

# SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS  
and DEMOGRAPHY

3/2020

ročník/volume 30

Recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov.

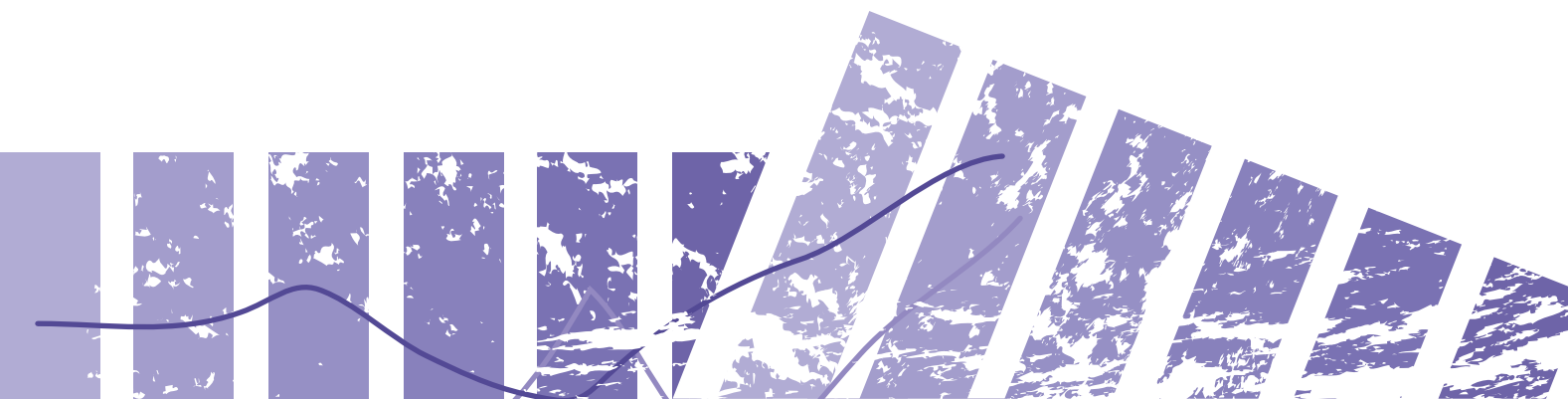
Scientific peer-reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures.

Článok/Article: 1

Typ článku/Type of article: editoriál/editorial

Strany/Pages: 3 – 8

Dátum vydania/Publication date: 15. júl 2020/July 15, 2020



## EDITORIÁL

Vážení čitatelia,



Doc. Ing. Iveta Stankovičová, PhD.

tretie číslo vedeckého časopisu *Slovenská štatistika a demografia* je tradične monotematické, snažíme sa v ňom prezentovať vybranú skupinu štatistických metód a postupov. V tomto čísle našim čitateľom predstavujeme vybrané viacrozmerné metódy a ich aplikácie. Viacrozmerné štatistické metódy sú vhodné na analýzu zložitých javov. Práve zložité javy a procesy sa v objektívnej realite vyskytujú oveľa častejšie ako javy a procesy jednoduché. Zložité javy a procesy dokážeme merať a charakterizovať iba pomocou viacerých merateľných ukazovateľov. Súbory rôznych ukazovateľov potom možno skúmať len pomocou viacrozmernej analýzy, pri ktorej sa využívajú práve tieto metódy.

Pod pojmom *metódy viacrozmernej analýzy* rozumieme široký okruh postupov, ktoré sa venujú viacrozmerným problémom z rôznych hľadísk. K najrozšírenejším metódam patrí: viacrozmerná regresná analýza, kanonická korelačná analýza, analýza hlavných komponentov, faktorová analýza, zhuková analýza, diskriminačná analýza, viacrozmerné škálovanie, viackriteriálne metódy hodnotenia ale aj iné. Jednotlivé metódy netvoria uzavreté okruhy, ale sú neustále obohacované o nové prístupy viacrozmerného riešenia a možnosti ich aplikácií v praxi.

Metódy viacrozmernej štatistickej analýzy zaznamenali prudký rozvoj najmä v posledných desaťročiach, hoci ich korene siahajú už do 2. polovice 19. storočia. Stretávame sa s nimi prakticky vo všetkých vedných odboroch, počínajúc prírodnými a technickými vedami a končiac spoločenskými vedami. Najčastejšie aplikácie týchto metód sa uskutočnili v oblasti psychológie a sociológie, medicíny, biológie, v rôznych populačných a regionálnych analýzach. Až neskôr sa viacrozmerné prístupy začali používať aj v hospodárskej praxi.

Súčasná doba je charakteristická explóziou dát, ktoré sa ukladajú v databázach. Tým rastie aj význam ich spracovania pomocou metód, ktoré dokážu z dát vyťažiť informácie. Práve informácie sú potrebné na efektívne rozhodovanie a riadenie podniku či spoločnosti. Význam viacrozmerných štatistických metód preto rastie nielen vo vede, ale aj pri riadení procesov v podnikoch a v spoločnosti. Vhodné využitie štatistických metód a správna interpretácia výsledkov získaných z analýz údajov prispievajú k takým rozhodnutiam, ktoré významne šetria čas a finančné prostriedky.

Štatistika ako veda sa rýchlo rozvíja a neustále ponúka nové metódy. Viacrozmerné štatistické metódy majú široké uplatnenie v rôznych oblastiach, avšak vyznačujú sa výpočtovou náročnosťou, čiže ich aplikácia je nemysliteľná bez využitia výpočtovej techniky. To sa na prvý pohľad javí ako prekážka ich širšieho používania. V konečnom dôsledku je to však jediný možný spôsob, ako zabezpečiť široké a systematické využitie

týchto metód v praxi. V súčasnosti je k dispozícii celý rad rôznych softvérových nástrojov, ktoré obsahujú širokú ponuku viacrozmerných metód a neustále vznikajú nové.

Štatistika ako vedná disciplína sa člení na mnoho oblastí a ako sme už uviedli, jej využitie v praxi je širokospektrálne. Aktuálne číslo časopisu *Slovenská štatistika a demografia* sa zameriava len na vybrané viacrozmerné štatistické metódy, ktoré boli aplikované pri analýzach vybraných ekonomických javov a problémov.

Prvý článok je rozsiahly a zaoberá sa v praxi často používanými viacrozmernými metódami: *analýza hlavných komponentov, faktorová analýza a zhluková analýza*. Autorky opisujú teoretické princípy a podmienky správneho použitia týchto metód. Teoretický koncept sa potom názorne aplikuje na hodnotenie a klasifikáciu 34 európskych krajín v oblasti ľudského rozvoja. Pomocou faktorovej analýzy bol vytvorený integrálny ukazovateľ ľudského rozvoja a na základe toho bolo zostavené poradie krajín. Všetky výpočty sa uskutočnili v programovom systéme SAS Enterprise Guide (verzia 7.1).

Ďalší príspevok sa venuje problematike výberu vysvetľujúcich premenných v *klasifikačných stromoch*. Týmto názvom je označovaná skupina metód, ktoré boli navrhnuté na riešenie klasifikačných úloh s vysvetľovanou premennou kategoriálneho typu na základe hodnôt vysvetľujúcich premenných. Obsah článku je ladený teoreticky, pretože obsahuje množstvo známych aj menej známych vzorcov na výpočet výberových kritérií. V závere autorka opísanú teóriu aplikovala aj na ilustračný príklad. Výpočty boli vykonané v programovom systéme IBM SPSS Decision Trees (verzia 26).

Tretí článok sa zaoberá jednoduchými metódami viacrozmerného hodnotenia (*metóda poradí, bodovacia metóda a metóda vzdialenosti od fiktívneho objektu*) a ich praktickou aplikáciou pri analýze zadlženosti domácností vo vybraných krajinách EÚ. V prvej časti autorka opisuje princíp metód a v druhej časti prezentuje aplikáciu týchto metód v priestorovej analýze dlhu domácností v krajinách, ktoré sa zúčastnili v druhej vlne prieskumu o financovaní a spotrebe domácností (HFCS). Výpočty sa takisto uskutočnili v programovom systéme SAS Enterprise Guide.

Posledný vedecký článok je venovaný *regresnej analýze* a jej aplikáciám v oblasti modelovania úmrtnosti v životnom poistení. Modely úmrtnosti sú štatistické (stochastické) modely, pre ktoré je potrebné odhadnúť parametre regresnej funkcie. Parametre vybraných regresných modelov úmrtnosti boli odhadované iteračnými metódami nelineárnej regresnej analýzy v softvéri SAS Enterprise Guide.

Všetky vedecké články pre toto číslo časopisu napísali ženy, ktoré patria medzi erudované vedecko-pedagogické pracovníčky renomovaných slovenských a českých univerzít. Vo svojej výskumnej činnosti a pedagogickom procese využívajú rôzne štatistické a analytické softvéry alebo aj programovacie jazyky, bez ktorých by bola aplikácia viacrozmerných štatistických metód na veľkých databázach údajov nepredstaviteľná a neuskutočniteľná. V predkladaných článkoch metodicky opísali a ukázali, ako treba v praxi pristupovať k hodnoteniu zložitých javov. Často sa na to využívajú viacrozmerné metódy a postupy, ktoré pri dodržaní podmienok ich použitia prinášajú nové, užitočné, logické a interpretovateľné výsledky. V prípade ich nedodržania,

výsledky môžu byť skreslené, nelogické až nesprávne. Preto je potrebné a dôležité v analytickej praxi ako prvý krok uskutočniť prieskum vstupných údajov a overiť podmienky použitia metód, ktoré plánujeme využiť v našej analýze.

Veríme, že vedecké články a ďalšie príspevky publikované v čísle 3/2020 *Slovenskej štatistiky a demografie* budú pre našich čitateľov obohacujúce a podnetné. Tým, ktorí sa zaujímajú o iné oblasti štatistiky, odporúčam do pozornosti vedecké články z niektorých minulých, ale perspektívne aj budúcich čísel nášho časopisu.

***Doc. Ing. Iveta STANKOVIČOVÁ, PhD.***

*Autorka pôsobí na Fakulte managementu Univerzity Komenského v Bratislave. Je členka redakčnej rady časopisu Slovenská štatistika a demografia a bola gestorkou prípravy monotematického čísla 3/2020.*

## EDITORIAL

Dear readers,

the third issue of the scientific Journal *Slovak Statistics and Demography* is traditionally monothematic, where we always try to present a selected group of statistical methods and procedures. In this issue, we present our readers with selected multidimensional methods and their applications. Multivariate statistical methods are suitable for the analysis of complex phenomena. It is the complex phenomena and processes that occur in objective reality much more frequently than simple phenomena and processes. We can measure and characterize complex phenomena and processes only with the help of several measurable indicators. Sets of diverse indicators can then only be examined using multivariate analysis, with the use of these methods.

By the *methods of multivariate analysis* we shall mean a wide range of procedures addressing multidimensional problems from different perspectives. The most common methods include: multidimensional regression analysis, canonical correlation analysis, principal components analysis, factor analysis, cluster analysis, discriminant analysis, multidimensional scaling, multicriteria evaluation methods, and others as well. The individual methods do not form closed circuits, but are constantly enriched with new approaches of multidimensional solutions and the possibility of their application in practice.

The methods of multivariate statistical analysis have developed rapidly, especially in recent decades, although their roots go back to the second half of the 19th century. We come across with them practically in all scientific disciplines, from natural and technical sciences to social sciences. These methods were the most frequently used in the field of psychology and sociology, medicine, biology, in various population and regional analyses. Only later the multidimensional approaches started to be used in the economic practice.

The current period is characterized by the explosion of data which are stored in databases. This also increases the importance of their processing, using methods enabling to extract information from the data. It is the information that is necessary for the effective decision-making and management of a company or a society. Therefore the importance of multidimensional statistical methods is growing not only in science, but also in process management in companies and in a society. The appropriate use of statistical methods and the correct interpretation of results obtained from data analyses contribute to a decisions that significantly save time and financial resources.

Statistics as a science is evolving rapidly and constantly offering new methods. Multivariate statistical methods have a wide range of application in various fields, but they are characterized by computational complexity, thus their application is unthinkable without the use of computational technology. At first glance, this seems to be an obstacle in their wider use. However, ultimately, this is the only possible way to ensure the wide and systematic use of these methods in practice. There are currently a number of different software tools available that include a wide range of multivariate methods and are constantly evolving.

Statistics as a scientific discipline is divided into many areas and, as we have already mentioned, it has a large-scale use in practice. The current issue of the Journal *Slovak Statistics and Demography* focuses only on selected multidimensional statistical methods that have been applied in the analysis of selected economic phenomena and problems.

The first article is extensive and focuses on the multidimensional methods frequently used in practice: *principal component analysis, factor analysis and cluster analysis*. The authors describe the theoretical principles and conditions of the correct use of these methods. The theoretical concept is then clearly applied to the evaluation and classification of 34 European countries in the field of human development. Using factor analysis, an integral indicator of human development was created, on the basis of which, the ranking of countries was compiled. All calculations are performed in the SAS Enterprise Guide software (version 7.1).

Another article deals with the issue of selecting explanatory variables in *classification trees*. This name refers to a group of methods that have been designed to solve classification problems with an explanatory variable of the categorical type, based on the values of the explanatory variables. The content of the article is theoretical, because it contains a number of known and less known formulae for calculating selection criteria. In the end, the author applied the described theory to an illustrative example. The calculations were performed on the IBM SPSS Decision Trees software (version 26).

The third article deals with simple methods of multidimensional evaluation (*order method, scoring method and method of distance from a fictitious object*) and their practical application in the analysis of household indebtedness in selected EU countries. In the first part, the author describes the principle of the methods and in the second part presents the application of these methods in the spatial analysis of household debt in countries participating in the second wave of the Survey on household finance and consumption (HFCS). The calculations are also performed in the SAS Enterprise Guide software.

The last scientific article is devoted to the *regression analysis* and its applications in the field of modelling mortality in life insurance. Mortality models are statistical (stochastic) models for which it is necessary to estimate the parameters of the regression function. The parameters of selected regression models of mortality were estimated by iterative methods of nonlinear regression analysis in the SAS Enterprise Guide software.

All scientific articles for this issue of the journal were written by women who are members of the erudite scientific and pedagogical staff of the prestigious Slovak and Czech universities. In their research activities and pedagogical process, they use various statistical and analytical software, or even programming languages, without which the application of multidimensional statistical methods on large data databases would be unimaginable and unfeasible. In the presented articles, they methodically described and showed how to evaluate the complex phenomena in practice. To this end, multidimensional methods and procedures are often used, which under the conditions of their use bring new, useful, logical and interpretable results. In the event of non-compliance, these results can be skewed, illogical or incorrect results. Therefore, it is

necessary and important in the analytical practice as a first step to conduct a survey of input data and verify the conditions of using the methods planned to use in our analysis.

We believe that the scientific articles and other contributions published in the issue 3/2020 of *Slovak Statistics and Demography* will be enriching and stimulating for our readers. I commend to those who are interested in other areas of statistics, the scientific articles from some of the past, but also future issues of our journal.

***Doc. Ing. Iveta STANKOVIČOVÁ, PhD.***

*The author works at the Faculty of Management of Comenius University in Bratislava. She is a member of the Editorial Board of the Journal Slovak Statistics and Demography and was responsible for the preparation of the monothematic issue 3/2020.*