typu kontrastov má vplyv na interpretáciu parametrov modelu, ale nie na výsledky testovania zložených hypotéz o parametroch modelu.

Ak preukážeme existenciu vplyvu faktora, tak môže nasledovať hlbšia analýza výsledkov, v ktorej zisťujeme, medzi ktorými skupinami existujú rozdiely. Okrem hypotézy  $H_0: \mu_i - \mu_j = 0$  môžeme testovať aj hypotézu o lineárnej kombinácii stredných hodnôt, t.j. o nulovom lineárnom kontraste.

$$\Psi_i = \sum_{i=1}^k c_i \mu \quad \text{kde } \sum_{i=1}^k c_i = 0 \quad (6)$$

Pri testovaní kontrastov je podstatné, či bol kontrast zvolený dopredu alebo až na základe výsledkov. Pri simultánnom teste viacerých kontrastov aj to, či sú kontrasty nezávislé. Hrozí zvýšenie chyby I. druhu, to znamená, že vo viacej ako  $100\alpha\%$  prípadov môže byť zamietnutie testovanej hypotézy dôsledkom náhodného kolísania a nie existujúceho rozdielu medzi strednými hodnotami. V teórii bolo odvodených veľa metód, ktoré kontrolujú chybu I. druhu ( $\alpha$ ) a ktoré sa označujú ako metódy viacnásobného porovnávania (multiple comparisons). Podstatou všetkých je konštrukcia  $100(1-\alpha)\%$  intervalu spoľahlivosti pre kontrast  $\Psi$ . Ak tento interval neobsahuje 0, preukázali sme nenulovosť kontrastu na hladine významnosti  $\alpha$ .

Jednotlivé metódy sa líšia jednak podmienkami použitia, jednak šírkou intervalu spoľahlivosti a sú súčasťou štatistických počítačových programov ako je aj systém SAS®.

### **Analýza rozptylu v systéme SAS®**

Systém SAS® umožňuje vykonávať analýzu rozptylu pomocou niekoľkých procedúr modulu SAS/STAT. Sú to procedúry:

- ANOVA - len pre využívané modely,
- GLM - pre využívané ale aj nevyužívané modely,
- MIXED - pre zmiešané modely,
- NESTED - pre hierarchické modely,
- NPAR1WAY - pre neparametrickú ANOVu.

Procedúra ANOVA je určená len pre vyvážené modely so špeciálnou štruktúrou a preto je rýchlejšia a vyžaduje menej pamäte ako procedúra GLM. Nie je však vhodná na jednorozmernú analýzu rozptylu experimentov v tvare latinských štvorcov, na čiastočne vyvážené neúplné blokové experimenty, na úplné hierarchické (nested) experimenty a na experimenty s takými početnosťami v skupinách, ktoré sú proporcionálne navzájom a tiež vo vzťahu k základnému súboru. Tieto výnimky majú taký dizajn, že všetky faktory sú navzájom ortogonálne. PROC ANOVA je vhodná pre experimenty, ktorých blokové diagonálne matice  $X^T X$  obsahujú prvky vo všetkých blokoch s rovnakými hodnotami a parciálne testuje túto požiadavku rovnosti stredných hodnôt v blokoch. Bohužiaľ, pre niektoré typy experimentov je tento test v PROC ANOVA chybný. Ak test zamietne hypotézu o rovnosti stredných hodnôt, tak procedúra vypíše varovanie, že analýza nie je platná a experiment nie je vyvážený a mali by sme radšej použiť procedúru GLM. Pre úplné overenie tejto podmienky je potrebná úplná matica  $X^T X$ , a preto ak nie ste si istý splnením tejto podmienky, tak by ste mali použiť PROC GLM.

Na nasledujúcom príklade demonštrujeme, že výsledky F štatistiky pomocou procedúr ANOVA a GLM môžu byť rôzne, keď dátá predstavujú nevyvážený model.

SAS kód na vytvorenie dátového súboru WORK.EXP:

```
data exp;
  input A $ B $ Y @@;
  datalines;
    A1 B1 12 A1 B1 14      A1 B2 11 A1 B2 9
    A2 B1 20 A2 B1 18      A2 B2 17 ;
  
```

Dátový súbor  
WORK.EXP:

| A  | B  | Y  |
|----|----|----|
| A1 | B1 | 12 |
| A1 | B1 | 14 |
| A1 | B2 | 11 |
| A1 | B2 | 9  |
| A2 | B1 | 20 |
| A2 | B1 | 18 |
| A2 | B2 | 17 |

SAS kód pre procedúru ANOVA:

```
PROC ANOVA DATA=WORK.EXP;
  CLASS A B;
  MODEL Y = A B;
RUN;
```

SAS výstup pre procedúru ANOVA (časť):

| Source          | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|----|----------------|-------------|---------|--------|
| Model           | 2  | 97,71          | 48,86       | Infty   | <,0001 |
| Error           | 4  | 0,00           | 0,00        |         |        |
| Corrected Total | 6  | 97,71          |             |         |        |

Výstraha v LOG okne pre nevyvážený model:

WARNING: PROC ANOVA has determined that the number of observations in each cell is not equal. PROC GLM may be more appropriate.

SAS kód pre procedúru GLM:

```
PROC GLM DATA=WORK.EXP;
  CLASS A B;
  MODEL Y = A B;
RUN;
```

SAS výstup pre procedúru GLM (časť):

| Source          | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|-----------------|----|----------------|-------------|---------|--------|
| Model           | 2  | 91,31          | 45,66       | 28,54   | 0,0043 |
| Error           | 4  | 6,40           | 1,60        |         |        |
| Corrected Total | 6  | 97,71          |             |         |        |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|--------|----|-----------|-------------|---------|--------|
| A      | 1  | 80,05     | 80,05       | 50,03   | 0,0021 |
| B      | 1  | 11,27     | 11,27       | 7,04    | 0,0568 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|--------|----|-------------|-------------|---------|--------|
| A      | 1  | 68,27       | 68,27       | 42,67   | 0,0028 |
| B      | 1  | 11,27       | 11,27       | 7,04    | 0,0568 |

## Určenie efektov v analýze rozptylu v systéme SAS®

V modeloch analýzy rozptylu môžeme určiť efekty, ktoré sú kombináciou faktorových premenných použitých na vysvetlenie variability závislej premennej v nasledovných formách:

- ako hlavné (main) efekty:

```
proc glm data=WORK.EXP;
  class A B;
```

| Dáta | Dizajn matice X |   |
|------|-----------------|---|
|      | A               | B |
|      |                 |   |

```
model Y=A B;
run;
```

| A | B | $\mu$ | A1 | A2 | B1 | B2 |
|---|---|-------|----|----|----|----|
| 1 | 1 | 1     | 1  | 0  | 1  | 0  |
| 1 | 2 | 1     | 1  | 0  | 0  | 1  |
| 2 | 1 | 1     | 0  | 1  | 1  | 0  |
| 2 | 2 | 1     | 0  | 1  | 0  | 1  |

- ako krížové (interaction) efekty:

```
proc glm data=WORK.EXP;
  class A B;
  model Y=A B A*B;
run;
```

| Dáta |   |       | Dizajn matice X |    |    |    |      |      |      |      |
|------|---|-------|-----------------|----|----|----|------|------|------|------|
| A    | B | $\mu$ | A               |    | B  |    | A*B  |      |      |      |
|      |   |       | A1              | A2 | B1 | B2 | A1B1 | A1B2 | A2B1 | A2B2 |
| 1    | 1 | 1     | 1               | 0  | 1  | 0  | 1    | 0    | 0    | 0    |
| 1    | 2 | 1     | 1               | 0  | 0  | 1  | 0    | 1    | 0    | 0    |
| 2    | 1 | 1     | 0               | 1  | 1  | 0  | 0    | 0    | 1    | 0    |
| 2    | 2 | 1     | 0               | 1  | 0  | 1  | 0    | 0    | 0    | 1    |

- ako hierarchické (nested) efekty:

```
proc glm data=WORK.EXP;
  class A B;
  model y=A B(A);
run;
```

| Dáta |   |       | Dizajn matice X |    |      |      |      |      |      |      |      |
|------|---|-------|-----------------|----|------|------|------|------|------|------|------|
| A    | B | $\mu$ | A               |    | B(A) |      |      | B1A1 | B2A1 | B1A2 | B2A2 |
|      |   |       | A1              | A2 | B1A1 | B2A1 | B1A2 |      |      |      |      |
| 1    | 1 | 1     | 1               | 0  | 1    | 0    | 0    | 0    | 0    | 0    | 0    |
| 1    | 2 | 1     | 1               | 0  | 0    | 1    | 0    | 1    | 0    | 0    | 0    |
| 2    | 1 | 1     | 0               | 1  | 0    | 0    | 0    | 1    | 0    | 0    | 0    |
| 2    | 2 | 1     | 0               | 1  | 0    | 0    | 0    | 0    | 0    | 1    | 1    |

Procedúry GLM a ANOVA umožňujú vypočítať všetky stredné hodnoty závislej premennej pre efekty, ktoré sa nachádzajú na pravej strane príkazu MODEL a to rôznymi metódami zadanými ako vol'ba v príkaze MEANS (vid' tab. 1 a 2). Je potrebné upozorniť, že procedúra ANOVA netestuje viacnásobné porovnania pre krízové efekty. Na tento účel je potrebné použiť príkaz LSMEANS v procedúre GLM.

## Viacnásobné porovnávania

V analýze rozptylu F štatistika v ANOVA tabuľke testuje globálnu hypotézu  $H_0: \mu_1=\mu_2=\dots=\mu_k$ . Netestuje však, ktoré stredné hodnoty sú navzájom rozdielne. K tomuto účelu sa používajú procedúry viacnásobného porovnávania (multiple comparisons), ktoré klasifikujeme dvoma spôsobmi:

1. podľa toho ako sa vykonávajú:
  - porovnávanie všetkých párov stredných hodnôt, t.j. aritmetických priemerov,
  - porovnávanie medzi kontrolným a všetkými ostatnými strednými hodnotami,
2. podľa toho aký silný úsudok poskytujú<sup>1</sup>:
  - individuálny: rozdiely medzi strednými hodnotami neupravené o multiplicity,
  - nehomogénnosť: stredné hodnoty sú rozdielne,
  - nerovnosť: ktoré stredné hodnoty sú rozdielne,
  - intervale: paralelné intervaly spoľahlivosti pre rozdiely stredných hodnôt.

Sila úsudku nám hovorí, ako je možné uvažovať o štruktúre stredných hodnôt, ked' F test je významný. Súvisí to s kontrolou chyby I. druhu ( $\alpha$ ) pri viacnásobnom porovnávaní. Metódy, ktoré kontrolujú len individuálne chybu I. druhu ( $\alpha$ ), t.j. kontrolujú chybu spôsobu porovnávania (CER = comparisonwise error rate), nie sú skutočné metódy viacnásobného porovnávania.

Problém nastáva pri opakovanej t-testoch. Ak predpokladáme, že máme  $k=10$  stredných hodnôt a každý t-test vykonávame na hladine významnosti  $\alpha=0,05$ , potom pri viacnásobnom porovnávaní máme až  $n_c=10(10-1)/2=45$  párov na porovnanie a každý s pravdepodobnosťou 0,05, že sa dopustíme chyby I. druhu (t.j. že nesprávne zamietneme nulovú hypotézu o zhode stredných hodnôt). Šanca, že najmenej raz sa pri týchto 45-tich pároch dopustíme chybu I. druhu je omnoho vyššia ako 0,05. Je ľažké presne vypočítať túto chybu I. druhu pri viacnásobných porovnávaniach, ale môžeme vypočítať aspoň pesimistický odhad tejto chyby za predpokladu, že porovnávania sú

<sup>1</sup> Úsudky sú zoradené od najslabšieho po najsilnejší.

nezávislé. Vypočítame hornú hranicu chyby, ktorú nazývame chybou spôsobu experimentu (EER = experimentwise error rate) podľa vzťahu  $1 - (1 - \alpha)^n$ , čiže v našom prípade pre 10 stredných hodnôt je to až  $1 - (1 - 0,05)^{10} = 0,90$ . Pri zvyšovaní počtu stredných hodnôt sa táto chyba blíži k číslu 1. V teórii štatistiky bolo odvodených veľa metód, ktoré môžu udržať túto chybu na nízkej hodnote. V nasledovných tabuľkách uvádzame prehľad metód umožňujúcich viacnásobné porovnávania v procedúre GLM v príkazoch MEANS a LSMEANS podľa sily úsudku.

Tab. 1: Porovnávanie všetkých párov stredných hodnôt

| Metóda               | Sila úsudku   | Syntax  |                       |
|----------------------|---------------|---------|-----------------------|
|                      |               | MEANS   | LSMEANS               |
| Student's <i>t</i>   | individuálny  | T       | PDIFF ADJUST=T        |
| Duncan               | individuálny  | DUNCAN  |                       |
| Student-Newman-Keuls | nehomogénnosť | SNK     |                       |
| REGWQ                | nerovnosť     | REGWQ   |                       |
| Tukey-Kramer         | intervaly     | TUKEY   | PDIFF ADJUST=TUKEY    |
| Bonferroni           | intervaly     | BON     | PDIFF ADJUST=BON      |
| Sidak                | intervaly     | SIDAK   | PDIFF ADJUST=SIDAK    |
| Scheffé              | intervaly     | SCHEFFE | PDIFF ADJUST=SCHEFFE  |
| SMM                  | intervaly     | SMM     | PDIFF ADJUST=SMM      |
| Gabriel              | intervaly     | GABRIEL |                       |
| Simulation           | intervaly     |         | PDIFF ADJUST=SIMULATE |

Tab. 2: Porovnávanie medzi kontrolným a všetkými ostatnými strednými hodnotami

| Metóda             | Sila úsudku  | Syntax  |                               |
|--------------------|--------------|---------|-------------------------------|
|                    |              | MEANS   | LSMEANS                       |
| Student's <i>t</i> | individuálny |         | PDIFF=CONTROL ADJUST=T        |
| Dunnett            | intervaly    | DUNNETT | PDIFF=CONTROL ADJUST=DUNNETT  |
| Bonferroni         | intervaly    |         | PDIFF=CONTROL ADJUST=BON      |
| Sidak              | intervaly    |         | PDIFF=CONTROL ADJUST=SIDAK    |
| Scheffé            | intervaly    |         | PDIFF=CONTROL ADJUST=SCHEFFE  |
| SMM                | intervaly    |         | PDIFF=CONTROL ADJUST=SMM      |
| Simulation         | intervaly    |         | PDIFF=CONTROL ADJUST=SIMULATE |

## Záver

Dnes si nevieme predstaviť, že by sme robili štatistické analýzy bez vhodného softvéru. Na druhej strane, treba užívateľov vystríhať od používania procedúr bez dôkladného preštudovania si dokumentácie. Často sa stáva, že užívatelia nepoužívajú vhodné procedúry, alebo nesprávne interpretujú výsledky počítačových výstupov.

## Použitá literatúra

- [1] Bohdalová M., Wimmer G., Witkovský V., Savin A: Interval spoľahlivosti pre referenčnú hodnotu v medzilaboratórnych porovnávacích štúdiách. Zborník konferencie SAS Forum 2004, Bratislava 17.5.2004, str.1-4.
- [2] Hebák Petr, Hustopecký Jiří, Jarošová Eva, Pecáková Iva: Vícerozměrné statistické metody 1. INFORMATORIUM. Praha 2004. ISBN 80-7333-025-3.
- [3] Keith E. Muller, Bethel A. Fetterman: Regression and ANOVA. An Integrated Approach Using SAS® Software. Copyright © 2002 by SAS Institute Inc., Cary, NC, USA, ISBN 1-58025-890-5.
- [4] Luha Ján a kol.: Niektoré možnosti využitia analýzy rozptylu pri skúmaní ekozložky životného prostredia. Príspevok na seminári "Ekologické problémy urbanizovaného prostredia". PFUK Bratislava 1988.
- [5] SAS OnLine Documentation: <http://support.sas.com/91doc/docMainpage.jsp>

## kontaktná adresa:

Ing. Iveta Stankovičová, PhD.  
 Katedra informačných systémov  
 Fakulta managementu UK  
 Odbojárov 10, 820 05 Bratislava

e-mail: [iveta.stankovicova@fm.uniba.sk](mailto:iveta.stankovicova@fm.uniba.sk)