

Štatistický úrad Slovenskej republiky
The Statistical Office of the Slovak Republic

SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS
and DEMOGRAPHY

vedecký časopis/scientific journal

2/2020
ročník 30



ŠTATISTICKÝ
ÚRAD
SLOVENSKEJ
REPUBLIKY

ISSN 1339-6854 (online)
ISSN 1210-1095 (tlačené vydanie)

SLOVENSKÁ ŠTATISTIKA A DEMOGRAFIA

Recenzovaný vedecký časopis založený v roku 1991. Od roku 2014 jednotlivé čísla časopisu zverejňujeme aj v elektronickej podobe na ssad.statistics.sk. Názory autorov článkov sa nemusia zhodovať s názormi vydavateľa.

Zahraniční poradcovia/Foreign Consultants

Gabriela Czanner

University of Liverpool
Veľká Británia/United Kingdom

Jitka Langhamrová

Vysoká škola ekonomická v Praze
University of Economics in Prague
Česká republika/Czech Republic

Estefanía Mourelle Espasandín

Universidade da Coruña
Španielsko/Spain

Michaela Potančoková

Joint Research Centre,
European Commission
Taliansko/Italy

Hana Řezanková

Vysoká škola ekonomická v Praze
University of Economics in Prague
Česká republika/Czech Republic

Milan Stehlík

Universidad Técnica Federico Santa María,
Čile/Chile
Johannes Kepler University Linz
Rakúsko/Austria

Výkonná redaktorka/Executive Editor

Silvia Hudecová

Jazykové redaktorky/Language Editors

Slovenský jazyk/Slovak Language

Silvia Duchková

Anglický jazyk/English Language

Andrea Okenková

SLOVAK STATISTICS AND DEMOGRAPHY

The scientific peer-reviewed journal founded in 1991. From 2014 individual copies of the journal are available to readers in electronic form at the website ssad.statistics.sk. The opinions of the authors do not necessarily correlate with the opinions of the publisher.

Redakčná rada/Editorial Board

Ľudmila Ivančíková

(predsedníčka/chairwoman)
Štatistický úrad SR
Statistical Office of the SR

Mikuláš Cár

Slovenská štatistická a demografická spoločnosť
Slovak Statistical and Demographic Society

Helena Glaser-Opitzová

Štatistický úrad SR
Statistical Office of the SR

Ján Haluška

INFOSTAT Bratislava

Iveta Stankovičová

Univerzita Komenského v Bratislave
Comenius University in Bratislava

Erik Šoltés

Ekonomická univerzita v Bratislave
University of Economics in Bratislava

Pavol Tišliar

Univerzita Komenského v Bratislave
Comenius University in Bratislava

Boris Vaňo

INFOSTAT - Výskumné demografické centrum
INFOSTAT - Demographic Research Centre

Obálka/Cover

Klára Smutná

Adresa redakcie/Address of Editorial Office

Slovenská štatistika a demografia
Štatistický úrad SR
Miletičova 3, 824 67 Bratislava
Slovenská republika

E-mailová adresa/E-mail address

SSaD@statistics.sk

ssad.statistics.sk
www.statistics.sk

OBSAH/CONTENTS

I. VEDECKÉ ČLÁNKY/SCIENTIFIC ARTICLES

Ondrej ORAVEC, Pavol ĎURČEK 3
VYUŽITIE HRUBÝCH MIER PRI TVORBE WEBBOVEJ TYPOLOGIE CELKOVÉHO PRÍRASTKU OBYVATEĽSTVA: APLIKÁCIA NA ÚROVEŇ OBCÍ SLOVENSKEJ REPUBLIKY V OBDOBÍ 1988 – 2016
THE USE OF GROSS RATES IN THE CREATION OF WEBB TYPOLOGY OF TOTAL INCREASE OF POPULATION: APPLICATION TO THE MUNICIPALITIES OF THE SLOVAK REPUBLIC IN THE PERIOD 1988 – 2016

Branislav ŠPROCHA 17
SKORÉ A NESKORÉ MATERSTVO NA SLOVENSKU V ČASOVOM A PRIESTOROVOM ASPEKTE
EARLY AND LATE MATERNITY IN SLOVAKIA IN TIME-SPACE PERSPECTIVE

Diana BÍLKOVÁ 35
ČTYŘPARAMETRICKÉ A TŘÍPARAMETRICKÉ LOGNORMÁLNÍ KŘIVKY JAKO MODELY MZDOVÝCH ROZDĚLENÍ
FOUR-PARAMETRIC AND THREE-PARAMETRIC LOGNORMAL CURVES AS WAGE DISTRIBUTION MODELS

II. INFORMATÍVNE ČLÁNKY, NÁZORY, RECENZIE, ROZHOVORY, INFORMÁCIE/ INFORMATIVE ARTICLES, OPINIONS, REVIEWS, INTERVIEWS, INFORMATION

Michal PÁLEŠ 57
MOŽNOSTI RIEŠENIA PROBLEMATIKY BIG DATA V JAZYKU R
ONE OF THE BIG DATA VIEWS IN R LANGUAGE
Informatívny článok/Informative article

Alžbeta RIDZOŇOVÁ 73
ASYMETRIE V ŠTATISTIKE ZAHRANIČNÉHO OBCHODU
ASYMMETRIES IN FOREIGN TRADE STATISTICS
Informatívny článok/Informative article

Eva KOTLEBOVÁ 77
Erik Šoltés
REGRESNÁ A KORELAČNÁ ANALÝZA S APLIKÁCIAMI V SOFTVÉRI SAS
Erik Šoltés
REGRESSION AND CORRELATION ANALYSIS WITH APPLICATIONS IN SAS SOFTWARE
Recenzia publikácie/Review of publication

Vladimíra LAUKO JACKOVÁ

79

Věra Kuchařová a kol.

ČESKÁ RODINA NA POČÁTKU 21. STOLETÍ

ŽIVOTNÍ PODMÍNKY, VZTAHY A POTŘEBY

Věra Kuchařová a kol.

CZECH FAMILY AT THE BEGINNING OF THE 21ST CENTURY

LIVING CONDITIONS, RELATIONSHIPS AND NEEDS

Recenzia publikácie/Review of publication

III. PRIPRAVUJEME/COMING SOON

83

Ondrej ORAVEC, Pavol ĎURČEK

Univerzita Komenského, Prírodovedecká fakulta, Katedra ekonomickej a sociálnej geografie, demografie a územného rozvoja

VYUŽITIE HRUBÝCH MIER PRI TVORBE WEBBOVEJ TYPOLÓGIE CELKOVÉHO PRÍRASTKU OBYVATEĽSTVA: APLIKÁCIA NA ÚROVEŇ OBCÍ SLOVENSKEJ REPUBLIKY V OBDOBÍ 1988 – 2016

THE USE OF GROSS RATES IN THE CREATION OF WEBB TYPOLOGY OF TOTAL INCREASE OF POPULATION: APPLICATION TO THE MUNICIPALITIES OF THE SLOVAK REPUBLIC IN THE PERIOD 1988 – 2016

ABSTRAKT

Hlavným cieľom príspevku je vypočítať hrubé miery prirodzeného prírastku a migračného salda a použiť ich pri zostrojení Webbovej typológie. Porovnáваме zmeny medzi tromi sledovanými obdobiami, konkrétne v rokoch 1988 – 1992, 2000 – 2004, 2012 – 2016 v obciach Slovenskej republiky. Lokalizujeme oblasti najväčších zmien alebo najsilnejšej rezistencie pri porovnávaní všetkých časových rezov s prihliadnutím na vývoj vybraných demografických ukazovateľov a na regionálnu diferenciáciu. Predovšetkým analýza prvého sledovaného obdobia by nebola možná bez digitalizácie analógového zdroja údajov. Tiež sme vyriešili problém rozdielneho územno-správneho členenia za jednotlivé obdobia. Poukázali sme na diferenciáciu jednotlivých oblastí a na oblasti s najväčšou rezistenciou a najväčšími premenami počas sledovaných období.

ABSTRACT

The main aim of the article is to calculate the intrinsic rates of natural increase and migration balance and use them in the construction of Webb typology. We compare the changes between the three periods, specifically in the years 1988 – 1992, 2000 – 2004, 2012 – 2016 in the municipalities of the Slovak Republic. We localize the areas of the greatest changes or the strongest resistance considering the development of selected demographic indicators and regional differentiation. In particular, the analysis of the first period would not have been possible without the digitization of the analogue data source. We also solved the problem of different territorial divisions for each period. We refer to the differentiation of different areas and the areas with the greatest resistance and the greatest changes during the reporting periods.

KLÚČOVÉ SLOVÁ

celkový prírastok obyvateľstva, pôrodnosť, migrácia, úmrtnosť, Webbova typológia

KEY WORDS

total increase of population, natality, migration, mortality, Webb's typology

1. ÚVOD

Hlavným cieľom práce je pomocou hrubých mier prirodzeného prírastku a migračného salda sledovať zmeny medzi tromi sledovanými obdobiami, konkrétne v rokoch 1988 – 1992, 2000 – 2004, 2012 – 2016 v obciach Slovenskej republiky. Hrubé miery následne použijeme pri vytvorení Webbovej typológie.

V nasledujúcom texte vysvetlíme, prečo sme dané časové rezy vybrali a pomocou odbornej literatúry vykreslíme ich špecifickú demografickú situáciu, ktorá počas ich trvania nastala.

Prvý časový rez sme (1988 – 1992) vybrali na základe toho, že sme chceli začiatok našej analýzy umiestniť do obdobia, keď bol ešte (pred transformačný) demografický režim nastavený minulým politickým zriadením. Väčšinu tohto obdobia platili pronatalitné a prorodinné opatrenia, ktoré boli zrušené až v roku 1992 [23]. Zároveň v tomto období vrcholí fáza stagnácie úmrtnostných pomerov, ktoré sa na začiatku 90. rokov 20. storočia začínajú pozvoľne zlepšovať [16]. Pred rokom 1989 prebieha vnútroštátna migrácia v kontúrach socialistickej urbanizácie. Po nežnej revolúcii tento efekt dobieha a migračné správanie začína viac dopĺňať zahraničné sťahovanie [24].

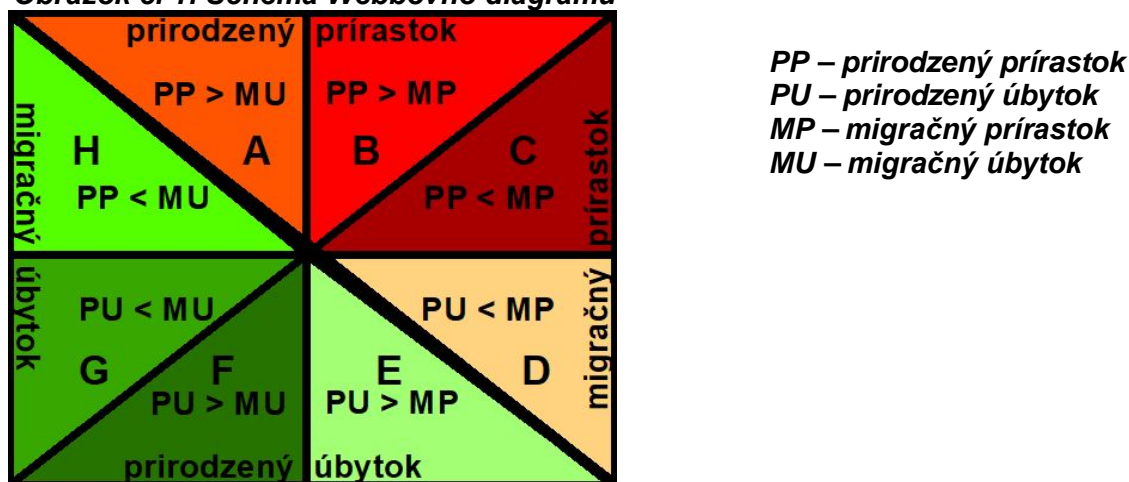
Druhý časový rez (2000 – 2004) zahŕňa práve obdobie, keď Slovenská republika zaznamenala prirodzený úbytok. V tomto období bolo dosiahnuté tzv. dno odkladania plodnosti [22]. Uzavretie manželstva a založenie rodiny už nie sú takou vysokou prioritou pre obyvateľov. Posúva sa vek uzavretia manželstva do vyššieho veku, taktiež sa zvyšuje vek, v ktorom matka porodí prvé dieťa, častejšie je spolužitie partnerov mimo manželstva, takzvaná kohabitácia. Klesá celková pôrodnosť obyvateľstva, čo má v kombinácii so zlepšujúcimi sa úmrtnostnými pomermi za následok rýchle starnutie obyvateľstva [22]. Časť týchto nových trendov je dôsledok druhého demografického prechodu [26]. Takisto sa v tomto období postupne prejavuje suburbanizácia a to najmä pri najväčších mestách [7].

Tretí časový rez (2012 – 2016) je blízky súčasnosti. Všetky spomínané javy sa berú ako bežné. Odkladanie plodnosti do vyššieho veku ústi do rekuperácie, čo spolu so stále sa zlepšujúcimi úmrtnostnými pomermi zvyšuje prirodzený prírastok [14]. Z hľadiska vnútornej migrácie možno sledovať predovšetkým dve základné tendencie. Suburbanizáciu, ktorá sa pomerne výrazne prejavuje aj v rade zázemí už stredne veľkých miest [7, 9]. Druhou tendenciou je silnejúce sťahovanie obyvateľov z východnej časti republiky na západ [8].

2. TEORETICKO-METODOLOGICKÉ UKOTVENIE PRÁCE

Naše teoreticko-metodologické ukotvenie práce vychádza z konceptu tzv. populačných syntéz [19]. To znamená, že naše hodnotenie sa vykonáva na úrovni jednotlivých demografických procesov, ale na základe vyjadrení, ktoré tieto procesy spájajú (bilancujú). Konkrétne využívame syntetizujúci efekt tzv. celkového prírastku obyvateľstva a kľúčovým prvkom našej práce je použitie Webbovej typológie, ktorú vytvoril J. W. Webb v roku 1963. Táto typológia predstavuje grafickú prezentáciu dvoch základných zložiek celkového prírastku obyvateľstva (prirodzeného prírastku a migračného salda) v dvojsovom grafe, ktorý je rozdelený na osem sektorov, tzv. oktantov. Jednotlivé sektory sú vyčlenené pomocou vzájomnej kombinácie hodnôt prirodzeného prírastku a migračného salda. Schematické znázornenie Webbovho grafu ponúkame na obrázku č. 1.

Obrázok č. 1: Schéma Webbovho diagramu



Zdroj: [20]

Oktanty Webbovho grafu sú označené písmenami A – H a môžeme ich interpretovať nasledovne. Oktanty A – D sú oktanty celkového prírastku obyvateľstva a oktanty E – H sú oktanty celkového úbytku obyvateľstva. Oktanty B a C predstavujú celkový prírastok obyvateľstva zložený z prírastkov migračných aj prírodných a oktanty F a G predstavujú celkový úbytok obyvateľstva, ktorý je kombináciou migračného a prírodného úbytku obyvateľstva [5].

Okrem zobrazenia v samotnom grafe využijeme aj mapové výstupy, v ktorých budeme prezentovať výsledky za každú obec samostatne. Kým samotný Webbov graf nám umožní určiť všeobecné tendencie pri zmene celkového prírastku obyvateľstva, tak mapy nám umožnia vykonať podrobné priestorové hodnotenie a prípadné identifikovanie oblastí, kde k zmenám dochádza rýchlejšie, a oblasti, ktoré zostávajú rezistentné.

3. DÁTA A ICH SPRACOVANIE

Štatistický úrad Slovenskej republiky poskytuje všetky dáta potrebné na vypočítanie zložiek celkového prírastku obyvateľstva od roku 1992 [4]. Údaje spred tohto roka sme museli digitalizovať z analógového zdroja [3]. Dôležitým prvkom pri digitalizácii bolo vyhnúť sa chybám pri duplicitných názvoch obcí a priradiť k obciam správne ZUJ kódy, pomocou ktorých sme následne vedeli vytvoriť mapové výstupy. Na správne zlučovanie a rozlíšenie duplicitných názvov obcí použijeme publikáciu od Majtána [15]. Do digitálnej podoby sme prepísali údaje o natalite, mortalite, migrácii a strednom stave. Výsledky tejto digitalizácie sú dostupné v práci [20].

Vzhľadom na rozdielny počet obcí Slovenskej republiky v sledovaných časových rezoch, bolo nevyhnutné vytvoriť jednotné vymedzenie obcí platné na všetky skúmané obdobia. Inak povedané, bolo potrebné vykonať na obecnej úrovni tzv. časovú harmonizáciu. Tento zámer sme vyriešili tak, že za celé sledované obdobie (1988 až 2016) sme našli rok, v ktorom bol počet obcí najnižší a podľa týchto obcí sme vytvorili mapovú vrstvu na príslušné obdobie. Najnižší počet obcí bol v roku 1988 a preto sme sa riadili týmto územno-správnym usporiadaním. Vo všetkých sledovaných obdobiach (aj 2000 – 2004, 2012 – 2016) sme zjednotili obecnú štruktúru na členenie platné z roku 1988. To sme dosiahli spájaním obcí, ktoré sme jednak vykonali v našej databáze zložiek celkového prírastku obyvateľstva

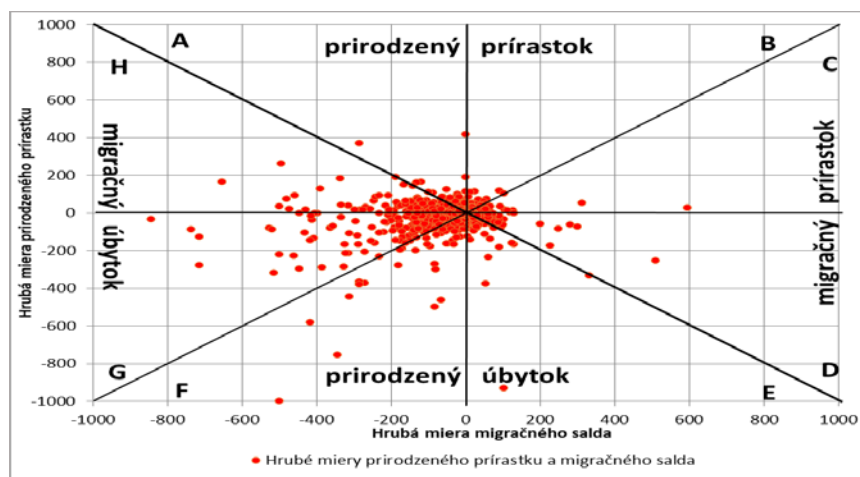
a jednak aj v mapovej vrstve. Zároveň upozorňujeme, že mestá Bratislava a Košice sme spracovali ako celok a nebudeme vykonávať analýzu za ich mestské časti.

Po časovej harmonizácii dátových sád pre jednotlivé obdobia sme pristúpili k výpočtu zložiek celkového prírastku obyvateľstva, ktoré sú potrebné na zostrojenie Webbovej typológie. Konkrétne je potrebné vypočítať hrubú mieru prirodzeného prírastku a hrubú mieru migračného salda. Hrubú mieru prirodzeného prírastku vypočítame ako rozdiel živonarodených a zomretých, ktorý vydáme stredným stavom obyvateľstva a následne vynásobíme číslom 1000. Hrubú mieru migračného salda vypočítame ako rozdiel prisťahovaných a vystahovaných, ktorý vydáme stredným stavom obyvateľstva a vynásobíme číslom 1000.

4. WEBBOVA TYPOLÓGIA V OBDOBÍ 1988 – 1992

V grafe č. 1 môžeme vidieť, že väčšina hodnôt, respektíve obcí Slovenskej republiky sa nachádza v oktantoch celkového úbytku obyvateľstva. Najpočetnejšie je zastúpený oktant H, v ktorom sa nachádza 651 obcí. Tento oktant vyjadruje prevahu migračného úbytku nad prirodzeným prírastkom. Druhým najpočetnejším oktantom je oktant G s 504 obcami. Tu migračný úbytok prevyšuje prirodzený úbytok. Ide o podobnú situáciu s rozdielom, že porovnávame dva úbytky. Môžeme tvrdiť, že oktanty F a G sú najnepriaznivejšie, keďže sa v nich porovnávajú úbytky. Oktanty E a F majú podobné zastúpenie, konkrétne 150 obcí v E a 215 obcí v F. Druhá polovica Webbovho grafu, ktorá zobrazuje celkový prírastok obyvateľstva, má približne rovnaké zastúpenie vo všetkých oktantoch. V grafe č. 1 si môžeme všimnúť, že hodnoty v týchto oktantoch, resp. v celej časti celkového prírastku obyvateľstva oscilujú pri nulovej hodnote na rozdiel od hodnôt, ktoré sú v druhej polovici – celkového úbytku obyvateľstva. Keď spojíme oktanty A a B vyjadrujúce prirodzený prírastok do jedného kvadrantu, konštatujeme, že je to po kvadrante migračného úbytku druhý najpočetnejší kvadrant. V konečnom dôsledku je to pozitívum, keďže migračný úbytok sa dopĺňa prirodzeným prírastkom. Oktant A zahŕňa 384 obcí a oktant B 286 obcí. Posledná časť Webbovho grafu je kvadrant migračného prírastku, respektíve oktanty C a D. V oktante C vidíme, že ani tu sa nevyskytujú extrémne hodnoty. Tento oktant porovnávajúci migračný a prirodzený prírastok obsahuje 269 obcí. Dve extrémne hodnoty vykazuje oktant D, ktorý komparuje migračný prírastok a prirodzený úbytok. V tomto oktante je identifikovaných 196 obcí.

Graf č. 1: Webbova klasifikácia v obciach Slovenskej republiky na základe hrubých mier za obdobie 1988 – 1992



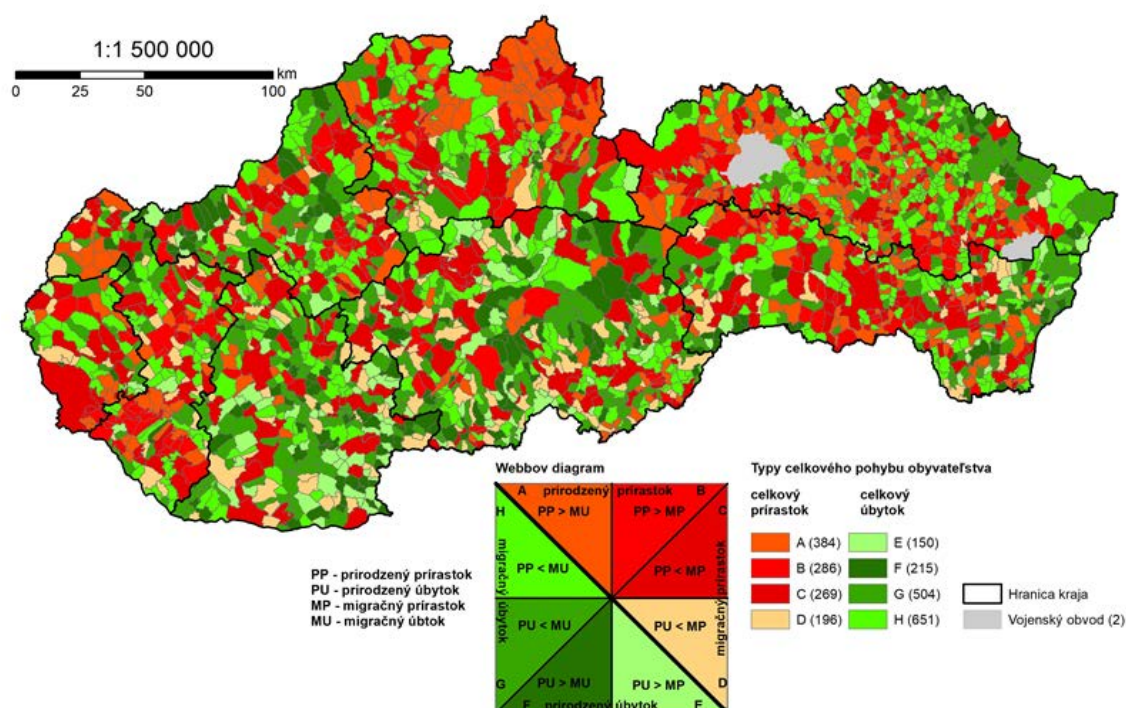
Zdroj: ŠÚ SR, Bilancia pohybu obyvateľstva 1989 – 1993; výpočty autora [20]

Tab. č. 1: Zastúpenie obcí v jednotlivých oktantoch za obdobie 1988 – 1992

Typ oktantu	Hrubé miery
A	384
B	286
C	269
D	196
E	150
F	215
G	504
H	651

Zdroj: [20]

Mapa č. 1: Webbova klasifikácia v obciach Slovenskej republiky na základe hrubých mier za obdobie 1988 – 1992



Zdroj: [20]

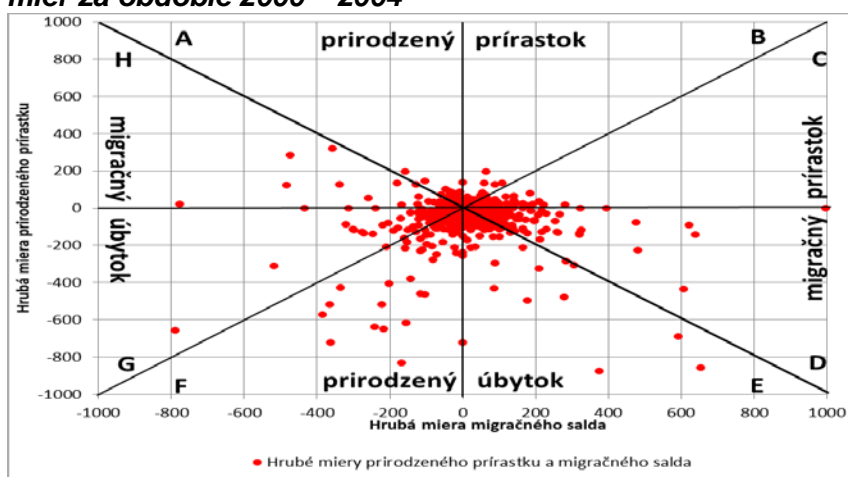
V mape č. 1 si ako prvé môžeme všimnúť kompaktný región Oravy, ktorý je typický svojou vyššou plodnosťou. Celkový prírastok obyvateľstva je viditeľný najmä na východnom Slovensku ale aj v okolí Bratislavy. Ako uvádza Korec [10], Bratislava a Košice boli najvyššími ekonomickými, hospodárskymi a administratívnymi centrami, a preto boli ziskové hlavne kvôli migrácii. Celkový prírastok obyvateľstva, respektíve prirodzený prírastok dominuje okrem Oravy aj na východnom Slovensku. Vyššia pôrodnosť je na východnom Slovensku dlhodobu typická a výrazne tomu dopomáha vysoká koncentrácia rómskeho obyvateľstva [25]. Najväčšie zastúpenie majú v okrese Kežmarok, respektíve v oblasti Spiša a Gemera. Na severovýchode Slovenskej republiky si môžeme všimnúť pás celkového úbytku obyvateľstva, v ktorom sa rovnomerne strieda migračný a prirodzený úbytok. Domnievame sa, že je to spôsobené tým, že v tomto regióne horného Zemplína, respektíve Spiša bola

(a stále je) nedostatočná infraštruktúra (okresy Svidník a Stropkov nie sú napojené na železnicu) a aj z tohto dôvodu tu absentujú pracovné príležitosti. Preto dochádza k odchodu obyvateľstva a zostávajú osoby, ktoré sú staršie (hlavne v poreprodukčnom veku) a tým pádom vo vyššom „riziku“ úmrtia. Práve v tejto oblasti sa v súčasnosti nachádzajú najmenšie obce na Slovensku, ako Príkra, Havranec či Šarbov (všetky okres Svidník). Ďalšie oblasti celkového úbytku obyvateľstva identifikujeme v oblasti Pohronia a Podunajskej roviny. V každom prípade diferenciacia obcí podľa Webbovej typológie je v tomto sledovanom období veľmi pestrá a je veľmi ťažké hľadať univerzálne vysvetľujúce modely.

5. WEBBOVA TYPOLÓGIA V OBDOBÍ 2000 – 2004

V druhom časovom reze nastávajú oproti prvému časovému rezu určité zmeny, dochádza k väčšiemu rozptylu hodnôt. Ide najmä o oktanty D, E, F, G a H. Pozorujeme vyššie zastúpenie obcí v časti celkového prírastku obyvateľstva oproti prvému sledovanému obdobiu. Zmena nastala najmä v zastúpení niektorých oktantov. Najvyššie zastúpenie má oktant D, ktorý symbolizuje prevahu migračného prírastku nad prirodzeným úbytkom. V tejto časti sa nachádza takmer tretina obcí, konkrétne 933. Druhý najpočetnejší oktant celkového prírastku obyvateľstva je oktant E, ktorý vyjadruje prevahu prirodzeného úbytku nad migračným prírastkom. Počet obcí v tomto oktante je 451. Môžeme tvrdiť, že pomerne vysoké zastúpenie v oboch týchto oktantoch nie je veľkým prekvapením, a to z dôvodu veľmi nízkej plodnosti v sledovanom období. Aprupo, ide o obdobie prirodzeného úbytku v Slovenskej republike. Posledné dva oktanty celkového prírastku obyvateľstva, teda A a B sú oktanty, ktoré vyjadrujú prirodzený prírastok a vyššie uvedený dôvod vysvetľuje ich málopočetné zastúpenie. Oktant A obsahuje 144 obcí a oktant B 257 obcí. Veľké zmeny nastávajú aj pri celkovom úbytku obyvateľstva. Zatiaľ čo v prvom sledovanom období bol najviac zastúpený oktant H a po ňom oktant G, ktoré dokopy obsahovali takmer polovicu obcí Slovenskej republiky, v druhom časovom reze sú zastúpené najmenej zo všetkých ôsmich oktantov. Obidva oktanty vyjadrujú migračný úbytok. Ich slabé zastúpenie môže byť spôsobené tým, že nastal prirodzený úbytok obyvateľstva a došlo k poklesu počtu obyvateľov, ale aj tým, že migračný úbytok vystriedal migračný prírastok. Postupne narastá efekt suburbanizácie [9]. Oktant G je zastúpený 136 obcami a oktant H 87 obcami. Ide o celkovo najnižšie zastúpenie niektorého z oktantov v našom výskume vo všetkých sledovaných obdobiach. Ako poslednú opíšeme časť prirodzeného úbytku. Napriek nízkej plodnosti, respektíve prirodzenému úbytku, ktorý nastal v troch rokoch zo skúmaných piatich v tomto časovom reze, nie sú oktanty E a F zastúpené tak početne, ako sme očakávali. Domnievame sa, že je to z dôvodu vyššieho migračného zisku, ktorý kompenzoval straty spôsobené prirodzeným úbytkom v tomto období [14].

Graf č. 2: Webbova klasifikácia v obciach Slovenskej republiky na základe hrubých mier za obdobie 2000 – 2004



Zdroj: ŠÚ SR, Bilancia pohybu obyvateľstva 2001 – 2005, výpočty autora [20]

Tab. č. 2: Zastúpenie obcí v jednotlivých oktantoch za obdobie 2000 – 2004

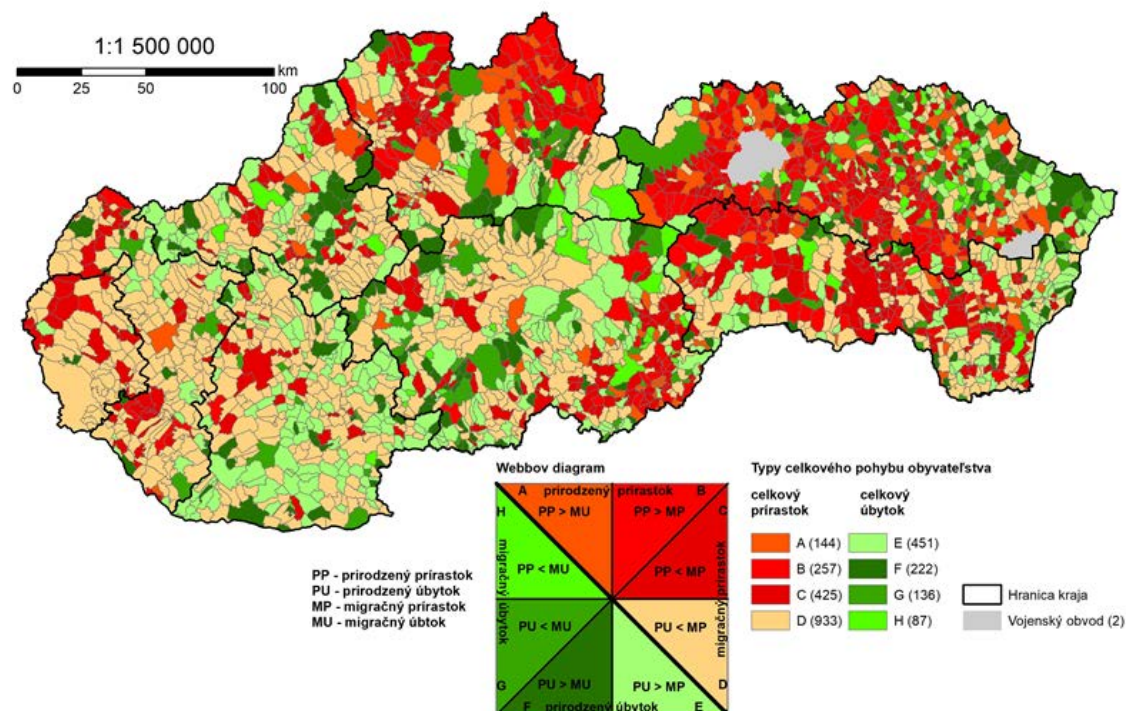
Typ oktantu	Hrubé miery
A	144
B	257
C	425
D	933
E	451
F	222
G	136
H	87

Zdroj: [20]

Najvýraznejšie zastúpenie migračného prírastku je viditeľné najmä na západnom Slovensku a v mestách Bratislava a Košice (mapa č. 2). Vzhľadom na už spomínanú polarizáciu medzi západným Slovenskom s dominanciou Bratislavy a zvyškom Slovenska to nie je prekvapujúce. Košice predstavujú regionálne centrum, ktoré v širokom okolí prevláda počtom obyvateľov, pracovným trhom a ďalšími faktormi. Napriek tomu, že v Bratislavskom kraji prevládajú dva oktanty (C a D), princíp zostáva rovnaký. Obidva oktanty totiž vyjadrujú migračný prírastok, čo vyjadruje už rozbehnutá suburbanizácia v zázemí Bratislavy a jej okolí. Ubehnutie 10 rokov od predchádzajúceho časového rezu neovplyvnilo zmenu oktantov v oblastiach, ktoré sú typické svojou vysokou plodnosťou. Ide tradične o región Oravy, kde je vysoká plodnosť ovplyvnená najmä religiozitou, a o veľkú časť východného Slovenska s dominanciou Spiša, Gemera a Šariša [18]. Príčiny pre túto oblasť sme vysvetľovali pri prvom sledovanom období a domnievame sa, že dané dôvody sa v týchto oblastiach ani po rokoch nezmenili. Celkovo si možno všimnúť, že na západnom a strednom Slovensku s výnimkou Oravy a Gemera, je veľmi nízka koncentrácia obcí s prirodzeným prírastkom. V tejto časti Slovenskej republiky prevláda migračný prírastok, respektíve prirodzený úbytok. Môžeme teda tvrdiť, že v tomto období nízkej plodnosti zabezpečovali reprodukciu obyvateľstva na Slovensku najmä obce na Orave a východnom Slovensku. Poslednou spomenutou oblasťou je oblasť Podunajskej roviny. Tak ako aj v predošlom časovom reze tu pozorujeme prirodzený úbytok. Dôvody možno hľadať v špecifickej etnicko-kultúrnej rovine. Táto oblasť je

typická obyvateľstvom maďarskej národnosti, ktoré už historicky vykazuje nižšiu plodnosť [13] a mierne zhoršené úmrtnostné pomery [17].

Mapa č. 2: Webbova klasifikácia v obciach Slovenskej republiky na základe hrubých mier za obdobie 2000 – 2004

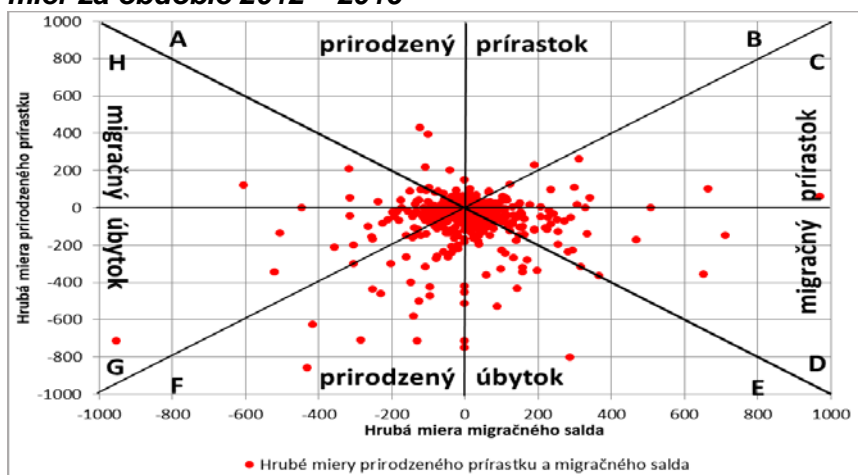


Zdroj: [20]

6. CELKOVÝ PRÍRASTOK OBYVATEĽSTVA V OBDOBÍ 2012 – 2016

Pri druhom sledovanom období sme spomínali rozdielne zastúpenie obcí v jednotlivých oktantoch medzi prvým a druhým časovým rezom. V tomto poslednom období môžeme konštatovať, že rozloženie obcí vo Webbovom grafe, respektíve v jeho oktantoch je veľmi podobné s druhým časovým rezom. To potvrdzuje, že najzávažnejšie zmeny z hľadiska zložiek celkového prírastku obyvateľstva sa udiali práve v deväťdesiatych rokoch dvadsiateho storočia a hneď na začiatku dvadsiateho prvého storočia. Zhutnene sa tieto zmeny nazývajú demografická transformácia [11]. Z Webbovho grafu č. 3 vyplýva, že najmenšie zastúpenie obcí je v časti prirodzeného prírastku. V tejto časti je aj najmenej extrémnych hodnôt. Konkrétny počet obcí v oktante A je 206, v oktante B je to 177 obcí. Pri analýze jednotlivých oktantov môžeme zhodnotiť, že najpočetnejší je oktant D, po ktorom nasleduje oktant E a C. Počet obcí v oktante D je 576. V oktante E je 425 obcí a oktant C obsahuje 422 obcí. Je potrebné podotknúť, že Slovensko zaznamenalo v tomto transformačnom období progresívny nárast strednej dĺžky života, ktorá do istej miery ovplyvňuje hodnoty prirodzeného úbytku obyvateľstva [6]. Konkrétne ide o oktanty E a F, ktoré obsahujú 425, respektíve 333 obcí. Nasleduje kvadrant migračného úbytku, a teda oktanty G a H. Počet obcí v oktante G je 300, v oktante H je to 216 obcí.

Graf č. 3: Webbova klasifikácia v obciach Slovenskej republiky na základe hrubých mier za obdobie 2012 – 2016



Zdroj: ŠÚ SR, Bilancia pohybu obyvateľstva 2013 – 2017, výpočty autora [20]

Tab. č. 3: Zastúpenie obcí v jednotlivých oktantoch za obdobie 2012 – 2016

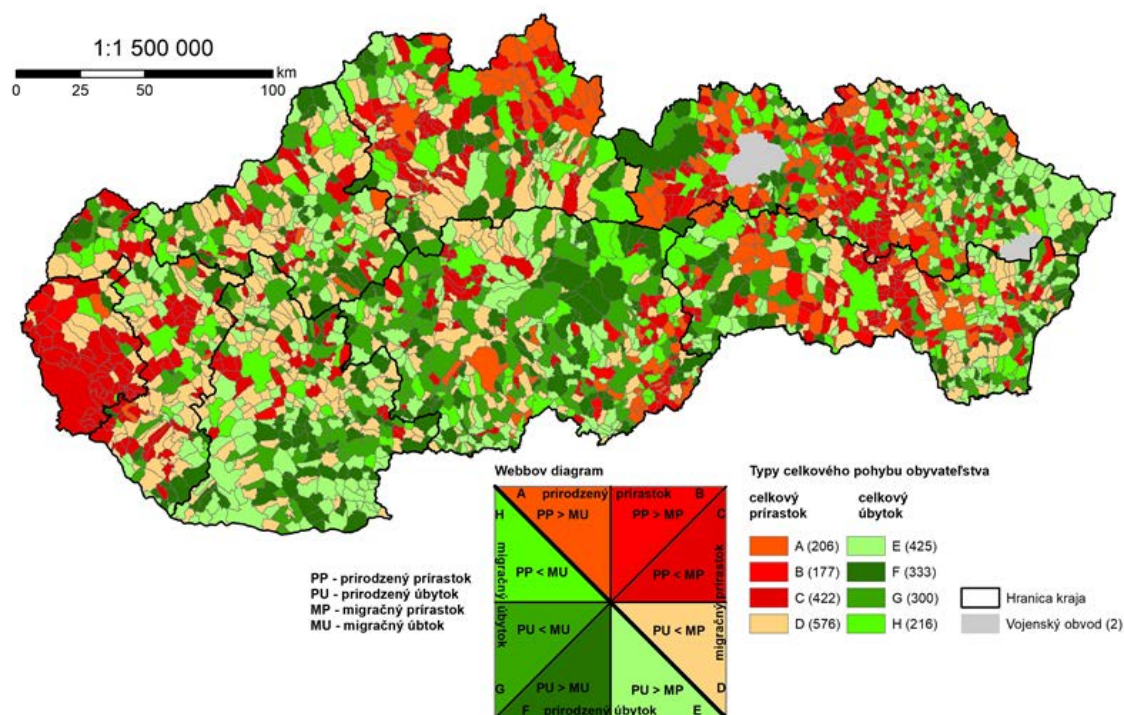
Typ oktantu	Hrubé miery
A	206
B	177
C	422
D	576
E	425
F	333
G	300
H	216

Zdroj: [20]

Môžeme tvrdiť, že mesto Košice trpí dlhšie obdobie migračnými stratami, ktoré zapríčiňuje proces suburbanizácie, obohacujúci obce v okolí tohto krajského mesta [18]. Avšak z výsledkov je zrejme, že aj v tomto období mali Košice migračné saldo kladné (mapa č. 3). Absolútnou dominantou v kontexte migračného prírastku je región Bratislavy a jeho okolie. Tu sa už naplno prejavuje jeho dominantné postavenie v rámci Slovenskej republiky. Početná migrácia do Bratislavy, respektíve jej suburbánnej zóny, či funkčného mestského regiónu pretrváva aj v tomto poslednom období. Myslíme si, že obce vykazujúce migračný úbytok, a to najmä v oblasti juhoslovenskej kotliny, ktoré spadajú do najmenej rozvinutých okresov spolu s najmenej rozvinutými okresmi nachádzajúcimi sa na východnom Slovensku, majú migračné väzby práve s funkčným mestským regiónom Bratislavy, a to najmä kvôli pracovným príležitostiam [8]. Migračne stratové sú niektoré krajské mestá, ako Prešov, Banská Bystrica, Nitra, Trnava alebo už spomínané Košice. V týchto mestách prebieha proces suburbanizácie a plodnosť v nich je na veľmi nízkej úrovni, keď nedosahuje ani záchovnú hodnotu. Už viackrát spomínaná oblasť Podunajskej roviny, začína vykazovať meniace sa tendencie, keď niekoľko obcí z tohto regiónu sa dostalo do oktantu D. Ako konštatuje Slavík a kol. [21], v západnejšej časti tohto regiónu je to dôsledok suburbanizácie. Zaujímavé výsledky možno identifikovať aj v oblasti južnej polovice stredného Slovenska. Väčšia časť regiónu vykazuje oktanty celkového úbytku obyvateľstva. Podotýkame, že v tejto oblasti, kde je zväčša hornatý terén, je celková nízka hustota zaľudnenia [12]. Výnimku tvoria nížiny a kotliny, ako

napríklad Zvolenská či Juhoslovenská kotlina. Práve tieto kotlinové oblasti majú charakter migračného prírastku okrem južnej časti Gemera, kde prevláda prirodzený prírastok. Za zmienku stojí aj situácia na slovensko-českých hraniciach resp. na strednom Považí (približne od mesta Holíč po mesto Púchov). Väčšina týchto obcí je migračne zisková. Pravdepodobnou príčinou tohto stavu je dostatok pracovných príležitostí [8]. Domnievame sa, že migračné zisky tejto oblasti prebiehajú aj na úkor stredného a východného Slovenska ale do istej miery môže táto oblasť získavať migrantov aj na úkor Českej republiky. Dôkazom týchto väzieb je aj vyšší podiel osôb českej národnosti v národnostnom zložení v týchto okresoch [2]. Pri úvahe nad dvojicou práca a migrácia dostávame odpoveď na migračné zisky obcí a ich zázemia nielen v týchto pohraničných obciach, ale aj v okolí Nitry, Trnavy či Žiliny, kde sa nachádzajú známe automobilky alebo elektrotechnické závody. Na severe Slovenskej republiky veľké zmeny oproti predošlému obdobiu nenastali. Orava ostáva kladná v kontexte prirodzeného prírastku. Menšie zmeny pozorujeme na Kysuciach. Niektoré obce získavajú charakter migračne kladnej. Domnievame sa, že to môže byť dôsledok amenitnej migrácie [1]. Na hranici Šariša a horného Zemplína môžeme vidieť rôznorodé zastúpenie jednotlivých oktantov. Domnievame sa, že táto výrazná diferenciacia je spôsobená tým, že počet obyvateľov je tu nižší a preto dokáže výsledky ovplyvniť aj efekt náhody. V každom prípade ostávajú pomery v oblasti Spiša nezmenené a dominuje tu prirodzený prírastok. Niektoré obce na severovýchode Slovenskej republiky nám vykazujú hodnoty migračného zisku, a tak môžeme predpokladať, že dôjde k zastaveniu vymierania obcí a nastane plnohodnotná reprodukcia.

Mapa č. 3: Webbova klasifikácia v obciach Slovenskej republiky na základe hrubých mier za obdobie 2012 – 2016



Zdroj: [20]

Priestorová analýza preukázala, že úroveň obcí vykazuje značný nesúlad s vývojom na úrovni štátu. Napríklad v prvom sledovanom období (1988 – 1992) sa väčšina obcí nachádza v oktantoch celkového úbytku obyvateľstva. Pritom ide o obdobie, v ktorom celkový prírastok obyvateľstva na úrovni Slovenskej republiky dosahoval najvyššie hodnoty. Ďalej v druhom období (2000 – 2004) pozorujeme vyššie zastúpenie obcí v časti celkového prírastku. Avšak celkový prírastok na celoštátnej úrovni dosahoval historicky (okrem obdobia prvej svetovej vojny) najnižšie hodnoty. Dôvody tohto efektu sú v princípe dva. V prvom rade predstavuje vysvetlenie veľkostná štruktúra obcí. Na Slovensku existuje až 1884 obcí menších ako 1000 obyvateľov (64 %), v ktorých žije 840 390 obyvateľov (15 %) [4]. Ak v týchto obciach nastane rozdielna situácia oproti republikovému priemeru, tak pri zobrazení vo Webbovom grafe to vyvoláva optický nesúlad. Druhú príčinu predstavuje jedna zo zložiek celkového prírastku – migračné saldo. Na úrovni republiky je migračné saldo tvorené len zahraničnou migráciou, ktorá je v porovnaní s vnútroštátnou migráciou výrazne menšia. Vďaka početnej vnútroštátnej migrácii dosahujú mnohé obce násobne rozdielne hodnoty hrubej miery migračného salda ako celá Slovenská republika.

7. ZÁVER

Cieľom práce bolo pomocou hrubých mier prirodzeného prírastku a migračného salda porovnať a sledovať zmeny v obciach Slovenskej republiky vo vybraných časových rezoch. Určite môžeme potvrdiť osobitosť nami skúmaných procesov, ktoré sa za sledované obdobie formovali až do dnešnej podoby a vplyvom rôznych faktorov sa budú formovať aj naďalej. Najviac rezistentné oblasti v kontexte prirodzeného prírastku sme identifikovali v oblastiach Spiša, kde je vyššia početnosť rómskeho obyvateľstva, a v regióne Oravy. Na druhej strane oblasť celkového úbytku obyvateľstva pozorujeme na severovýchode Slovenskej republiky a v regióne Podunajska. Najmä v poslednom sledovanom období si všimame kombináciu prirodzeného a migračného prírastku v zázemí Bratislavy. Predpokladáme, že proces suburbanizácie nie je ukončený a v zázemí bude stúpať celkový prírastok obyvateľstva.

Naša práca môže poslúžiť jednotlivým samosprávam ako ukážka ich vývoja za nami sledované obdobie. Vďaka digitalizácii dát z obdobia socializmu je možno prácu využiť aj na tvorbu mapových výstupov v rôznych atlasoch. Vybrané oblasti by sa dali skúmať aj detailnejšie a s pridaním ďalších procesov, čomu by sme sa mohli venovať v ďalšom výskume. Takisto by sme mohli skúmať staršie obdobia pridaním nových časových rezov, čím by sme dostali širší pohľad na vývoj transformácie obcí Slovenskej republiky.

Podarilo sa nám vyriešiť problém s rozdielnym územno-správnym členením obcí počas troch časových rezov a vytvorili sme jednotný počet obcí aj novú mapovú vrstvu prislúchajúcu k novému usporiadaniu. Taktiež sa nám podarilo digitalizovať značnú časť údajov z prvého skúmaného obdobia.

LITERATÚRA

[1] BARTOŠ, M. – KUŠOVÁ, D.: Amenitní migrace jako specifická forma globální migrace obyvatel a její vliv na kvalitu života. In: Životné prostredie, 2005, č. 6, s. 315 – 318.

- [2] BENŽA, M. a i.: Atlas obyvateľstva Slovenska. Bratislava: Univerzita Komenského, 2006. 170 s.
- [3] Bilancia obyvateľstva obcí v roku 1988/1989/1990/1991. Bratislava: ŠÚ SR (analogový zdroj).
- [4] Bilancia pohybu obyvateľstva podľa obcí v období 1993 – 2016. Bratislava: ŠÚ SR, 2018. [online]. [cit. 1. 8. 2019]. Dostupné na: <http://datacube.statistics.sk/>
- [5] BLEHA, B. – NOVÁKOVÁ, G.: Praktikum z demogeografie a demografie 2. Geografika. Bratislava, 2010. 82 s. ISBN 9978-80-89317-14-1.
- [6] BLEHA, B. – ŠPROCHA, B. – VAŇO, B.: Prognóza obyvateľstva Slovenska do roku 2060: Revízia poznatkov a predpokladov v kontexte pokračujúcej transformácie. 1. vyd. Bratislava: INFOSAT, 2018. 78.s. ISBN 97880-89398-37-9.
- [7] GAZDA, M. – NOVOTNÝ, L.: Vývoj základných komponentov pohybu obyvateľstva na Slovensku vo vzťahu k vybraným ekonomickým ukazovateľom. In: Geographia Cassoviensis, 2014, č. 1, s. 5 – 17.
- [8] JURČOVÁ, D.: Migračné toky v Slovenskej republike. Bratislava: Inštitút informatiky a štatistiky, 2010. 129 s.
- [9] JURÍK, J.: Vplyvy migrácie na demografický vývoj Slovenska. In: Inštitút pre paradigmatické reformy, 2012. [online]. [cit. 1. 8. 2019]. Dostupné na: <https://ippr.sk/s/159-vplyvy-migracie-na-demograficky-vyvoj-slovenska>
- [10] KOREC, P. – BYSTRICKÁ, S.: Dynamika hierarchickej organizácie regionálnej štruktúry Slovenska. In: Geographia Cassoviensis, 2008, č. 1, s. 74 – 85.
- [11] LANGHAMROVÁ, J. – VAŇO, B. (ed.): 20 rokov samostatnosti z pohľadu demografie ČR, SR, ČSR. Bratislava: Výskumné demografické centrum INFOSAT, Vysoká škola ekonomická Praha, 2014. 92 s. ISBN 978-80-89398-25-6.
- [12] LAUKO, V.: Geografia Slovenskej republiky: humánna geografia. Bratislava: Geo-grafika, 2013. 289 s. ISBN 978-80-89317-23-3.
- [13] MAJO, J.: Niekoľko poznámok k fenoménu etnicity v súčasnej slovenskej humánnej geografii. In: Acta Geographica Universitatis Comenianae, 2014, č. 2, s. 149 – 172.
- [14] MAJO, J. – ŠPROCHA, B.: Storočie populačného vývoja Slovenska I.: demografické procesy. Bratislava: INFOSAT, Univerzita Komenského v Bratislave, 2016. 185 s. ISBN 978-80-89398-30-0.
- [15] MAJTÁN, M.: Názvy obcí Slovenskej republiky: vývin v rokoch 1773 – 1997. Veda, 1998. 600 s. ISBN: 978-80-224-0530-2.
- [16] MESLÉ, F.: Mortality in Central and Eastern Europe: Long-term trends and recent upturns. In: Demographic Research, Special Collection 2, 2004, s. 45 – 70.
- [17] MÉSZÁROS, J.: Atlas úmrtnosti Slovenska 1993 – 2007. Bratislava: INFOSAT, 2008. 108 s. ISBN 978-80-89398-03-4.
- [18] MLÁDEK, J.: Priestorová diferencovanosť populačného vývoja Slovenska. In: Ekonomická teória a prax – dnes a zajtra. Banská Bystrica: Ekonomická fakulta Univerzity M. Bela, 2018, s. 41 – 45.
- [19] MLÁDEK, J.: Demografické syntézy. In: Mládek – Kusendová, D. (ed.): Demografická analýza Slovenska Bratislava. Univerzita Komenského v Bratislave, 2006, s. 137 – 151.
- [20] ORAVEC, O.: Demografická typizácia obcí Slovenskej republiky: Komparácia troch časových rezov [bakalárska práca]. Bratislava: Univerzita Komenského, Prírodovedecká fakulta, 2018. 59 s.
- [21] SLAVÍK, V. – GRÁC, R. – KLOBUČNÍK, M. – KOHÚTOVÁ, K.: Development of Suburbanizations of Slovakia on the Example of the Bratislava Region. In: Marszal,

T. (ed.): Urban Regions as engines of Development. Warsaw: Polish Academy of Science, 2011, Committee for Spatial Economy and regional Planning, s. 35 – 58.

[22] ŠPROCHA, B. – ŠTASTNÁ, A. – ŠÍDLO, L.: Bezdetnosť – nový fenomén v životných dráhach žien na Slovensku. In: Sociológia, 2017, č. 5, s. 561 – 587.

[23] ŠPROCHA, B. – TIŠLIAR, P.: 100 rokov obyvateľstva Slovenska: od vzniku Československa po súčasnosť. Bratislava: Muzeológia a kultúrne dedičstvo, 2018. 445 s. ISBN 978-80-89881-14-7.

[24] ŠTEFANČÍK, R.: Migrácia v Európe v II. polovici 20. storočia. Migračná politika národných štátov, 2011. 33 s.

[25] ŠUVADA, M. – SLAVÍK, V.: Prínos slovenskej geografie do výskumu rómskej minority v Slovenskej republike. In: Acta Geographica Universitatis Comenianae, 2016, č. 2, s. 207 – 236.

[26] VAN DE KAA, D. J.: Europe's second demographic transition. In: Population bulletin, 1987, č. 1, s. 1 – 59.

Príspevok je súčasťou riešenia grantového projektu APVV VV-17-0079 s názvom „Analýza a prognóza demografického vývoja Slovenskej republiky v horizonte 2080: identifikácia a modelovanie dopadov na sociálno-ekonomickú sféru v rozličných priestorových mierkach“.

RESUMÉ

Vo výsledkoch práce sme zistili rozmanitosť obcí v Slovenskej republike z hľadiska celkového prírastku obyvateľstva. Využili sme Webbovu typológiu, ktorá nám rozmanitosť potvrdila. Táto typológia je grafickým znázornením dvoch základných zložiek celkového prírastku obyvateľstva (prirodzeného a migračného). Mapy sme vytvorili pomocou hrubých mier a Webbovej typológie. Od roku 1988 sme digitalizovali veľkú časť údajov z konca Československej federatívnej republiky a vytvorili sme jednu mapovú vrstvu pre rovnaký počet obcí od roku 1988 pre všetky časové obdobia. Jednoznačne môžeme potvrdiť migračnú dominanciu Bratislavy z dôvodu pracovných príležitostí a migračných strát východného Slovenska, ktoré sú nahradené nadpriemerným prirodzeným rastom. Našli sme výrazné oblasti zmien, ale aj oblasti rezistencie, ako sú regióny Orava a Spiš. Potvrdili sme vplyv suburbanizácie v okolí hlavného mesta a zvýšenie celkového prírastku obyvateľstva v okolitých obciach.

RESUME

In the outcomes of the work we found a diversity of municipalities in the Slovak Republic from the perspective of the overall population increase. We used the Webb's typology which confirmed the diversity. This typology is a graphical representation of the two basic components of the total increase (natural and migratory). We created maps using intrinsic rates and Webb typology. We digitized a considerable part of data from the end of the Czechoslovak Federal Republic and created a single map layer for the same number of municipalities from 1988 for all time periods. We can clearly confirm the migration dominance of Bratislava due to job opportunities and migration losses of Eastern Slovakia, which are replaced by the above-average natural growth. We found highlighted areas of change, but also areas of resistance such as the Orava and Spiš regions. We confirmed the impact of suburbanization around the capital city and the increasing total population increase in the surrounding municipalities.

PROFESIJNÝ ŽIVOTOPIS

Mgr. Pavol Ďurček, PhD., vyštudoval odbor geografia a demogeografia v štátnej správe a samospráve na Prírodovedeckej fakulte Univerzity Komenského v Bratislave. Medzi rokmi 2011 až 2015 absolvoval doktorandské štúdium na katedre humánnej geografie a demografie. V súčasnosti je vedeckým pracovníkom na katedre humánnej geografie a demografie. Venuje sa výskumu geografickej, ako aj štatistickej diferenciácie demografických procesov a demografických štruktúr.

Bc. Ondrej Oravec je študentom Univerzity Komenského v Bratislave na Prírodovedeckej fakulte v odbore geografia a demografia v štátnej správe a samospráve.

KONTAKT

durcek@fns.uniba.sk

ondrej.oravec2@gmail.com

Branislav ŠPROCHA
INFOSTAT – Výskumné demografické centrum
Centrum spoločenských a psychologických vied SAV

SKORÉ A NESKORÉ MATERSTVO NA SLOVENSKU V ČASOVOM A PRIESTOROVOM ASPEKTE¹

EARLY AND LATE MATERNITY IN SLOVAKIA IN TIME-SPACE PERSPECTIVE

ABSTRAKT

Skoré materstvo na Slovensku bolo historicky dôležitým charakteristickým prvkom reprodukčného správania. Od začiatku 90. rokov minulého storočia sme však svedkami výrazného a dynamicky prebiehajúceho odklonu od tohto modelu a naopak dochádza, k procesu odkladania materského štartu a tým aj celého prechodu k materstvu a rodičovstvu. Hlavným cieľom príspevku je analyzovať skoré a neskoré materstvo na Slovensku v dlhšom časovom aspekte, poukázať na hlavné zmeny a vývojové trendy a súčasne identifikovať prípadné regionálne rozdiely existujúce v prvej polovici 90. rokov a v poslednom období.

ABSTRACT

Early motherhood was a historically important characteristic feature of reproductive behavior in Slovakia. Since the early 1990s, we have witnessed, however, a significant and dynamic deviation from this model and conversely, we are experiencing a process of postponement of maternal starts and thus of the whole transition to motherhood and parenthood. The main aim of the paper is to analyze early and late maternity in Slovakia in the longer term, to point out the main changes and development trends and at the same time to identify the possible regional differences existing in the first half of the 1990s and in the recent period.

KLÚČOVÉ SLOVÁ

skoré a neskoré materstvo, odkladanie plodnosti, vývoj, regionálne rozdiely, Slovensko

KEY WORDS

early and late maternity, fertility postponement, development, regional differences, Slovakia

1. ÚVOD

Populácia Slovenska sa dlhodobo vyznačovala úzkym previazaním života v manželstve a realizáciou reprodukčných zámerov. Len malá časť detí sa rodila mimo manželský zväzok [11]. Súčasne bol charakteristickou črtou v zmysle Hajnalovho rozdelenia [3] tzv. neeurópsky model sobášneho správania, ktorý sa vyznačoval nízkym vekom žien pri vstupe do manželstva a vysokou intenzitou sobášnosti [3, 12]. Skoré časovanie prechodov do manželstva tak vytváralo predpoklad na skorý začiatok reprodukcie. Špecifické podmienky, ktoré vznikli po druhej svetovej vojne a najmä v období budovania reálneho socializmu, prispeli

¹ Príspevok je výsledkom projektu VEGA č. 2/0064/20 „Pokračujúca transformácia rodinného a reprodukčného správania na Slovensku v časovom a priestorovom aspekte“ a projektu APVV-017-0079 „Analýza a prognóza demografického vývoja Slovenskej republiky v horizonte 2080: identifikácia a modelovanie dopadov na sociálno-ekonomickú sféru v rozličných priestorových mierkach“.

k upevneniu tohto modelu. Sme svedkami povojnového poklesu už tak nízkeho priemerného veku pri prvom sobáši a následne aj priemerného veku pri narodení prvého dieťaťa [12]. Model skorej sobášnosti, plodnosti orientovanej výhradne do prvej polovice reprodukčného obdobia so skorým ukončením reprodukčnej dráhy pri medzigeneračne čoraz viac sa prehlbujúcej orientácii na dve deti môžeme identifikovať v podstate až do začiatku 90. rokov. K stabilite tohto modelu prispieval celý komplex vonkajších faktorov. Jednak to bolo skoré ukončovanie vzdelanostnej dráhy, plná zamestnanosť mladých, mzdová nivelizácia, systém populačnej politiky zvýhodňujúci mladé manželské páry s deťmi, prednostné pridelovanie bytov, ako aj obmedzená sexuálna výchova a možnosti spoľahlivého bránenia počatiu [7, 11]. Tento systém nevytváral tlak na odkladanie materského štartu a v mnohých smeroch ho v podstate, naopak, podporoval. Mladí ľudia tak necítili potrebu odkladať manželstvo a materstvo [7], pričom takéto reprodukčné správanie by ich navyše v mnohých smeroch penalizovalo. Výhodnosť skorého začiatku reprodukčnej dráhy sa postupne dostala aj do spoločenského diskurzu a premietol sa v normách časovania prechodov do manželstva a rodičovstva. Nastavenie rodinných opatrení v kombinácii s potrebou dvojpríjmovej rodiny a tak vznikajúceho duálneho tlaku na ženy prispieval k snahám v pomerne mladom veku realizovať všetky reprodukčné zámery. Okrem toho v povojnovom období je potrebné ešte pripomenúť dokončovanie procesu demografickej revolúcie a s ňou spojeného vedomého obmedzovania plodnosti a veľkosti rodín. Výsledkom tak bola postupná orientácia na dvojdetný (prípadne maximálne trojdetný) model rodiny a výrazná koncentrácia plodnosti do veku 30 rokov [11].

Obdobie po roku 1989 a s ním prichádzajúca hospodárska, celospoločenská, politická a kultúrna transformácia priniesli celkovú diskontinuitu životných podmienok mladým mužom a ženám a jedným z hlavných aspektov ich rodinného a reprodukčného správania sa stáva odkladanie prechodov do manželstva a rodičovstva [6, 7]. To je súčasťou celkovej komplexnej premeny časovania a nastavenia prechodov k dospelosti. Dochádza k uvoľneniu dovtedy tesného prepojenia medzi manželstvom a materstvom a k časovej heterogenizácii plodnosti. Jedným z hlavných znakov reprodukcie sa stáva jej čoraz väčšia koncentrácia do druhej polovice reprodukčného veku. Odkladanie plodnosti je typickou črtou reprodukčného správania mladých generácií žien, pričom sa to týka nielen celkovej plodnosti, ale postupne v čoraz väčšej miere aj rodenia prvého dieťaťa [8]. Viacerí autori [1, 4] uvádzajú, že súčasný proces odkladania predstavuje tak špecifickú a v mnohých aspektoch historicky jedinečnú kapitolu v premene ľudskej reprodukcie, že ju identifikujú ako samostatnú tranzíciu plodnosti (postponement transition). Na druhej strane však viaceré domáce analýzy [9, 10] poukazujú na skutočnosť, že v populácii Slovenska stále zostáva určitá podskupina žien, ktoré sa aj napriek týmto transformačným zmenám stávajú matkami v mladom až veľmi mladom veku.

Hlavnou snahou príspevku bude analyzovať skoré aj neskoré materstvo na Slovensku, ich miesto v reprodukčnom správaní v dlhšom časovom aspekte, poukázať na hlavné zmeny a vývojové trendy, ktorými prešli zhruba od konca prvej svetovej vojny, a súčasne identifikovať prípadné regionálne rozdiely, ktoré sa vytvorili na začiatku 90. rokov a v poslednom období.

2. KONCEPT SKORÉHO A NESKORÉHO MATERSTVA A ICH VÝVOJ NA SLOVENSKU

Nikdy v histórii sa ženy v Európe nestávali matkami tak neskoro, ako je to od začiatku 21. storočia [5]. Kohler a kol. [4] dopĺňajú, že súčasné trendy vo vývoji plodnosti sú v postmoderných spoločnostiach charakterizované rastúcim trvalým, nereverzibilným odkladaním plodnosti naprieč širokou škálou socioekonomických podmienok. Proces odkladania je analyzovaný v desiatkach špecializovaných štúdií [napr. 1, 2, 5, 6, 8], ktoré sa venujú rôznym aspektom (napr. historickým trendom, teoretickým ukotvením neskorého a veľmi neskorého materstva, ich dôsledkami a pod.) Z hľadiska vekového vymedzenia Toulemon [13] definuje neskorú plodnosť ako plodnosť žien vo veku 35 a viac rokov. Mierne odlišné vymedzenie poskytuje Billari a kol. [2], ktorí uvádzajú, že ako neskorú plodnosť (*late-fertility*) je možné označiť plodnosť žien vo veku 40 a viac rokov, pričom plodnosť žien vo veku 45 a viac rokov definujú ako veľmi neskorú plodnosť (*very late fertility*). S tým je spojené označenie veku žien 40 a viac rokov ako pokročilý reprodukčný vek (*advanced childbearing age*) a 45 a viac rokov ako veľmi pokročilý reprodukčný vek (*very advanced childbearing age*). Keďže na Slovensku v minulom režime bola plodnosť vo veku nad 35 rokov veľmi nízka a celkovo reprodukcia bola výrazne koncentrovaná do prvej polovice reprodukčného veku, za neskoré materstvo budeme považovať už plodnosť vo veku 35 a viac rokov, pričom sa zvlášť pozrieme aj na úroveň a zmeny v intenzite plodnosti v pokročilom a veľmi pokročilom reprodukčnom veku.

V prípade mladého materstva sa vymedzenie opiera o vek 20, resp. 25 rokov. Keďže ešte donedávna bola reprodukcia vo veku 20 – 24 rokov na Slovensku kľúčová pre celkovú plodnosť žien, budeme mladé materstvo ohraničovať vekom 20 rokov.

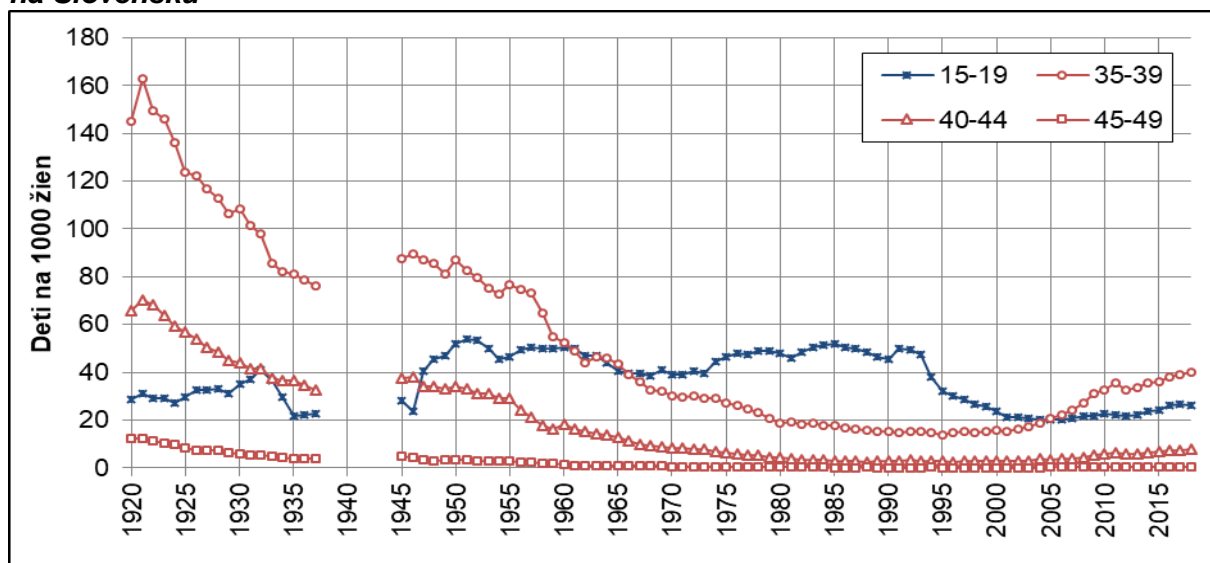
Po prvej svetovej vojne sa Slovensko z pohľadu transformácie reprodukčného správania v rámci demografickej revolúcie nachádzalo stále ešte len na polceste k modernému charakteru reprodukcie. Navyše povojnová kompenzačná fáza a s ňou spojený prudký nárast plodnosti dočasne zastavili kontinuálny pokles intenzity rodenia detí. Aj preto na začiatku 20. rokov minulého storočia vidíme pomerne výrazné zvýšenie plodnosti vo veku 35 – 39 rokov. V starších vekoch bol tento jav už menej výrazný a u najmladších žien vo veku do 20 rokov sa v podstate neprejavil vôbec. Po tom, ako odzneli priaznivé reprodukčné podmienky vytvorené v krátkom kompenzačnom období, nastal prudký pokles plodnosti, ktorý sa opätovne najviac týkal žien vo veku 35 – 39 rokov. Klesala však aj intenzita rodenia detí vo veku 40 a viac rokov. Tieto zmeny boli v prevažnej miere odzrkadlením ďalšieho šírenia a prehľbovania vedomého obmedzovania plodnosti a veľkosti rodiny v rámci demografickej revolúcie. Ako ukázali vo svojich prácach Šprocha – Tišliar [11, 12], práve medzivojnové obdobie predstavovalo kľúčový míľnik na ceste k modernej forme reprodukčného správania populácie Slovenska, ktorý sa dotýkal najmä redukcie detí vyššieho poradia a teda aj starších vekov reprodukčného obdobia. Na druhej strane v najmladšom veku identifikujeme spočiatku stagnáciu, ktorú vystriedal mierny rast plodnosti v druhej polovici 20. rokov. Ten pravdepodobne zastavili nepriaznivé životné podmienky, ktoré vytvorila hospodárska kríza v 30. rokoch. V staršom veku sa pokles plodnosti ešte viac prehĺbil a najmä vo veku 35 – 39 rokov sa pomerne výrazne medziročne urýchlil. V najmladšom veku bol spomínaný pokles tiež pomerne značný, a preto už v polovici 30. rokov dosahovala plodnosť žien vo veku do 20 rokov najnižšiu úroveň v medzivojnovom období.

Z hľadiska príspevkov jednotlivých sledovaných vekových skupín bol vývoj v podstate identický s celkovým vývojom intenzity plodnosti. Po prvej svetovej vojne klesala váha plodnosti žien vo veku 35 – 39 rokov z takmer jednej pätiny pod hranicu 15 %. Vo veku nad 40 rokov sa v povojnovom období celkovo realizovala ešte takmer desatina plodnosti, ale v druhej polovici 30. rokov to už nebolo ani 7 %. Hlavnou zložkou pritom však bola plodnosť žien vo veku 40 – 44 rokov, keďže vo veku nad 45 rokov len v prvej dekáde povojnového obdobia prekračovala hranicu 1 %. Na druhej strane vekového spektra sa príspevky najmladších žien k celkovej plodnosti postupne zvýšili z 3 % až nad 6 %, no spomínaný pokles po roku 1933 spôsobil prudké zníženie na hranicu približne 4 %. Uvedené zmeny spôsobili, že kým na začiatku 20. rokov sledované vekové skupiny sa spoločne podieľali na približne 30 % plodnosti, na konci existencie medzivojnového Československa to už nebola ani jedna štvrtina.

Pre obdobie rokov 1938 – 1944 nedisponujeme potrebnými údajmi na konštrukciu vekovo špecifických mier plodnosti, preto miery, ako aj príspevky analyzovaných vekových skupín k celkovej reprodukcií žien na Slovensku môžeme sledovať opätovne až od roku 1945. V porovnaní s obdobím pred koncom existencie medzivojnového Československa je zrejmé, že intenzita rodenia detí vzrástla viac-menej vo všetkých analyzovaných vekových skupinách. Najmenšie zmeny vidíme v najstaršom a najmladšom veku. Celkovo sa sledované veky v roku 1945 podieľali na približne 26 % z celej plodnosti. Ďalší vývoj sa niesol najprv v pozvoľnom poklese intenzity rodenia detí vo veku 35 a viac rokov, ktorý sa v druhej polovici 50. rokov zrýchlil. V podstate kontinuálne prebiehal až do polovice 90. rokov, keď intenzita plodnosti v týchto vekových skupinách dosiahla svoje minimálne hodnoty. Z grafu č. 1 je tiež zrejmé, že prijatie komplexu pronatalitných opatrení koncom 60. a v prvej polovici 70. rokov sa žien vo veku 35 a viac rokov v podstate nedotklo. Len na začiatku 70. rokov identifikujeme dočasné medziročné zastavenie poklesu, ktoré však veľmi rýchlo prešlo do ďalšej fázy znižovania reprodukcie. Uvedený trend sa následne prejavil aj vo váhe plodnosti žien vo veku 35 a viac rokov na celkovej úrovni rodenia detí. Kým ešte v druhej polovici 40. rokov sa v tomto veku realizovala približne pätina, o približne dve desaťročia to už nebola ani desatina a od prvej polovice 80. do prvej polovice 90. rokov váha tejto vekovej skupiny klesla dokonca pod hranicu 5 %. Plodnosť žien starších ako 35 rokov tak na celkovej reprodukcií zohrávala v podstate marginálnu úlohu. To bol výsledok spomínaného dokončenia vedomého obmedzovania reprodukcie a veľkosti rodiny, čoraz väčšej koncentrácie plodnosti do prvej polovice reprodukčného obdobia a s tým súvisiaceho mladnutia vekového profilu mier plodnosti. Tieto zmeny sa tak následne odzrkadlili aj vo vývoji mier plodnosti a jej váhy u žien vo veku do 20 rokov. Na rozdiel od najstarších vekových skupín intenzita rodenia detí u -náštročných dievčat mala na Slovensku po druhej svetovej vojne rastúci trend. Ten bol prerušený miernym poklesom začiatkom 50. rokov a následne aj v 60. rokoch, keď v ich druhej polovici došlo k určitej stagnácii na úrovni približne 40 detí na 1000 žien. Práve v tomto období nastala situácia, keď intenzita, s akou sa matkami stávali -náštročné dievčatá, prekonalala úroveň rodenia detí dospelých žien vo veku 35 a viac rokov (graf č. 1). Na rozdiel od týchto starších vekových skupín však komplex pronatalitných opatrení priniesol určité oživenie plodnosti u najmladších žien, vďaka čomu sa daný rozdiel ešte viac prehĺbil. Vďaka tomuto vývoju došlo tiež k zvýšeniu podielu plodnosti žien vo veku do 20 rokov na celkovej plodnosti. Kým po druhej svetovej vojne sa ich príspevky pohybovali pod hranicou 5 %, od konca 70. rokov prekonalali desatinu a na začiatku

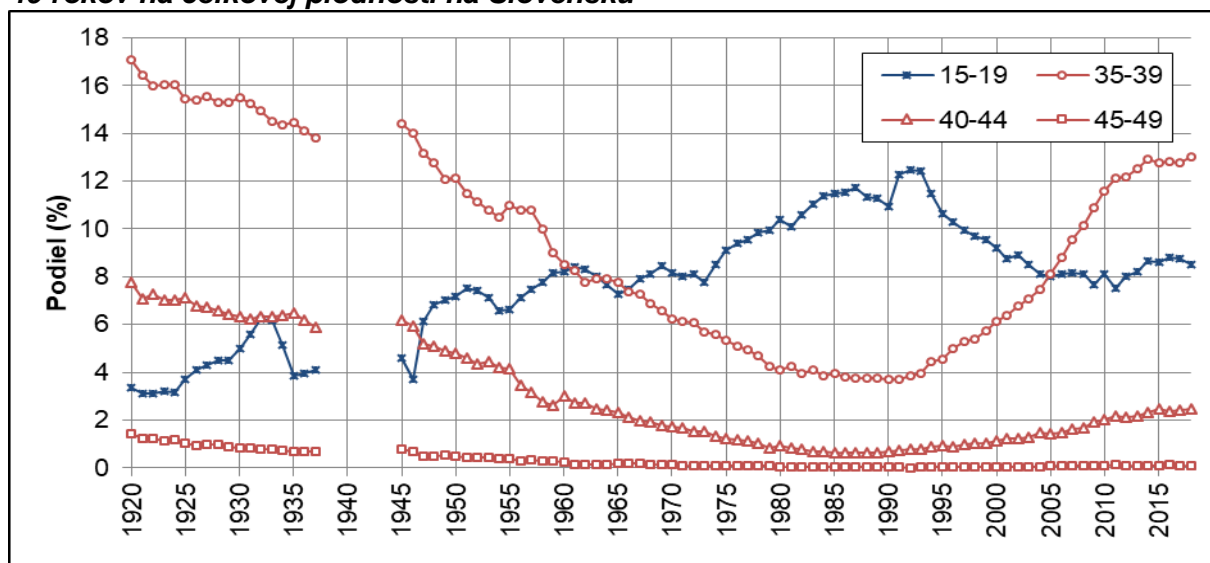
90. rokov sa dostali na svoje historické maximum s viac ako 12 % (graf č. 2). Ako sme spomenuli v úvode, nasledujúce obdobie sa nesie v historicky jedinečných transformačných zmenách reprodukčného správania, pričom jedným z hlavných znakov je odkladanie materského štartu a posun maximálnej úrovne plodnosti do vyššieho veku. To sa odzrkadilo aj na prudkom poklese intenzity a váhy plodnosti žien vo veku do 20 rokov. Približne od začiatku nového milénia však došlo k stabilizácii približne na úrovni 20 detí na 1000 žien, resp. 8 %. V posledných rokoch dokonca identifikujeme mierny nárast intenzity, ako aj samotného príspevku k celkovej úrovni plodnosti. Podľa posledných dostupných údajov sa tak v priemere vo veku 15 – 19 rokov narodí niečo viac ako 26 detí na 1000 žien daného veku, čo predstavuje približne 8,5 % z hodnoty úhrnnej plodnosti.

Graf č. 1: Vývoj mier plodnosti vo veku 15 – 19 rokov a 35 – 39, 40 – 44 a 45 – 49 rokov na Slovensku



Zdroj údajov: Pohyb obyvateľstva v Československu v rokoch 1920 – 1937 a 1945 – 1991 a Pohyb obyvateľstva na Slovensku 1992 – 2018; výpočty autora

Graf č. 2: Vývoj podielu plodnosti žien vo veku 15 – 19 rokov a 35 – 39, 40 – 44 a 45 – 49 rokov na celkovej plodnosti na Slovensku



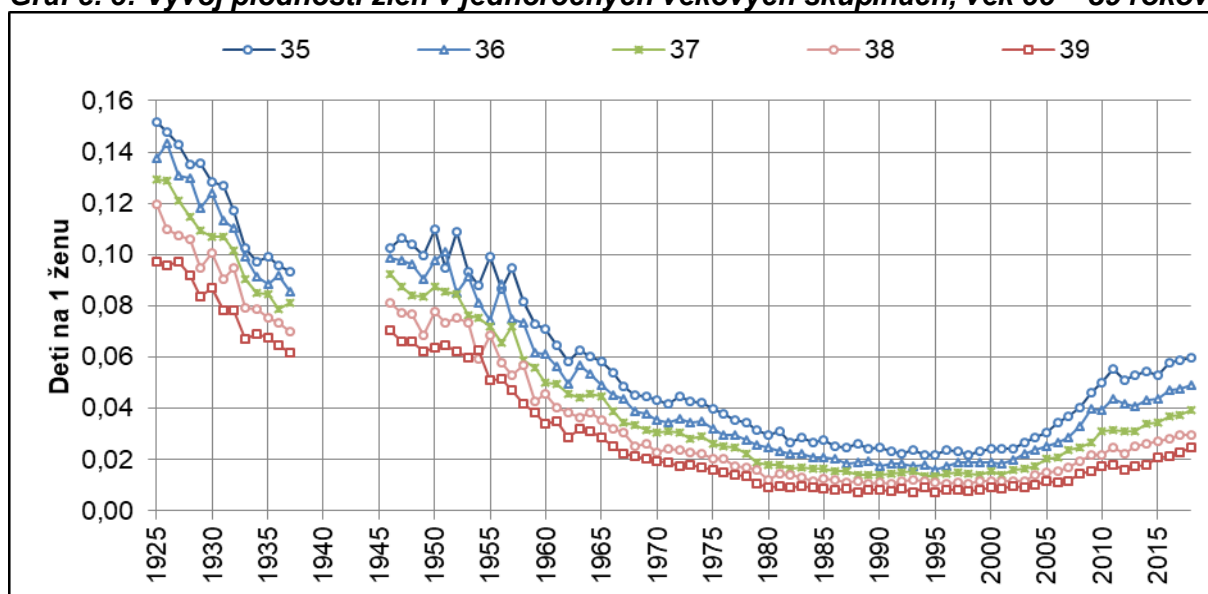
Zdroj údajov: Pohyb obyvateľstva v Československu v rokoch 1920 – 1937 a 1945 – 1991 a Pohyb obyvateľstva na Slovensku 1992 – 2018; výpočty autora

V neskorom a veľmi neskorom veku však situácia bola opačná. Od svojho minima z prvej polovice 90. rokov došlo v spojitosti s odkladaním a procesom rekuperácie týchto odložených pôrodov k nárastu intenzity rodenia detí vo veku 35 a viac rokov a tým postupne aj k zvyšovaniu váhy plodnosti v tomto vekovom spektre. Aj vzhľadom na biologické danosti a čiastočne aj vekové normy sa to týka najmä vekového intervalu 35 – 39 rokov. Približne od roku 2005 je v tejto vekovej skupine už vyššia plodnosť ako u najmladších žien a v súčasnosti dosahuje približne 40 detí na 1000 žien. Ako je však zrejmé z grafu č. 1, ani zďaleka to nie je úroveň, akú sme na Slovensku pozorovali v medzivojnovom či povojnovom období. Rovnako aj vo veku 40 a viac rokov je posledná známa úroveň plodnosti ďaleko za hodnotami, ktoré sa na Slovensku dali identifikovať ešte v 50. či na začiatku 60. rokov. Vzhľadom na zmeny v rozložení plodnosti však váha týchto vekových skupín (s výnimkou najstaršej) už nie je nezanedbateľná. Ak napríklad vo veku 35 – 39 rokov na začiatku 90. rokov sa realizovalo ani nie 5 % z celkovej plodnosti, v súčasnosti je to 13 %. Vo veku 40 – 44 rokov je to 2,5 % v porovnaní s necelým 1 % z prvej polovice 90. rokov.

Bližší pohľad na vývoj intenzity plodnosti v sledovaných vekových skupinách nám ukazuje, že v nich existujú tiež pomerne veľké diferencie v úrovni, ako aj vývojových trendoch. Vzhľadom na dlhodobu všeobecne veľmi nízku plodnosť žien vo veku nad 45 rokov sa zameriame len na vek 35 – 39 rokov a 40 – 44 rokov.

Grafy č. 3 a č. 4 potvrdzujú hlavné vývojové trendy, ktoré sme už pozorovali pri hodnotení celých 5-ročných vekových skupín. Ukazuje sa tiež, že na Slovensku dlhodobo platí negatívny vekový gradient, keď s každým jedným starším vekom klesala intenzita plodnosti. Súčasne je však zrejmé, že tieto diferencie medzi jednotlivými vekovými skupinami mali dlhodobu klesajúci trend spolu s celkovým poklesom úrovne plodnosti. Najmenšie rozdiely tak nachádzame v období najnižšej úrovne plodnosti – v prvej polovici 90. rokov. S postupným presadzovaním procesu rekuperácie a s ňou spojeného oživenia plodnosti nastáva tak pomerne značná divergencia v úrovni plodnosti medzi jednotlivými jednotkami vekov, pričom platí, že s každým ďalším vekom intenzita rodenia detí signifikantne klesá.

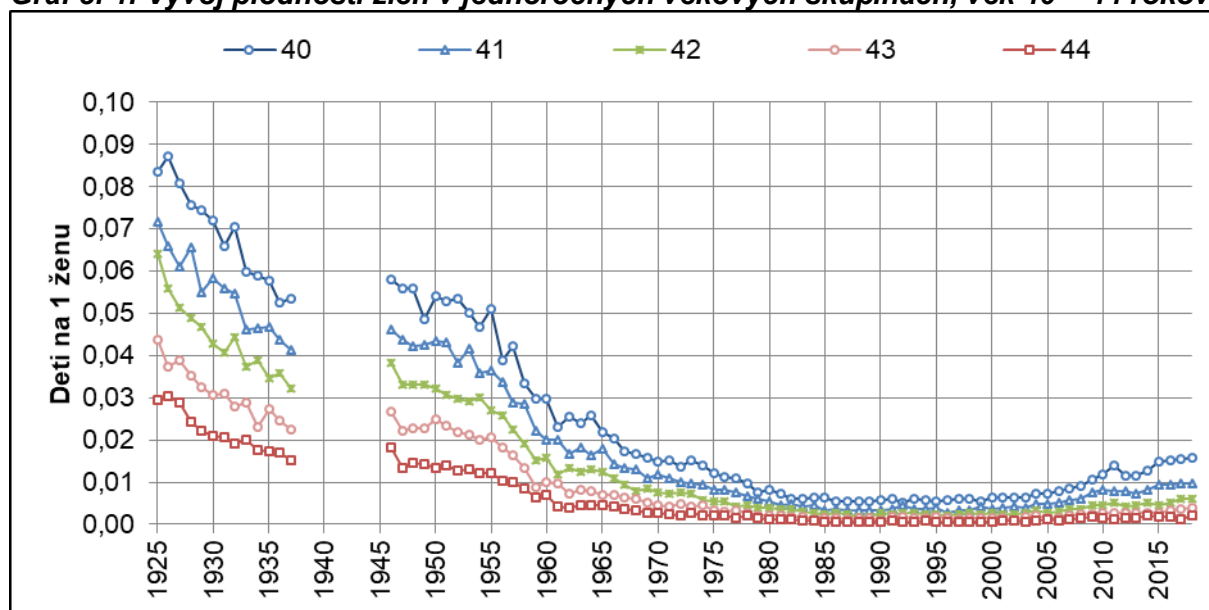
Graf č. 3: Vývoj plodnosti žien v jednoročných vekových skupinách, vek 35 – 39 rokov



Zdroj údajov: Pohyb obyvateľstva v Československu v rokoch 1920 – 1937 a 1945 – 1991 a Pohyb obyvateľstva na Slovensku 1992 – 2018; výpočty autora

To je veľmi dôležitá informácia najmä v súvislosti s procesom odkladania a predlžovania obdobia bezdetnosti, resp. realizáciou cieleného počtu narodených detí. V prípade zväčšovania rozdielov medzi jednotkami veku, teda keď s každým ďalším vekom nedochádza k naplneniu reprodukcie, sa zvyšuje a prehľbuje riziko, že tie sa nemusia napokon realizovať. Je potrebné si pritom uvedomiť, že s presunom čoraz väčšej časti plodnosti (a najmä rodenia prvého dieťaťa) do druhej polovice reprodukčného veku sa pri tomto stave výrazne skracuje reprodukčný interval a pri snahe stať sa matkou viackrát sa tieto reprodukčné zámery čoraz viac musia koncentrovať do úzkeho vekového rozpätia. To so sebou, samozrejme, prináša riziko ich nenaplnenia a najmä otázky spojené s rekuperáciou druhého a prípadne ďalšieho dieťaťa.

Graf č. 4: Vývoj plodnosti žien v jednoročných vekových skupinách, vek 40 – 44 rokov



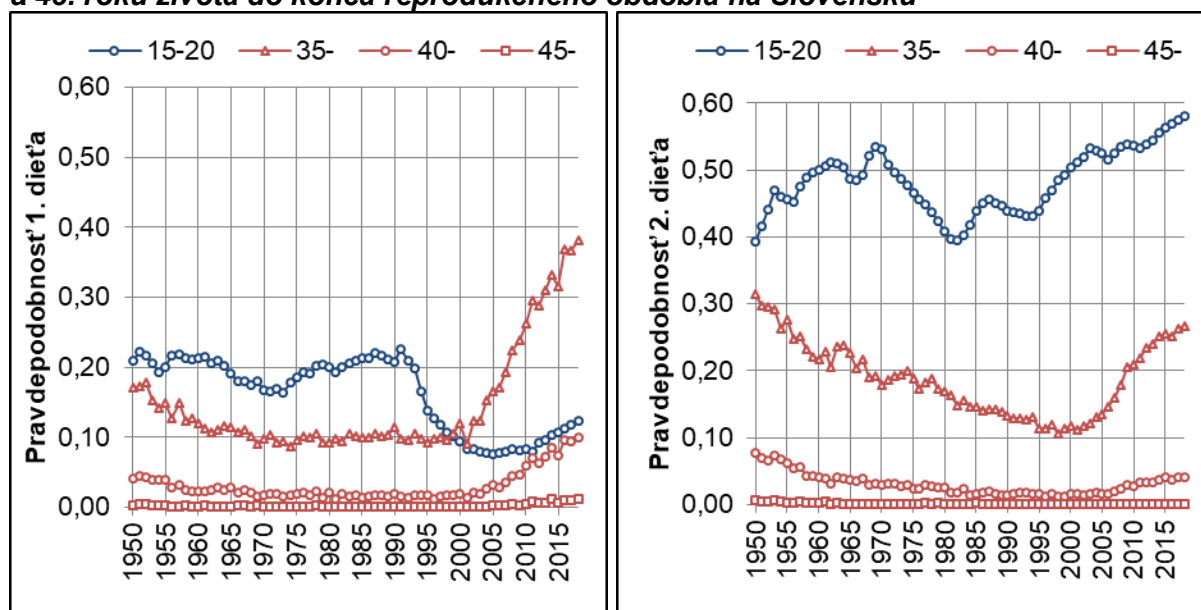
Zdroj údajov: Pohyb obyvateľstva v Československu v rokoch 1920 – 1937 a 1945 – 1991 a Pohyb obyvateľstva na Slovensku 1992 – 2018; výpočty autora

Najlepšie je možné pozorovať vývoj intenzity plodnosti v sledovaných vekových skupinách v spojitosti s poradím prostredníctvom podmienených pravdepodobností. Tie vychádzajú z prierezových inkrementno-dekrementných tabuliek plodnosti, ktoré sú pre Slovensko konštruované od roku 1950. Grafy č. 5 a 6 umožňujú sledovať pravdepodobnosť narodenia prvého dieťaťa bezdetnej žene a druhého dieťaťa žene s jedným dieťaťom v príslušných vekových skupinách. Pri materskom štarte sa potvrdzuje povojnový pokles šancí stať sa matkou vo vyššom veku a následná stabilizácia na pomerne nízkej úrovni. Dlhodobo pritom mali bezdetné -násťročné dievčatá vyššiu pravdepodobnosť stať sa matkou medzi 15. a 20. rokom života ako ženy od 35., 40. a 45. narodenín do konca reprodukčného veku (50. roku života). V poslednom desaťročí 20. storočia sa však táto stabilita končí a sme svedkami výrazných zmien, ktoré už predznamenal vývoj vekovo špecifických mier plodnosti. Predovšetkým dochádza k výraznému nárastu pravdepodobnosti, že sa bezdetná žena vo veku 35 rokov stane ešte aspoň raz matkou pred ukončením reprodukčného obdobia. V súčasnosti už táto pravdepodobnosť dosahuje takmer 40 %, kým počas takmer celého minulého politického režimu to bolo len približne 10 % (pozri graf č. 5). Pomerne výrazne vzrástla tiež pravdepodobnosť, že sa bezdetná žena stane matkou vo veku nad 40 rokov. Dlhodobo pritom bola pravdepodobnosť vzniku tejto

demografickej udalosti na Slovensku v podstate zanedbateľná (graf č. 5). Podľa posledných dostupných údajov už dosahuje hranicu 10 %. Ako je zrejmé z grafu č. 5, po dovŕšení 45. roku života však pravdepodobnosť narodenia prvého dieťaťa stále zostáva minimálna.

U najmladších žien sa pravdepodobnosť vzhľadom na prudké odkladanie materského štartu najprv pomerne výrazne znížila. Z viac ako 20 % na začiatku 90. rokov klesla o viac ako polovicu. V poslednej dekáde však dochádza k nárastu a pravdepodobnosť, že sa bezdetná žena stane matkou medzi 15. a 20. rokom života už prekračuje hranicu 12 %. Vysvetlenie tohto špecifického vývojového trendu by sa mohlo opierať o rastúci vplyv žien z marginalizovaných rómskych komunit na celkovú plodnosť [bližšie pozri 10], u ktorých skorý začiatok reprodukčnej dráhy stále zostáva veľmi dôležitou súčasťou nastavenia ich prokreatívneho správania [10].

Graf č. 5 a 6: Pravdepodobnosť narodenia prvého dieťaťa bezdetnej žene a druhého dieťaťa žene s jedným dieťaťom medzi 15. a 20. rokom života a po dovŕšení 35., 40. a 45. roku života do konca reprodukčného obdobia na Slovensku



Zdroj údajov: ŠÚ SR, výpočty autora

Koncentrácia plodnosti do úzkeho vekového intervalu a s tým spojené časté ukončenie reprodukčnej dráhy ešte pred dovŕšením 30. roku života sa premietlo aj do postupného kontinuálneho poklesu pravdepodobností narodenia druhého dieťaťa ženám, ktoré sa už raz stali matkou po dovŕšení 40. a najmä 35. roku. Kým na začiatku 50. rokov jej hodnota ešte prekračovala hranicu 30 %, v druhej polovici 90. rokov dosahovala minimálnu úroveň. Približne len desiatim ženám zo 100, ktoré mali jedno dieťa sa po 35. roku narodilo ďalšie dieťa. Vo vyššom veku táto pravdepodobnosť bola v podstate zanedbateľná. Približne od začiatku nového milénia však pozorujeme pomerne prudký rast týchto pravdepodobností. Ten súvisí tiež s naštartovaním procesu rekuperácie a to aj v prípade druhých detí. V súčasnosti tak žena s jedným dieťaťom v presnom veku 35 rokov má už takmer 26 % šancu, že sa jej do konca reprodukčného obdobia narodí ešte jedno dieťa. Vo veku 40 rokov táto pravdepodobnosť však stále zostáva veľmi nízka, a to zhruba na úrovni 4 %. Je preto zrejmé, že ak sa ženám nepodarí stať matkou a tiež matkou dvoch detí pred

dovršením 40. roku života, je zatiaľ na Slovensku veľmi malá šanca, že sa na tom niečo ešte zmení do dovŕšenia konca reprodukčného obdobia.

Vývoj podmienených pravdepodobností druhého dieťaťa tiež poukazuje na dôležitý jav platný pre populáciu Slovenska v podstate dlhodobo. Ak sa žena stala matkou vo veľmi mladom veku, existovala pomerne vysoká šanca, že sa jej narodí v mladom veku aj druhé dieťa. Z grafu č. 6 je zrejmé, že táto pravdepodobnosť oscilovala na úrovni 40 – 50 %, pričom v poslednom období dokonca pomerne prudko rastie a už takmer dosahuje hranicu 60 %. Opätovne však musíme upozorniť, že je to reprodukčný model len určitej špecifickej populačnej podskupiny, pretože skoré až veľmi skoré materstvo a rodičovstvo sa stávajú na Slovensku marginálnymi reprodukčnými stratégiami. Ak však k tomu predsa len dôjde, vývoj pravdepodobností ukazuje, že sa posilňuje šanca na to, že sa takáto žena stane matkou aj druhýkrát a to ešte pred dovŕšením 20. roku života. Posledné vývojové zmeny v reprodukčnom správaní tak na jednej strane poukazujú na pokračujúci trend presunu plodnosti do vyššieho veku s čoraz vyššími šancami rodenia prvého i druhého dieťaťa až po 35 rokoch, no súčasne tiež poukazujú na zvyšujúcu sa váhu a intenzitu rodenia detí určitej podskupiny -násťročných žien. To sa následne prejavuje aj na priebehu vekovo špecifických mier plodnosti, keď okolo 20. roku života vzniká a postupne sa formuje určité sekundárne maximum. Spoločne tieto trendy prispievajú k prehĺbovaniu celkovej vekovej diverzifikácii reprodukčných dráh žien na Slovensku.

3. PRIESTOROVÁ DIFERENCIÁCIA SKORÉHO A NESKORÉHO MATERSTVA NA SLOVENSKU

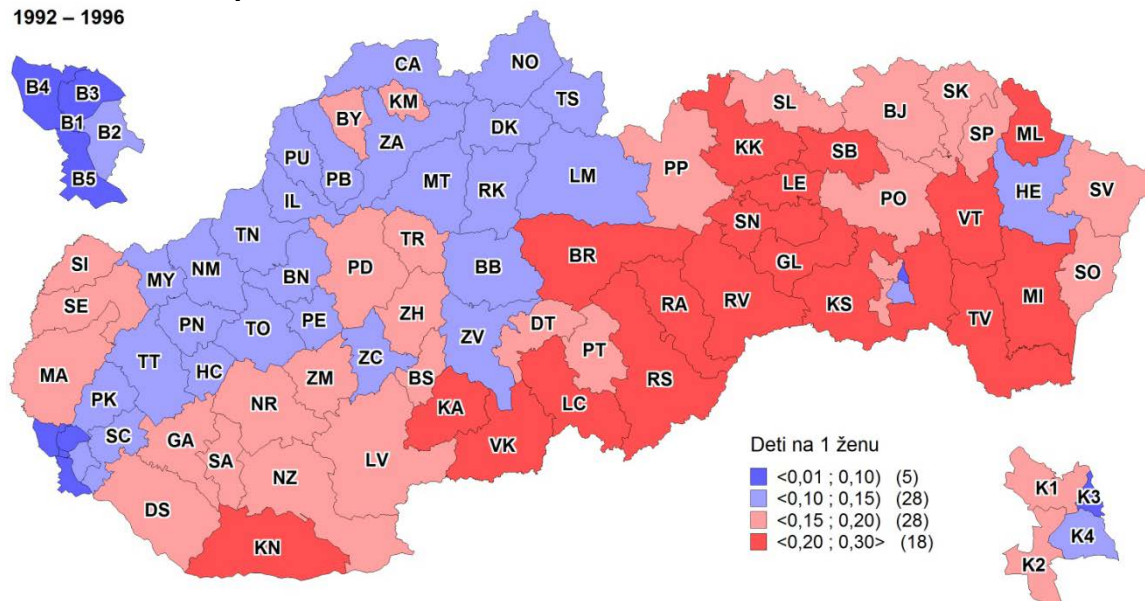
Z hľadiska priestorovej diferenciácie plodnosti na Slovensku sa pri určitom zovšeobecnení dá povedať, že na západnom a z veľkej časti aj strednom Slovensku (s výnimkou okresov krajného severu a juhu) sa postupne vytvára model poznačený výrazným odkladaním rodenia detí do vyššieho veku, pričom pre celkovú intenzitu plodnosti je rozhodujúcim faktorom úspešnosť žien pri realizácii týchto odložených reprodukčných zámerov. Na druhej strane stojí väčšina okresov východného a severovýchodného Slovenska (s výnimkou krajných prihraničných regiónov), kde pozorujeme pretrvávanie skoršieho začiatku reprodukčnej dráhy a vo všeobecnosti aj celkovo vyššiu úroveň plodnosti. Tá je v mnohých smeroch ovplyvnená najmä prítomnosťou rómskeho etnika a predovšetkým žien žijúcich v špecifickom prostredí marginalizovaných rómskych komunít.

Priestor s najvyššou celkovou intenzitou plodnosti žien vo veku do 20 rokov sa od 90. rokov zmenšil a obmedzuje sa na okresy ležiace na juhu stredného a východného Slovenska v páse od Lučenca po Michalovce. Na severe do tejto skupiny v súčasnosti patria ešte ďalšie 4 okresy. Znamená to, že v porovnaní s prvou polovicou 90. rokov došlo k značnej redukcii plodnosti najmladších žien na juhu a juhozápade Slovenska, pričom sa vytvoril v podstate kompaktný priestor západného a väčšiny okresov stredného Slovenska s veľmi nízkou intenzitou rodenia detí pred dovŕšením 20. roku života. V prvom sledovanom období pritom išlo len o niektoré mestské okresy Bratislavy a Košíc. Vo všeobecnosti tak môžeme v súčasnosti už Slovensko rozdeliť na oblasť západu a stredú s nízkou až veľmi nízkou plodnosťou -násťročných dievčat a naopak, časť juhu stredného a väčšinu okresov východného Slovenska tvoriacich oblasť s nadpriemernou až výrazne nadpriemernou intenzitou rodenia detí vo veku do 20 rokov.

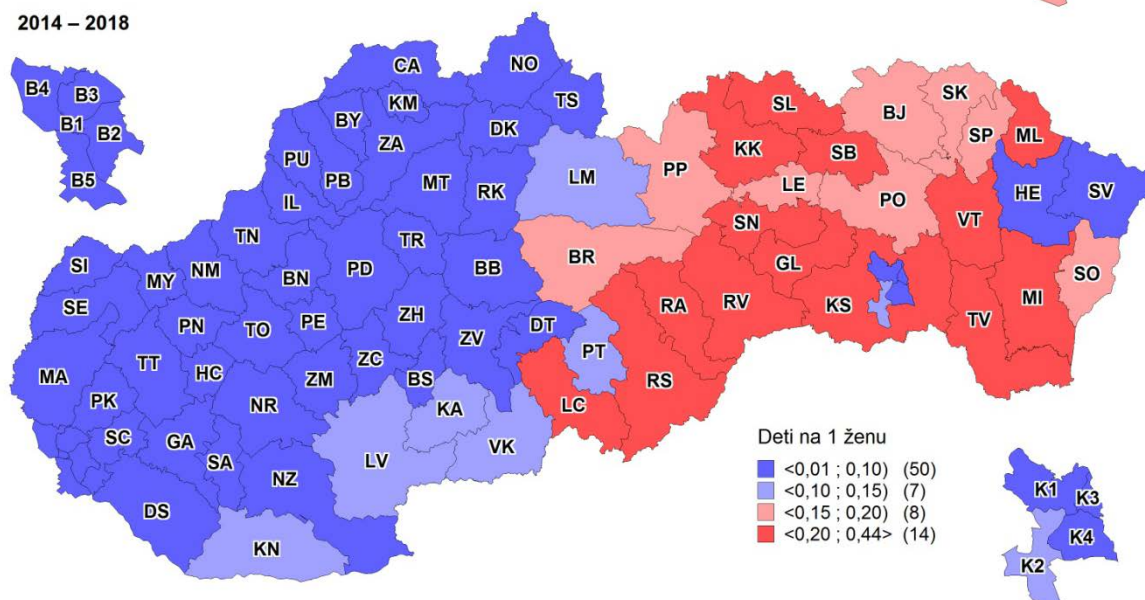
Obdobný priestorový obraz poskytuje aj rozloženie príspevkov plodnosti žien vo veku do 20 rokov k celkovej plodnosti v jednotlivých okresoch Slovenska. Jednak sledujeme redukciu samotnej úrovne podielu plodnosti danej skupiny žien, ako aj zmenšenie priestoru s maximálnymi príspevkami a jeho redukciu na okresy východného Slovenska. V dvoch z nich (Rimavská Sobota, Gelnica) sa ženy vo veku do 20 rokov podieľajú viac ako pätinou na celkovej plodnosti. V ďalších 9 okresoch v tomto priestore sa príspevky pohybujú na úrovni 15 – 20 % (pozri obr. č. 2).

Obr. č. 1: Celková plodnosť žien vo veku do 20 rokov v okresoch Slovenska

1992 – 1996



2014 – 2018



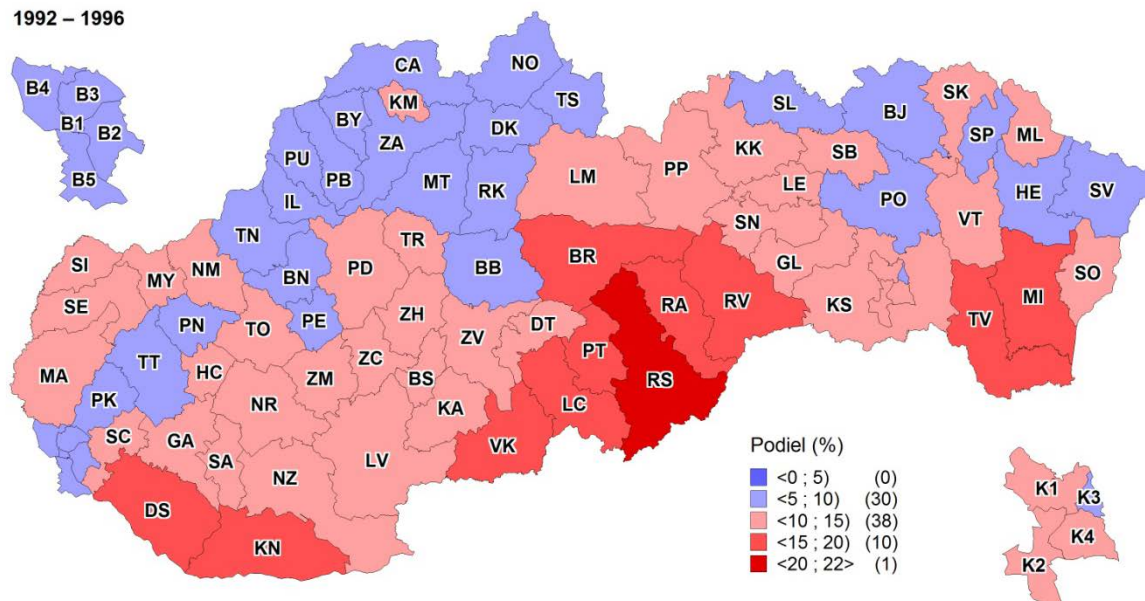
Zdroj údajov: ŠÚ SR, výpočty autora

Vo väčšine okresov západného a stredného Slovenska však podiel plodnosti -náštročných žien neprekračuje hranicu 5 %. Také nízke podiely v prvej polovici 90. rokov nedosahovali pritom žiadne okresy na Slovensku. Aj to svedčí o dynamickej úrovni odkladania plodnosti v tomto veku, jeho marginálnej pozícií z pohľadu reprodukcie a naopak, špecifickom charaktere reprodukcie okresov, v ktorých je jej váha nadpriemerná.

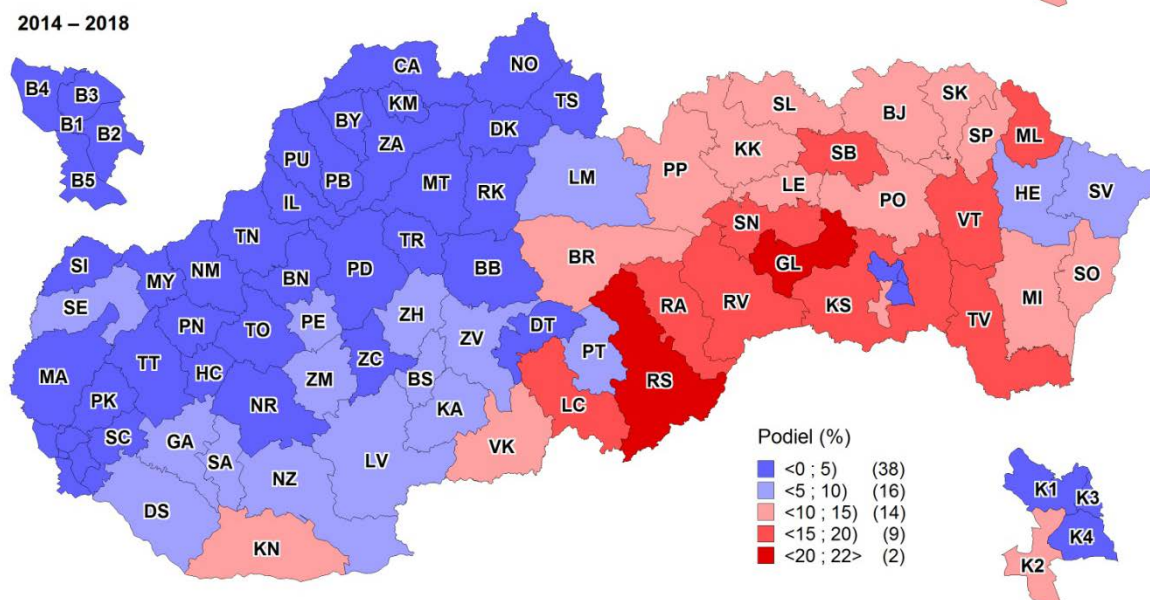
Ako je zrejmé z obr. č. 3 a č. 4, vyššie spomínaná výrazná koncentrácia plodnosti do prvej polovice reprodukčného veku sa ešte aj v 90. rokoch minulého storočia odrážala na veľmi nízkej intenzite reprodukcie vo veku 35 – 39 rokov a ich príspevkoch na celkovej plodnosti v podstate vo väčšine okresoch Slovenska. Len na severe stredného a východného Slovenska môžeme vzhľadom na častejšie rodenie tretieho a ďalších detí identifikovať mierne vyššiu úroveň a podiel plodnosti v tomto veku.

Obr. č. 2: Podiel plodnosti žien vo veku do 20 rokov k celkovej plodnosti žien v okresoch Slovenska

1992 – 1996



2014 – 2018

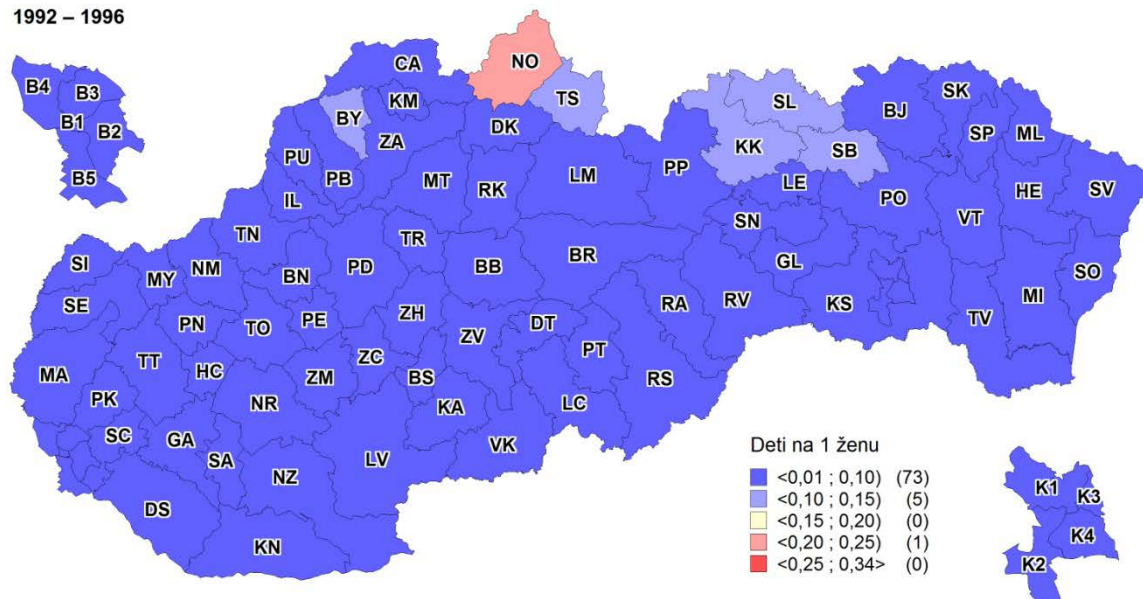


Zdroj údajov: ŠÚ SR, výpočty autora

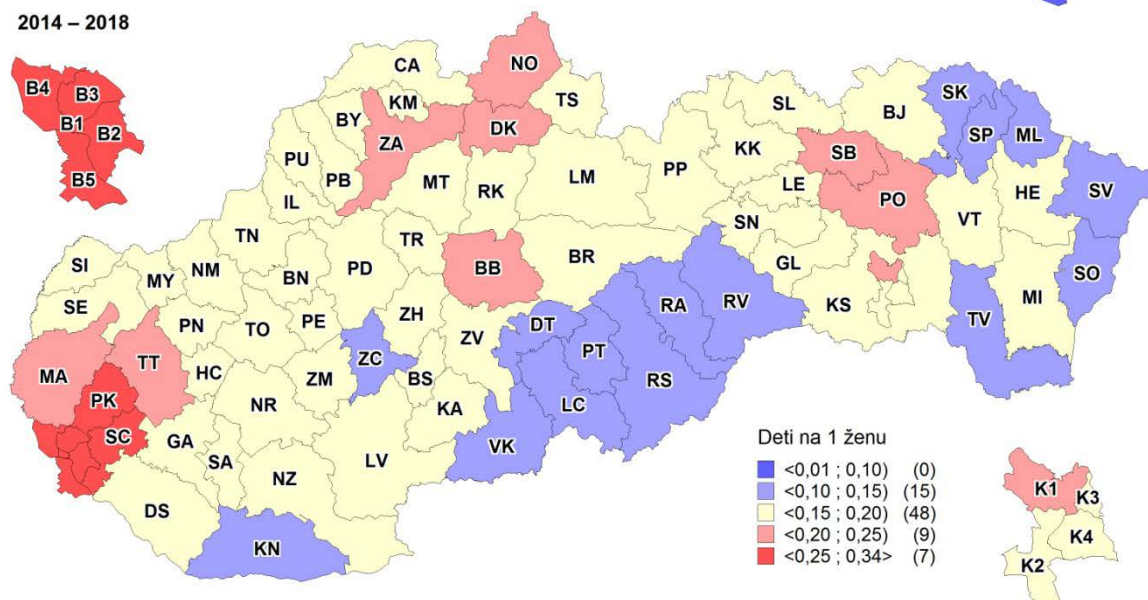
Pomerne intenzívne odkladanie reprodukcie a naštartovanie procesu rekuperácie prispeli, že vo veku 35 – 39 rokov sa v podstate vo všetkých okresoch Slovenska plodnosť do súčasnosti zvýšila a rovnako vzrástla aj jej váha na celkovej hodnote úhrnnej plodnosti. Hlavnú úlohu plodnosť v tomto veku zohráva v okresoch Bratislavy a v jej zázemí, ďalej v okrese Trnava, Piešťany a Trenčín na západnom, v okrese Zvolen a Banská Bystrica na strednom Slovensku a v okrese Košice I na východe.

Obr. č. 3: Celková plodnosť žien vo veku 35 – 39 rokov v okresoch Slovenska

1992 – 1996



2014 – 2018



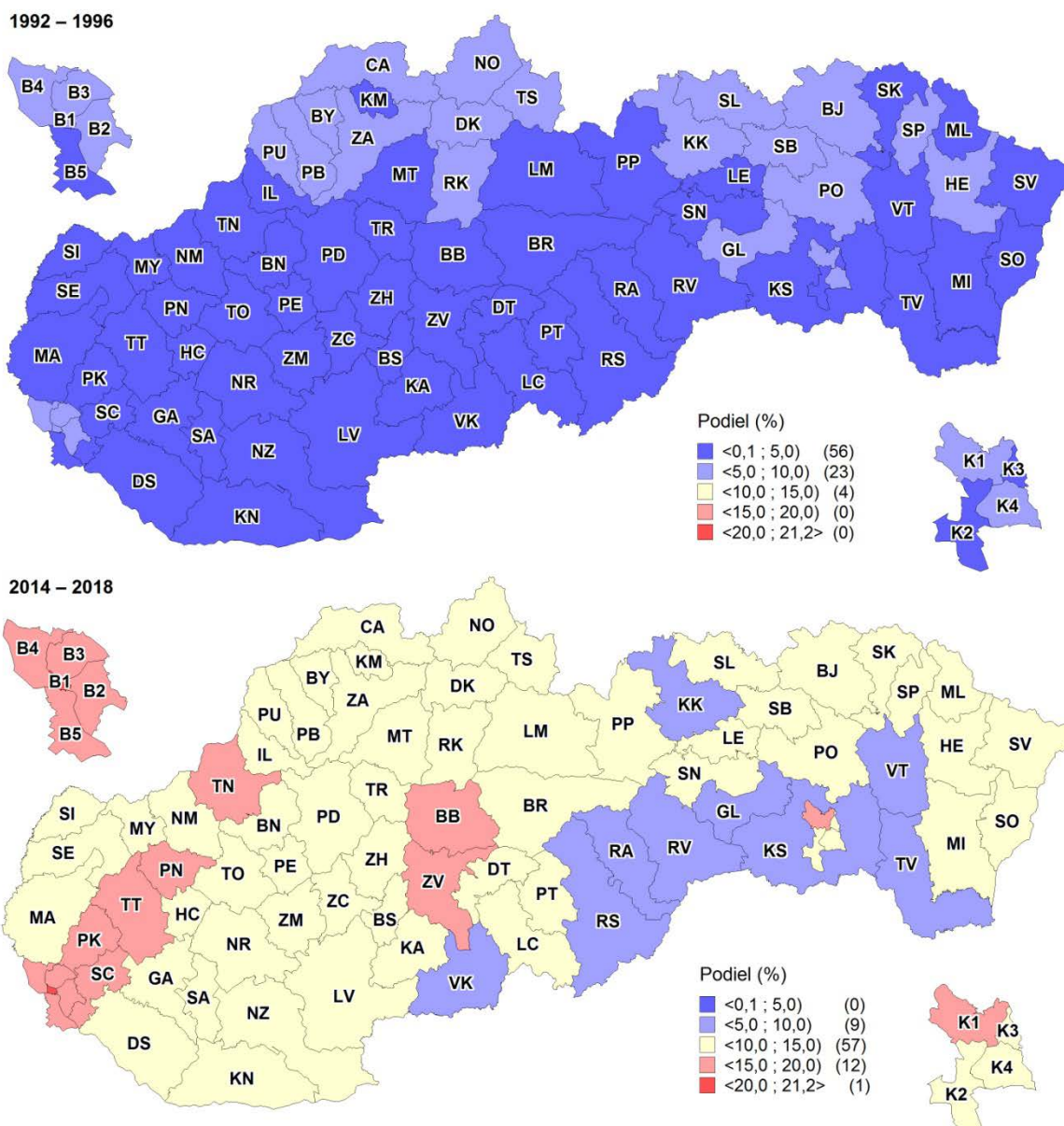
Zdroj údajov: ŠÚ SR, výpočty autora

V týchto celkoch sa plodnosť žien vo veku 35 – 39 rokov podieľa na 15 – 22 % z celkovej reprodukcie. Vo väčšine okresov Slovenska je jej príspevok na úrovni 10 – 15 %, kým v polovici 90. rokov to nebolo ešte ani 5 %. Podpriemernú intenzitu, ako aj váhu plodnosti predmetnej vekovej skupiny nachádzame v súčasnosti len na juhu stredného a východného Slovenska v páse prihraničných okresov (pozri obr. č. 3 a č. 4).

Veľmi neskorá plodnosť aj napriek určitému zvýšeniu od prvej polovice 90. rokov stále zostáva vo väčšine regiónoch Slovenska veľmi nízka. Ako ukazujú posledné dostupné údaje z rokov 2014 – 2018, s výnimkou okresu Bratislava I sa nepodieľa ani na 5 % z celkovej úrovne reprodukcie. Nadpriemernú plodnosť žien vo veku 40 a viac rokov môžeme tak z celoslovenského hľadiska identifikovať len v Bratislave, v okrese Pezinok a na severe Slovenska v okrese Námestovo (pozri obr. č. 5).

Aj napriek len miernemu nárastu celkovej plodnosti vo veku 40 a viac rokov a jej stále pomerne nízkej úrovni môžeme na Slovensku vidieť postupné utváranie niektorých oblastí, v ktorých tento proces prebieha o niečo rýchlejšie a tým sa pokročilý a veľmi pokročilý reprodukčný vek využíva o niečo častejšie na realizáciu reprodukčných plánov. Okrem vyššie spomínaných okresov ide tiež o zázemie hlavného mesta, ďalej mestské okresy Košíc, viaceré okresy na Považí a pomerne súvislý priestor severného Slovenska, ku ktorému sa pripája z pohľadu intenzity aj väčšina okresov východného Slovenska. Určité rozdiely pritom pozorujeme z hľadiska samotného formovania týchto priestorov. Kým okresy s veľkými hospodárskymi centrami a čiastočne aj viaceré celky na západe Slovenska sú priestorom, kde sa v pokročilom veku realizuje aj veľká časť materských štartov, na severe a východe Slovenska sú dôležitým pre vyššiu úroveň a váhu plodnosti v tomto veku najmä deti vyššieho poradia.

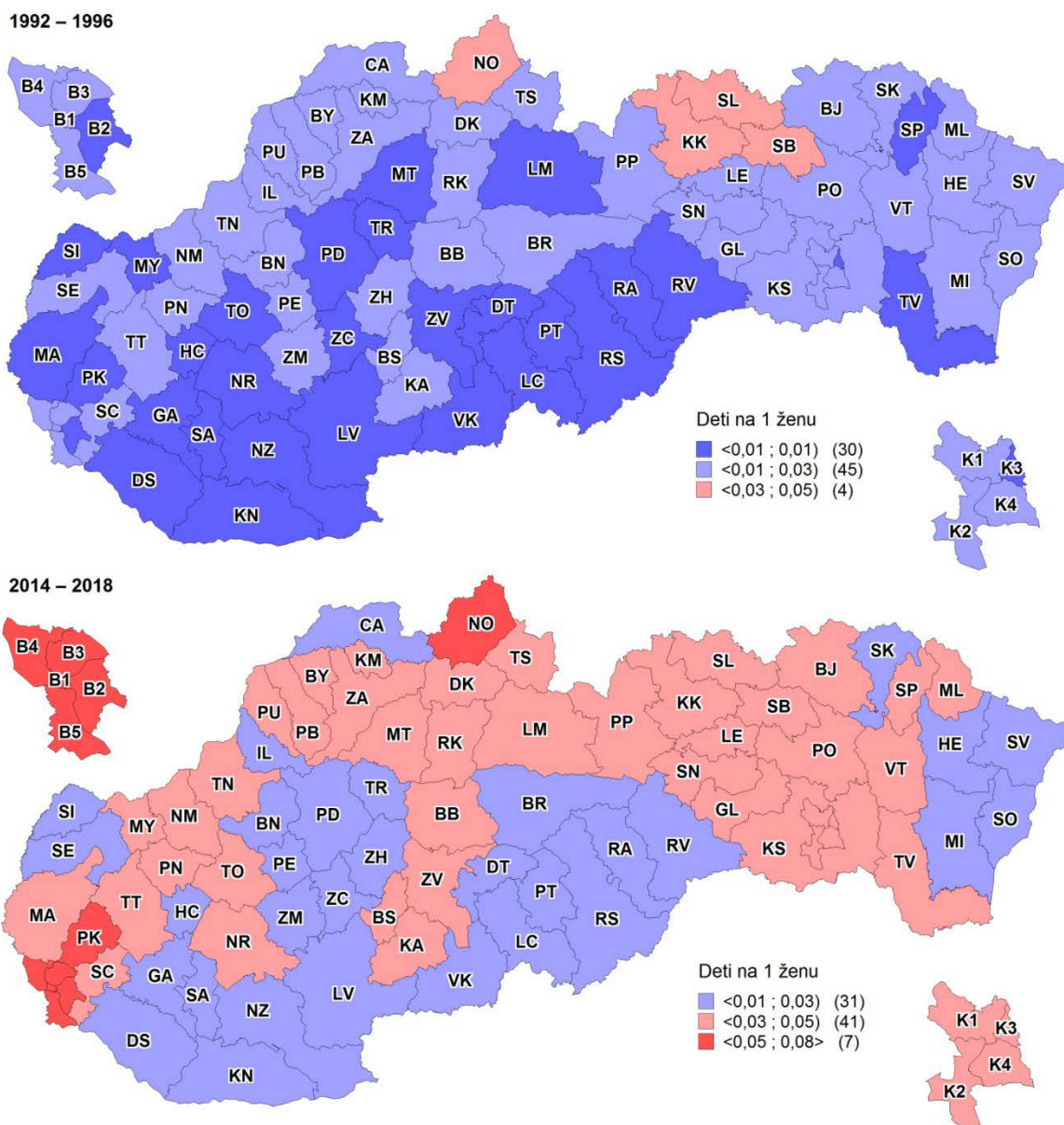
Obr. č. 4: Podiel plodnosti žien vo veku 35 – 39 rokov z celkovej plodnosti žien v okresoch Slovenska



Zdroj údajov: ŠÚ SR, výpočty autora

Vo všeobecnosti je však zrejmé, že aj na regionálnej úrovni nie je tento pokročilý a veľmi pokročilý reprodukčný vek zatiaľ výraznejšie využívaný na realizáciu reprodukčných zámerov a drvivá väčšina z nich sa odohráva ešte pred dovŕšením 40. roku života. Na jednej strane môžeme hovoriť o prirodzenom biologickom poklese fekundity, no pri porovnaní s úrovňou plodnosti vo veku 40 a viac rokov s medzivojnovým obdobím, ako aj niektorými európskymi krajinami s najpokročilejším procesom odkladania reprodukcie je zrejmé, že Slovensko a väčšina jeho regiónov má v tomto veku ešte značné rezervy.

Obr. č. 5: Celková plodnosť žien vo veku 40 a viac rokov v okresoch Slovenska

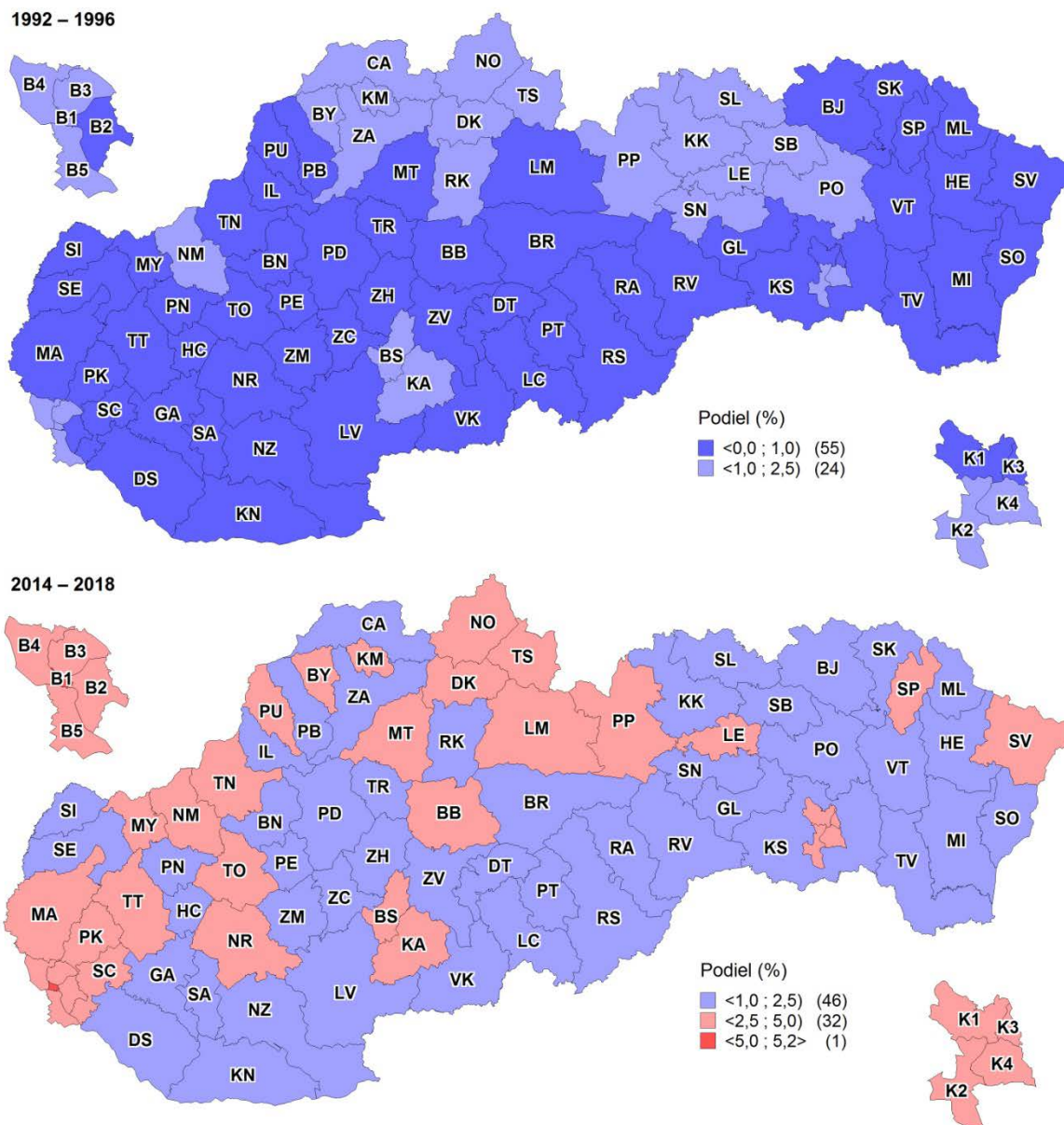


Zdroj údajov: ŠÚ SR, výpočty autora

Postupný nárast plodnosti v pokročilom reprodukčnom veku signalizuje, že do budúcnosti budeme svedkami ďalšieho zvyšovania nielen samotnej intenzity, ale aj váhy plodnosti žien vo veku nad 40 rokov. V spojitosti s tým je potrebné ešte spomenúť existenciu vekových noriem na nevhodné materstvo a z biologického hľadiska zatiaľ obmedzené možnosti (v porovnaní so zahraničím) umelého

oplodnenia. Práve to sa môže stať jedným z kľúčových faktorov rýchlejšieho zvyšovania intenzity reprodukcie v pokročilom reprodukčnom období

Obr. č. 6: Podiel plodnosti žien vo veku 40 a viac rokov z celkovej plodnosti žien v okresoch Slovenska



Zdroj údajov: ŠÚ SR, výpočty autora

4. ZÁVER

Plodnosť žien na Slovensku sa v dlhodobom kontexte po prvej svetovej vojne vyznačovala pomerne výraznou redukciou intenzity v pokročilom reprodukčnom veku, spočiatku najmä ako dôsledok pokračujúcej demografickej revolúcie a vedomého obmedzovania veľkosti rodiny a neskôr v období reálneho socializmu ako prejav reakcie na špecifické reprodukčné podmienky vytvorené minulým politickým režimom. S tým je úzko spojená výrazná koncentrácia plodnosti do prvej polovice reprodukčného veku a pokročilý a veľmi pokročilý reprodukčný vek sa stali v podstate marginálnymi a nepodstatnými na realizáciu reprodukčných zámerov. Na druhej strane však kombinácia viacerých aspektov prispela k čoraz väčšej preferencii

skorého a veľmi skorého materstva, čo sa odzrkadlilo na náraste plodnosti žien vo veku do 20 rokov. Tento špecifický reprodukčný vek tak zohrával až do konca minulého storočia dôležitejšiu úlohu pri celkovej reprodukcii na Slovensku ako plodnosť žien vo veku nad 35 rokov.

Výrazná, dynamicky prebiehajúca diskontinuita životných podmienok mladých ľudí po roku 1989 však výrazným spôsobom prispela k celkovej premene modelov reprodukcie. Do popredia sa dostáva proces odkladania manželského a materského štartu a čoraz väčšia časť z nich sa orientuje do polovice alebo až druhej polovice reprodukčného veku. S tým sa zvyšuje aj intenzita rodenia detí vo veku 35 a viac rokov a súčasne klesá plodnosť v mladom veku. Tento proces navyše ešte urýchlilo oživenie reprodukcie spojené s procesom dobiehania odložených reprodukčných zámerov. Z hľadiska mladého veku však vidíme, že v posledných rokoch dochádza k zvyšovaniu intenzity plodnosti a tiež k nárastu pravdepodobnosti stať sa matkou a porodiť druhé dieťa pred dovŕšením 20. roku života. Môžeme predpokladať, že v tomto vekovo špecifickom reprodukčnom období sa v čoraz väčšej miere prejavuje špecifický charakter reprodukčného správania žien pochádzajúcich najmä z marginalizovaných rómskych komún. Do určitej miery by to potvrdzovalo aj priestorové rozloženie celkov s maximálnou úrovňou a podielom plodnosti žien vo veku do 20 rokov. V podstate sa na Slovensku vytvoril súvislý priestor väčšej časti východného Slovenska s nadpriemernou úrovňou plodnosti v tomto veku a naopak, západné a stredné Slovensko sa vyznačujú podpriemernou intenzitou a váhou rodenia detí v najmladšom veku. Na druhej strane pokročilý a veľmi pokročilý reprodukčný vek stále zostávajú marginálnymi v procese reprodukcie. Platí to nielen na národnej, ale aj regionálnej úrovni, kde existujúce rozdiely sú v podstate veľmi malé. Vzhľadom na rastúci trend však môžeme predpokladať, že do budúcnosti aj na Slovensku budeme svedkami ďalšieho zvyšovania intenzity a váhy plodnosti žien vo veku nad 40 rokov.

LITERATÚRA

- [1] BEAUJOUAN, E. – SOBOTKA, T.: Late Motherhood in Low-Fertility Countries: Reproductive Intentions, Trends and Consequences. Human Fertility Database Research Report HFD RR-2017-002, 2017.
- [2] BILLARI, F. C. – KOHLER, H.-P. – ANDERSSON, G. – LUNDSTRÖM, H.: Approaching the Limit: Long-Term Trends in Late and Very Late Fertility. In: Population and Development Review, 2007, č. 1, s. 149 – 170.
- [3] HAJNAL, J.: European marriage pattern in historical perspective. In: Glass, D. V. – Eversley, D.E.C. (ed.): Population in History. Arnold, Londres, 1965, s. 101 – 143.
- [4] KOHLER, H. P. – BILLARI, F. – ORTEGA, J.A.: The emergence of Lowest-Low Fertility in Europe during the 1990s. In: Population and Development Review, 2002, č. 4, s. 641 – 680.
- [5] PRIOUX, F.: Late fertility in Europe: some comparative and historical data. In: Revue d'Épidémiologie et de Santé Publique, 2005, n. HS2, s 3 – 11.
- [6] SOBOTKA, T.: Postponement of Childbearing and Low Fertility in Europe. Groningen: Rijksuniversiteit Groningen, 2004.
- [7] SOBOTKA, T.: Fertility in Central and Eastern Europe after 1989: Collapse and Gradual Recovery. In: Historical Social Research, 2011, č. 2, s. 246 – 296.
- [8] SOBOTKA, T.: Shifting Parenthood to Advanced Reproductive Ages: Trends, causes and Consequences. In: Tremmel, J. (ed.): A Young Generation Under

Pressure? The Financial Situation and the „Rush Hour“ of the Cohorts 1970-1985 in a Generational Comparison. Springer, 2010, s. 129 – 154.

[9] MÉSZÁROS, J. – VAŇO, B.: Reprodukčné správanie obyvateľstva v obciach s nízkym životným štandardom. Bratislava: INFOSTAT, 2004.

[10] ŠPROCHA, B.: Reprodukcia rómskeho obyvateľstva na Slovensku a prognóza jeho populačného vývoja. Bratislava: OFPRINT JH, 2014. 177 s. ISBN 978-80-89037-38-4.

[11] ŠPROCHA, B. – TIŠLIAR, P.: Transformácia plodnosti žien Slovenska v 20. a na začiatku 21. storočia. Bratislava: Centrum pre historickú demografiu a populačný vývoj Slovenska FiF UK v Bratislave, 2016. 376 s. ISBN 978-80-89881-02-4.

[12] ŠPROCHA, B. – TIŠLIAR, P.: 100 rokov obyvateľstva Slovenska. Od vzniku Československa po súčasnosť. Centrum pre historickú demografiu a populačný vývoj Slovenska FiF UK v Bratislave, 2018. 445 s. ISBN 978-80-89881-14-7.

[13] TOULEMON, L.: Who are the late mothers. In: Revue d'Épidémiologie et de Santé Publique, 2005, n. HS2, s. 13 – 24.

RESUMÉ

Cieľom príspevku bolo analyzovať vývoj a podiel plodnosti žien v mladom (do 20 rokov) a staršom veku (35 a viac rokov) so zameraním na pokročilý a veľmi pokročilý vek. Výsledky sú zasadené do širšieho konceptu premeny reprodukčného správania prebiehajúceho v rámci demografickej revolúcie, po druhej svetovej vojne počas špecifického obdobia minulého politického režimu a po roku 1989 v kontexte celospoločenských zmien a tranzície plodnosti odkladaním. Okrem toho sme analyzovali aj priestorové rozdiely (na úrovni okresov) a zmeny v intenzite a zastúpení plodnosti mladých a starších žien.

Získané výsledky ukazujú na dlhodobý pokles plodnosti žien v staršom veku a naopak posilňovanie intenzity a váhy v mladom veku. Tento scenár sa však po roku 1989 výrazne mení a do popredia sa dostáva proces odkladania reprodukcie a s ním spojený nárast intenzity a podielu plodnosti v staršom veku. Pokročilý a veľmi pokročilý reprodukčný vek sa na Slovensku zatiaľ využívajú len značne obmedzene. Platí to aj pre regionálnu úroveň. Na druhej strane v prípade plodnosti v mladom veku existujú pomerne výrazné regionálne rozdiely. Zdá sa, že sú spojené najmä s existenciou špecificky sa reprodukčne správajúcich žien z marginalizovaných rómskych komunít. Ich zvyšujúca sa váha v populácii by mohla tiež vysvetľovať súčasnú rastúcu intenzitu a podiel plodnosti žien vo veku do 20 rokov.

RESUME

The aim of the paper was to analyze the development and proportion of women's fertility in the young (up to 20 years) and older (35 and over) age with a focus on advanced and very advanced age. The results are set in a broader concept of transformation of the ongoing reproductive behavior in the demographic revolution, after World War II during a specific period of the previous political regime, and after 1989 in the context of societal changes and the post-transitional fertility. In addition, we have also analyzed spatial differences (at district level) and changes in the intensity and proportion of the young and older women's fertility.

The obtained results show a long-term decline in the fertility of women in older age and conversely, strengthening of the intensity and weight in young age. However, this scenario has been changing considerably after 1989 and the process of postponing reproduction and the associated increase in the intensity and proportion of fertility at an older age have come to the forefront. Advanced and very advanced

reproduction age are used in Slovakia only to a limited extent. This also applies to the regional level. On the other hand, there are relatively significant regional differences in fertility at a young age. In particular, they appear to be associated mainly with the existence of specifically reproductive women from marginalized Roma communities. Their increasing weight in the population could also explain the current growing fertility rate of women under the age of 20.

PROFESIJNÝ ŽIVOTOPIS

RNDr. Branislav Šprocha, PhD., absolvoval magisterské štúdium na Prírodovedeckej fakulte Univerzity Karlovej v Prahe v odbore demografia a geodemografia (2006). V roku 2011 ukončil doktorandské štúdium v programe demografia. Od roku 2007 je vedeckovýskumným pracovníkom Výskumného demografického centra pri INFOSTAT-e a od roku 2009 vedeckým pracovníkom Prognostického ústavu Centra spoločenských a psychologických vied SAV. V roku 2015 sa stal vedúcim Výskumného demografického centra. V oblasti demografie sa špecializuje na problematiku rodinného a reprodukčného správania a ich vplyvov na spoločnosť. Okrem toho sa zameriava na analýzu vybraných populačných štruktúr, reprodukčného správania rómskeho obyvateľstva na Slovensku a otázky konštrukcie populačných prognóz.

KONTAKT

branislav.sprocha@gmail.com

Diana BÍLKOVÁ

Fakulta informatiky a statistiky, Vysoká škola ekonomická v Praze

ČTYŘPARAMETRICKÉ A TŘÍPARAMETRICKÉ LOGNORMÁLNÍ KŘIVKY JAKO MODELY MZDOVÝCH ROZDĚLENÍ

FOUR-PARAMETRIC AND THREE-PARAMETRIC LOGNORMAL CURVES AS WAGE DISTRIBUTION MODELS

ABSTRAKT

Tento článek se zabývá konstrukcí modelů rozdělení mezd pomocí čtyřparametrových a tříparametrových lognormálních křivek. Hlavním cílem výzkumu je porovnat přesnost získaných modelů s využitím obou typů lognormálního rozdělení jako modelů rozdělení mezd. Minimální mzda v daném roce představuje počátek čtyřparametrických a tříparametrických lognormálních křivek. Odhady zbývajících tří, respektive dvou parametrů, jsou konstruovány pomocí kvantilové metody. Testové kritérium chí-kvadrát je použito k vyhodnocení přesnosti získaných modelů. Téměř ve všech případech mzdových modelů poskytly čtyřparametrové lognormální křivky přesnější výsledky než tříparametrové lognormální křivky. Výsledky z hlediska přesnosti obou typů lognormálních křivek jsou stejně přesné pouze v několika případech. Rozdíly v přesnosti čtyřparametrových a tříparametrových mzdových modelů však nejsou kritické.

ABSTRACT

This paper deals with the construction of wage distribution models using four-parameter and three-parameter lognormal curves. The main objective of the research is to compare the accuracy of the models obtained using both types of lognormal distribution as wage distribution models. The minimum wage in a given year represents the beginning of four-parameter and three-parameter lognormal curves. The estimates of the remaining three or two parameters are constructed using the quantile method. The chi-square testing criterion is used for the evaluation of the accuracy of the models obtained. In almost all cases of wage models, the four-parameter lognormal curves yielded more accurate results than the three-parameter lognormal curves. The results in terms of accuracy of both types of lognormal curves are equally accurate only in a few cases. However, the differences in the accuracy of the four-parameter and three-parameter wage models are not critical.

KLÍČOVÁ SLOVA

modely mzdových rozdělení, čtyřparametrické lognormální křivky, tříparametrické lognormální křivky, kvantilová metoda bodového odhadu parametrů, testové kritérium chí-kvadrát, přesnost bodového odhadu parametrů

KEY WORDS

wage distribution models, four-parameter lognormal curves, three-parameter lognormal curves, quantile method of point parameter estimation, chi-square testing criterion, accuracy of point parameter estimation

1. ÚVOD

Lognormální rozdělení je po mnoho desetiletí nástrojem pro konstrukci teoretických modelů rozmanitých analyzovaných proměnných v různých technických,

fyzikálních nebo ekonomických oblastech. Například [4] používá toto teoretické pravděpodobnostní rozdělení v oblasti kontroly jakosti nebo dvojice autorů [3] již využívá toto pravděpodobnostní rozdělení při modelování personálních příjmů. Toto pravděpodobnostní rozdělení má různé modifikace od dvouparametrické až po čtyřparametrickou alternativu. Příkladem použití čtyřparametrické alternativy tohoto pravděpodobnostního rozdělení je studie [6], kde autoři používají čtyřparametrickou verzi tohoto rozdělení ve spojení s modelováním rozdělení atmosférických aerosolových částic. Již v polovině šedesátých let minulého století používá [5] čtyřparametrickou verzi tohoto pravděpodobnostního rozdělení v kontextu ztrát z rozsahu. O něco později [2] provádí výzkum problematiky odhadování parametrů čtyřparametrické verze lognormálního rozdělení.

Tento příspěvek se zabývá problematikou odhadu parametrů čtyřparametrického lognormálního rozdělení kvantilovou metodou v souvislosti s konstrukcí modelů rozdělení mezd. Definice a popis čtyřparametrického a tříparametrického lognormálního rozdělení byly čerpány ze zdrojů [1] a [7], konstrukce odhadů parametrů kvantilovou metodou je dosti podrobná a vychází z vlastního odvození.

Hlavním cílem této studie je srovnání přesnosti modelů mzdových rozdělení konstruovaných s využitím čtyřparametrických a tříparametrických lognormálních křivek. Pro získání odhadů parametrů byla použita v obou případech kvantilová metoda odhadu parametrů. V obou případech lognormálního rozdělení počátek lognormálních křivek představovala minimální mzda v příslušném roce. Testové kritérium chí-kvadrát bylo použito při vyhodnocení přesnosti získaných modelů. Hlavní vědecká hypotéza spočívá v tvrzení, že při použití kvantilové metody odhadování parametrů vedou čtyřparametrické lognormální modely k přesnějším výsledkům než tříparametrické lognormální modely.

Tabulka č. 1: Rozsahy výběrových souborů zaměstnanců podle velikosti jednotky (tisíce zaměstnanců)

Velikost jednotky	Rok				
	2014	2015	2016	2017	2018
Méně než 10 zaměstnanců	527,5	571,4	570,5	576,5	593,0
Od 10 do 49 zaměstnanců	738,0	693,9	702,0	697,2	674,5
Od 50 do 249 zaměstnanců	795,7	811,7	830,5	842,0	866,4
Od 250 do 999 zaměstnanců	657,6	677,2	697,6	714,8	725,5
Od 1 000 do 4 999 zaměstnanců	657,6	1 340,3	494,8	518,3	525,6
Od 5 000 zaměstnanců	305,6	298,7	409,6	320,4	329,4

Zdroj: www.czso.cz

Tabulka č. 2: Vývoj minimální mzdy (v Kč) v období let 2014 – 2018

Rok	2014	2015	2016	2017	2018
Minimální mzda	8 500	9 200	9 900	11 000	12 200

Zdroj: www.mpsv.cz

Data pro tento výzkum a rozsahy výběrových souborů pocházejí z oficiálních webových stránek Českého statistického úřadu (ČSÚ). Data zahrnují mzdy zaměstnanců České republiky, hrubá měsíční nominální mzda v Kč představuje hlavní zkoumanou proměnnou. Odpovídající data ČSÚ byla ve formě intervalového rozdělení četností s krajními otevřenými intervaly, individuální data nejsou běžně

dostupná. Analyzovaná data byla roztržena podle velikosti jednotky a pokrývala období od roku 2014 do roku 2018, viz tabulka č. 1. Tabulka č. 2 představuje vývoj minimální mzdy v příslušném období. Výběrové kvantily použité v rámci kvantilové metody odhadu parametrů byly odhadnuty z příslušného intervalového rozdělení četností. Celkově bylo takto zkoumáno 30 mzdových rozdělení.

Data zahrnují odměnu za vykonanou práci zaměstnanců jak v podnikatelském, tak nepodnikatelském sektoru ekonomiky. Mzda se vyplácí zaměstnanci za vykonanou práci v soukromé (podnikatelské) sféře, zatímco plat se vydělává ve sféře státní (nepodnikatelské). V této studii, stejně jako data prezentovaná ČSÚ, jsou pod termín mzda zahrnuty mzdy v podnikatelské sféře i platy v nepodnikatelské sféře.

2. TEORIE A METODY

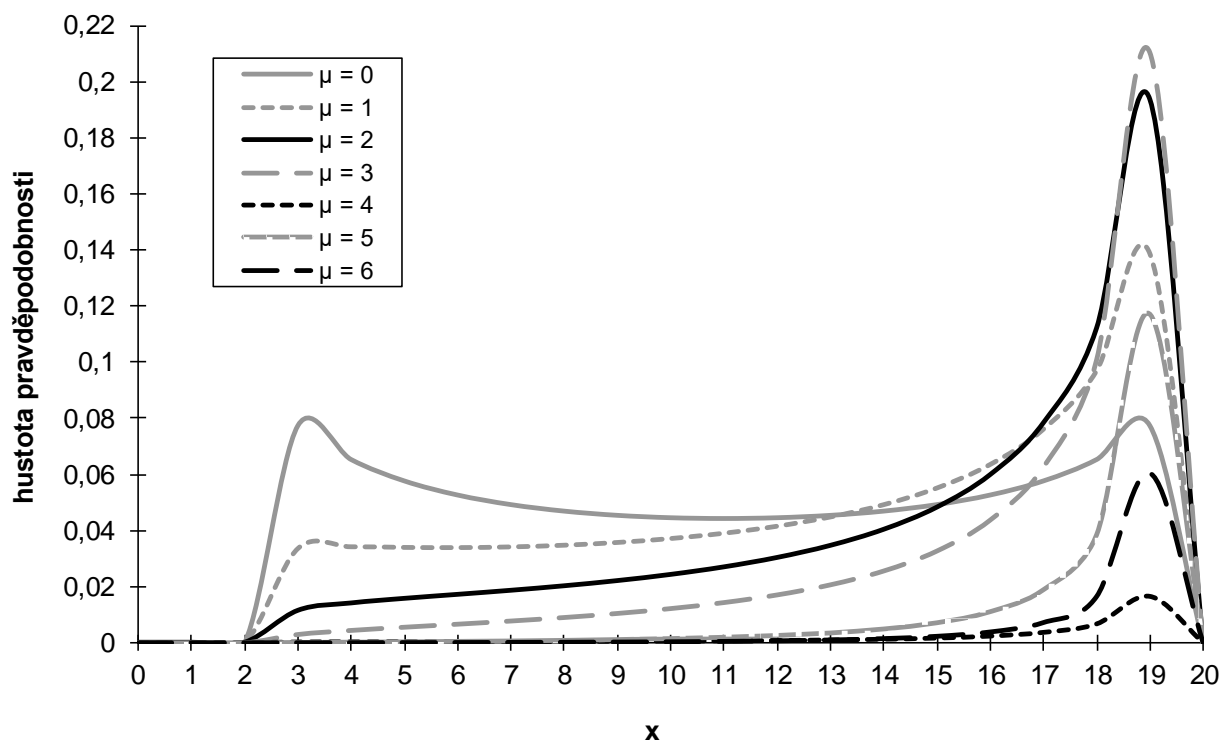
2.1 Čtyřparametrické lognormální rozdělení

Náhodná veličina X má čtyřparametrické lognormální rozdělení s parametry μ , σ^2 , θ a τ , kde $-\infty < \mu < \infty$, $\sigma^2 > 0$, $-\infty < \theta < \tau < \infty$, jestliže její hustota pravděpodobnosti má tvar:

$$f(x; \mu, \sigma^2, \theta, \tau) = \frac{(\tau - \theta)}{\sigma \cdot (x - \theta) \cdot (\tau - x) \cdot \sqrt{2\pi}} \cdot \exp \left[-\frac{\left(\ln \frac{x - \theta}{\tau - x} - \mu \right)^2}{2\sigma^2} \right], \quad \theta < x < \tau, \quad (1)$$

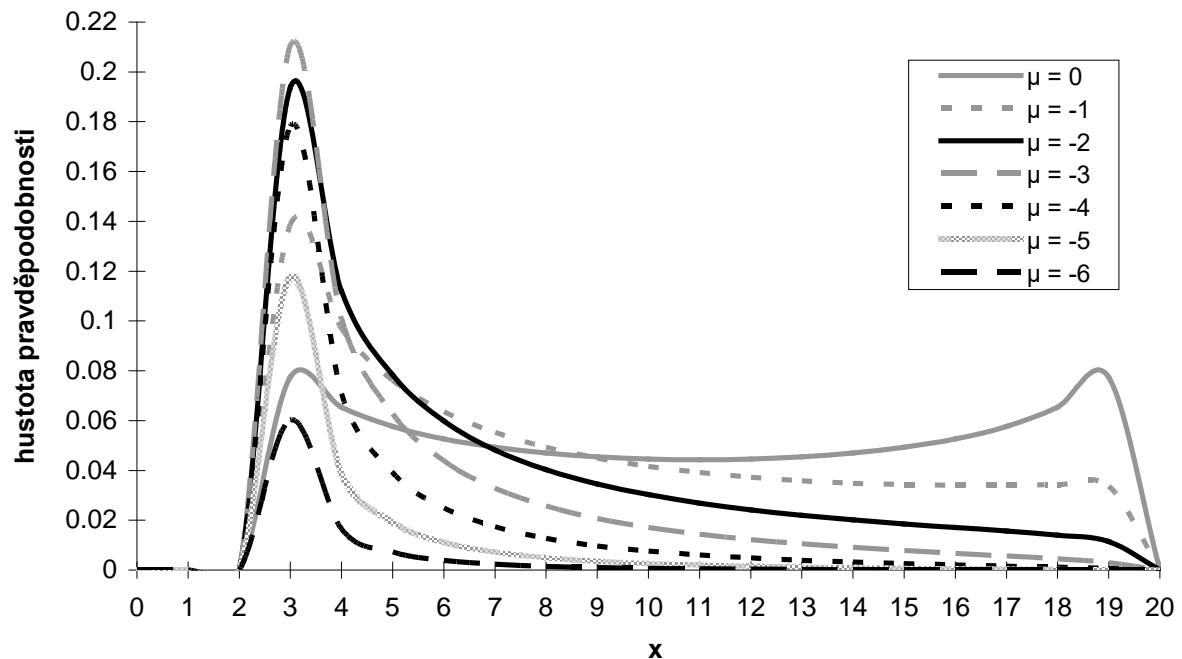
$$= 0, \quad \text{jinak.}$$

Obrázek č. 1: Hustota pravděpodobnosti čtyřparametrického lognormálního rozdělení pro hodnoty parametrů $\sigma = 2$ ($\sigma^2 = 4$); $\theta = 2$; $\tau = 20$



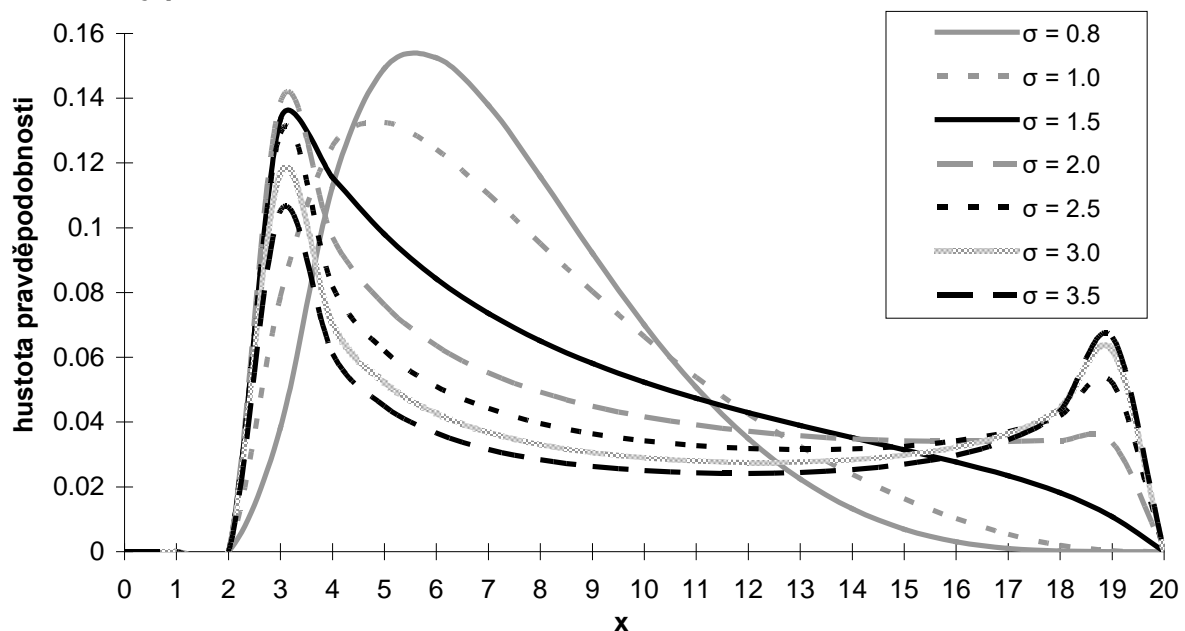
Zdroj: vlastní konstrukce

Obrázek č. 2: Hustota pravděpodobnosti čtyřparametrického lognormálního rozdělení pro hodnoty parametrů $\sigma = 2$ ($\sigma^2 = 4$); $\theta = 2$; $\tau = 20$



Zdroj: vlastní konstrukce

Obrázek č. 3: Hustota pravděpodobnosti čtyřparametrického lognormálního rozdělení pro hodnoty parametrů $\mu = -1$; $\theta = 2$; $\tau = 20$



Zdroj: vlastní konstrukce

Lognormální rozdělení s parametry μ , σ^2 , θ a τ budeme značit $LN(\mu, \sigma^2, \theta, \tau)$. Funkce hustoty pravděpodobnosti čtyřparametrického lognormálního rozdělení $LN(\mu, \sigma^2, \theta, \tau)$ může mít velmi odlišné tvary v závislosti na hodnotách parametrů tohoto rozdělení. Rozdělení může být rovněž dvouvrcholové pro $\sigma^2 > 2$ a $|\mu| < \sigma^2 \cdot \sqrt{(1 - 2/\sigma^2)} - 2 \tanh^{-1} \sqrt{(1 - 2/\sigma^2)}$. Obrázky č. 1 – 3 představují ilustrační

příklady tvarů hustoty pravděpodobnosti čtyřparametrického lognormálního rozdělení v závislosti na hodnotách parametrů.

Jestliže má náhodná veličina X čtyřparametrické lognormální rozdělení s parametry μ , σ^2 , θ a τ , potom náhodná veličina:

$$Y = \ln \frac{X - \theta}{\tau - X} \quad (2)$$

má normální rozdělení s parametry μ a σ^2 a náhodná veličina:

$$U = \frac{\ln \frac{X - \theta}{\tau - X} - \mu}{\sigma} \quad (3)$$

má normované normální rozdělení.

Parametr μ tedy představuje střední hodnotu náhodné veličiny (2) a parametr σ^2 je rozptyl této náhodné veličiny. Parametr θ je počátek lognormálního rozdělení (teoretické minimum) a parametr τ představuje koncový bod lognormálního rozdělení (teoretické maximum).

2.2 Tříparametrické lognormální rozdělení

Náhodná veličina X má tříparametrické lognormální rozdělení s parametry μ , σ^2 a θ , kde $-\infty < \mu < \infty$, $\sigma^2 > 0$, $-\infty < \theta < \infty$, jestliže její hustota pravděpodobnosti má tvar:

$$f(x; \mu, \sigma^2, \theta) = \frac{1}{\sigma \cdot (x - \theta) \cdot \sqrt{2\pi}} \cdot \exp \left[-\frac{[\ln(x - \theta) - \mu]^2}{2\sigma^2} \right], \quad x > \theta, \quad (4)$$

$$= 0, \quad \text{jinak.}$$

Lognormální rozdělení s parametry μ , σ^2 a θ budeme značit $LN(\mu, \sigma^2, \theta)$. Funkce hustoty pravděpodobnosti tříparametrického lognormálního rozdělení $LN(\mu, \sigma^2, \theta)$ je vždy kladně zešikmená. Obrázky č. 4 – 6 představují ilustrační příklady tvarů hustoty pravděpodobnosti tříparametrického lognormálního rozdělení v závislosti na hodnotách parametrů. Tyto obrázky kopírují hodnoty parametrů vždy příslušného obrázku č. 1 – 3, přičemž vynechávají parametr τ .

Jestliže má náhodná veličina X tříparametrické lognormální rozdělení s parametry μ , σ^2 a θ , potom náhodná veličina:

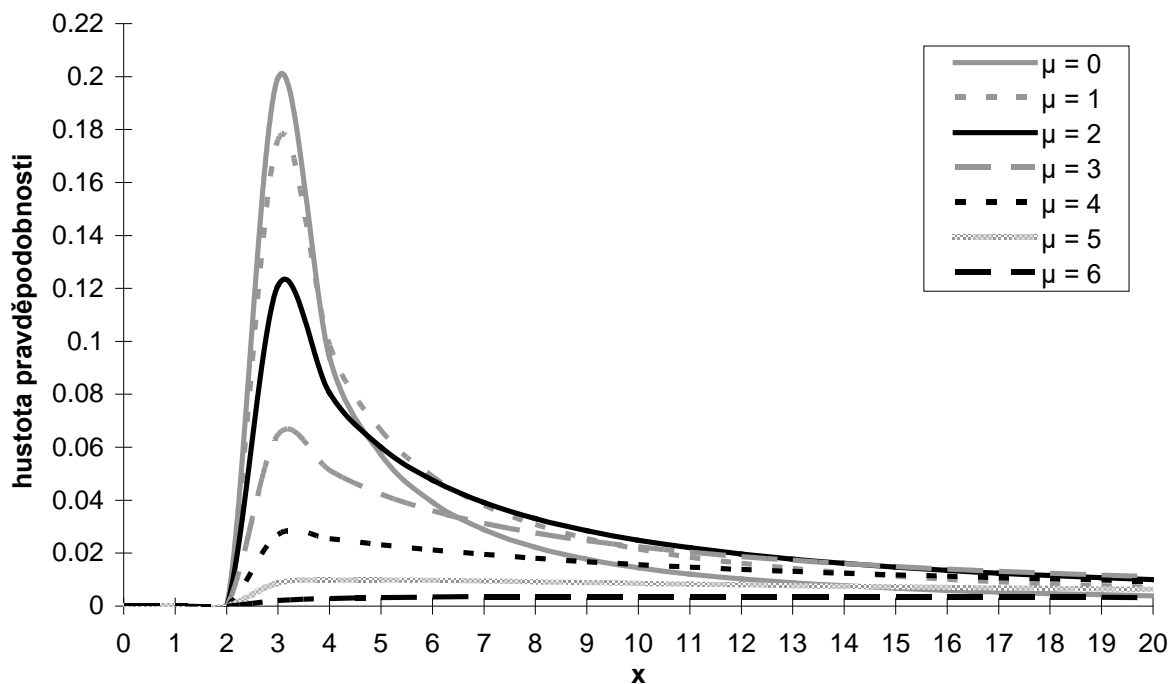
$$Y = \ln(X - \theta) \quad (5)$$

má normální rozdělení s parametry μ a σ^2 a náhodná veličina:

$$U = \frac{\ln(X - \theta) - \mu}{\sigma} \quad (6)$$

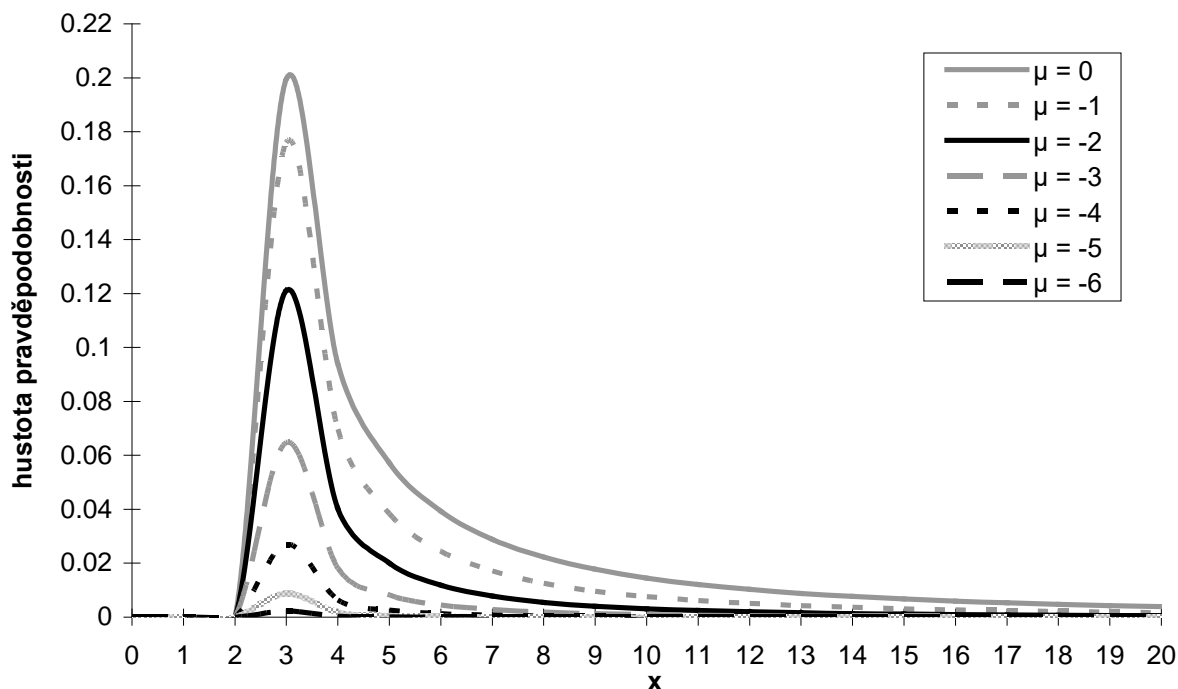
má normované normální rozdělení.

Obrázek č. 4: Hustota pravděpodobnosti tříparametrického lognormálního rozdělení pro hodnoty parametrů $\sigma = 2$ ($\sigma^2 = 4$); $\theta = 2$



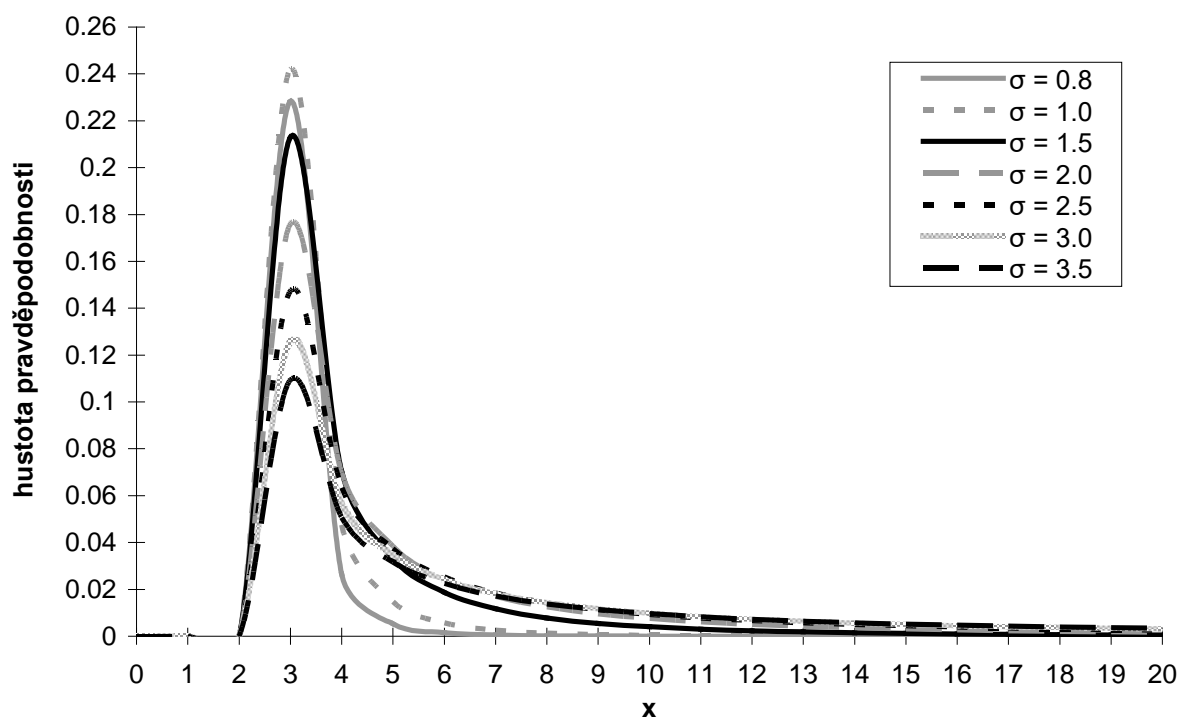
Zdroj: vlastní konstrukce

Obrázek č. 5: Hustota pravděpodobnosti tříparametrického lognormálního rozdělení pro hodnoty parametrů $\sigma = 2$ ($\sigma^2 = 4$); $\theta = 2$



Zdroj: vlastní konstrukce

Obrázek č. 6: Hustota pravděpodobnosti tříparametrického lognormálního rozdělení pro hodnoty parametrů $\mu = -1$; $\theta = 2$



Zdroj: vlastní konstrukce

Parametr μ tedy představuje střední hodnotu náhodné veličiny (5) a parametr σ^2 je rozptyl této náhodné veličiny. Parametr θ je počátek lognormálního rozdělení (teoretické minimum).

2.3 Kvantilová metoda bodového odhadu parametrů pro čtyřparametrické lognormální rozdělení

Předpokládáme, že známe parametr θ z důvodu existence institutu minimální mzdy. Jak již bylo uvedeno, náhodná veličina (2) má normální rozdělení s parametry μ a σ^2 . $100 \cdot P\%$ kvantil normálního rozdělení s parametry μ a σ^2 má tvar:

$$x_p = \mu + \sigma u_p, \quad (7)$$

kde u_p je $100 \cdot P\%$ kvantil normovaného normálního rozdělení. Pro medián normovaného normálního rozdělení platí, že $u_{0,50} = 0$.

Celkem odhadujeme tři parametry čtyřparametrového lognormálního rozdělení (parametr θ je známý jako minimální mzda), takže potřebujeme soustavu tří kvantilových rovnic. Nechť $x_{0,25}$ představuje dolní kvartil, $x_{0,50}$ představuje medián a $x_{0,75}$ představuje horní kvartil čtyřparametrického lognormálního rozdělení, $y_{0,25}$ je dolní kvartil, $y_{0,50}$ je medián a $y_{0,75}$ je horní kvartil normálního rozdělení s parametry μ a σ^2 . Teoretické kvantily lognormálního rozdělení odhadneme odpovídajícími výběrovými kvantily. Analogicky, $u_{0,25}$ je dolní kvartil, $u_{0,50}$ je medián a $u_{0,75}$ je horní kvartil normovaného normálního rozdělení. Pak s ohledem na symetrii hustoty pravděpodobnosti normovaného normálního rozdělení podle nuly platí vztah $u_{0,25} = -u_{0,75}$. Získáváme soustavu tří kvantilových rovnic:

$$y_{0,25} = \ln \frac{x_{0,25} - \theta}{\hat{\tau} - x_{0,25}} = \hat{\mu} + \hat{\sigma} u_{0,25} = \hat{\mu} - \hat{\sigma} u_{0,75}, \quad (8)$$

$$y_{0,50} = \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} = \hat{\mu} + \hat{\sigma} u_{0,50} = \hat{\mu}, \quad (9)$$

$$y_{0,75} = \ln \frac{x_{0,75} - \theta}{\hat{\tau} - x_{0,75}} = \hat{\mu} + \hat{\sigma} u_{0,75}. \quad (10)$$

Rovnici (9) vložíme do rovnic (8) a (10), čímž se soustava tří rovnic zredukuje na soustavu rovnic dvou. Získáváme:

$$\ln \frac{x_{0,25} - \theta}{\hat{\tau} - x_{0,25}} = \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} - \hat{\sigma} u_{0,75}, \quad (11)$$

$$\ln \frac{x_{0,75} - \theta}{\hat{\tau} - x_{0,75}} = \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} + \hat{\sigma} u_{0,75}. \quad (12)$$

Úpravou získáváme:

$$\hat{\sigma} = \frac{\ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} - \ln \frac{x_{0,25} - \theta}{\hat{\tau} - x_{0,25}}}{u_{0,75}}, \quad (13)$$

$$\hat{\sigma} = \frac{\ln \frac{x_{0,75} - \theta}{\hat{\tau} - x_{0,75}} - \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}}}{u_{0,75}}. \quad (14)$$

Získáváme jednu rovnici s jedním neznámým parametrem τ :

$$\hat{\sigma} = \frac{\ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} - \ln \frac{x_{0,25} - \theta}{\hat{\tau} - x_{0,25}}}{u_{0,75}} = \frac{\ln \frac{x_{0,75} - \theta}{\hat{\tau} - x_{0,75}} - \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}}}{u_{0,75}}, \quad (15)$$

takže:

$$\ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} - \ln \frac{x_{0,25} - \theta}{\hat{\tau} - x_{0,25}} = \ln \frac{x_{0,75} - \theta}{\hat{\tau} - x_{0,75}} - \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}}. \quad (16)$$

Tudíž:

$$\begin{aligned} & \ln(x_{0,50} - \theta) - \ln(\hat{\tau} - x_{0,50}) - \ln(x_{0,25} - \theta) + \ln(\hat{\tau} - x_{0,25}) = \\ & = \ln(x_{0,75} - \theta) - \ln(\hat{\tau} - x_{0,75}) - \ln(x_{0,50} - \theta) + \ln(\hat{\tau} - x_{0,50}), \end{aligned} \quad (17)$$

$$\ln \frac{(\hat{\tau} - x_{0,25})(\hat{\tau} - x_{0,75})}{(\hat{\tau} - x_{0,50})(\hat{\tau} - x_{0,50})} = \ln \frac{(x_{0,25} - \theta)(x_{0,75} - \theta)}{(x_{0,50} - \theta)(x_{0,50} - \theta)}, \quad (18)$$

$$\frac{(\hat{\tau} - x_{0,25})(\hat{\tau} - x_{0,75})}{(\hat{\tau} - x_{0,50})^2} = \frac{(x_{0,25} - \theta)(x_{0,75} - \theta)}{(x_{0,50} - \theta)^2}. \quad (19)$$

Spočítáme konstantu C:

$$C = \frac{(x_{0,25} - \theta)(x_{0,75} - \theta)}{(x_{0,50} - \theta)^2}, \quad (20)$$

takže:

$$\frac{(\hat{\tau} - x_{0,25})(\hat{\tau} - x_{0,75})}{(\hat{\tau} - x_{0,50})^2} = C. \quad (21)$$

Získáváme kvadratickou rovnici:

$$(1 - C)\hat{\tau}^2 + (2Cx_{0,50} - x_{0,25} - x_{0,75})\hat{\tau} - Cx_{0,50}^2 + x_{0,25} \cdot x_{0,75} = 0. \quad (22)$$

Získáváme kvantilový odhad parametru τ :

$$\hat{\tau} = \frac{-(2Cx_{0,50} - x_{0,25} - x_{0,75}) + \sqrt{(2Cx_{0,50} - x_{0,25} - x_{0,75})^2 + 4(1 - C)(Cx_{0,50}^2 - x_{0,25} \cdot x_{0,75})}}{2(1 - C)}. \quad (23)$$

Kvantilové odhady všech tří neznámých parametrů čtyřparametrického lognormálního rozdělení mají tedy tvar:

$$\hat{\tau} = \frac{-(2Cx_{0,50} - x_{0,25} - x_{0,75}) + \sqrt{(2Cx_{0,50} - x_{0,25} - x_{0,75})^2 + 4(1 - C)(Cx_{0,50}^2 - x_{0,25} \cdot x_{0,75})}}{2(1 - C)}, \quad (24)$$

$$\hat{\sigma} = \frac{\ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}} - \ln \frac{x_{0,25} - \theta}{\hat{\tau} - x_{0,25}}}{u_{0,75}} = \frac{\ln \frac{x_{0,75} - \theta}{\hat{\tau} - x_{0,75}} - \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}}}{u_{0,75}}, \quad (25)$$

$$\hat{\mu} = \ln \frac{x_{0,50} - \theta}{\hat{\tau} - x_{0,50}}, \quad (26)$$

kde konstanta C je určena vztahem (17).

2.4 Kvantilová metoda bodového odhadu parametrů pro tříparametrické lognormální rozdělení

Kvantilový odhad parametrů v případě tříparametrického lognormálního rozdělení je výrazně jednodušší. Opět předpokládáme, že parametr θ známe z důvodu existence institutu minimální mzdy.

Jak je uvedeno výše, náhodná veličina (5) má normální rozdělení s parametry μ a σ^2 . Získáváme soustavu dvou kvantilových rovnic:

$$y_{0,50} = \ln(x_{0,50} - \theta) = \hat{\mu} + \hat{\sigma}u_{0,50} = \hat{\mu}, \quad (27)$$

$$y_{0,75} = \ln(x_{0,75} - \theta) = \hat{\mu} + \hat{\sigma}u_{0,75}. \quad (28)$$

Rovnici (24) vložíme do rovnice (25), čímž se systém dvou rovnic zredukuje na jednu rovnici. Získáváme rovnici:

$$\ln(x_{0,75} - \theta) = \ln(x_{0,50} - \theta) + \hat{\sigma}u_{0,75}. \quad (29)$$

Tudíž:

$$\hat{\sigma} = \frac{\ln \frac{x_{0,75} - \theta}{x_{0,50} - \theta}}{u_{0,75}}. \quad (30)$$

Kvantilové odhady obou neznámých parametrů tříparametrického lognormálního rozdělení mají tedy tvar:

$$\hat{\sigma} = \frac{\ln \frac{x_{0,75} - \theta}{x_{0,50} - \theta}}{u_{0,75}}, \quad (31)$$

$$\hat{\mu} = \ln(x_{0,50} - \theta). \quad (32)$$

3. VÝSLEDKY

Tabulka č. 3 uvádí výběrové kvantily mzdových rozdělení použité k odhadu parametrů čtyřparametrických lognormálních křivek pro mzdové intervaly o šířce 5 000 Kč, přičemž počátek těchto křivek představuje minimální mzda v daném roce. Výběrové horní kvantily a mediány z tabulky č. 3 jsou potom rovněž použity k odhadu parametrů tříparametrických lognormálních křivek pro mzdové intervaly opět se šířkou 5 000 Kč, přičemž minimální mzda v daném roce představuje rovněž počátek těchto křivek. Tabulka č. 3 ukazuje nejnižší úroveň mezd pro nejmenší podniky s méně než 10 zaměstnanci. Úroveň mezd se pak zvyšuje s velikostí společnosti až na 1 000 – 4 999 zaměstnanců, kde dosahuje svého vrcholu. Úroveň mzdy se s růstem počtu zaměstnanců pro podniky s 5 000 a více zaměstnanci dále již zásadně nemění.

Tabulka č. 4 představuje odhady parametrů pro čtyřparametrové lognormální křivky představující modely mzdových rozdělení a tabulka č. 5 zobrazuje odhady parametrů pro tříparametrové lognormální křivky jako modely mzdových rozdělení.

Obrázky č. 7 – 11 představují čtyřparametrové modely mzdových rozdělení v období let 2014 – 2018. Tato obrázky ukazují změnu tvaru modelů rozdělení mezd s růstem velikosti podniku, zatímco tvar modelů rozdělení mezd je jasně odlišný pro nejmenší podniky s méně než 10 zaměstnanci. Obrázky č. 12 – 16 představují tříparametrové modely mzdových rozdělení v období let 2014 – 2018. Z těchto obrázků je rovněž patrný jasný rozdíl ve tvaru modelů rozdělení mezd pro nejmenší podniky s méně než 10 zaměstnanci.

Tabulka č. 3: Výběrové kvartily mzdových rozdělení (v Kč) v období let 2014 – 2018

Rok	Velikost jednotky	Kvartily		
		Dolní kvartil	Medián	Horní kvartil
2014	méně než 10 zaměstnanců	9 385	15 177	20 478
	od 10 do 49 zaměstnanců	15 301	21 659	28 297
	od 50 do 249 zaměstnanců	17 434	23 161	30 411
	od 250 do 999 zaměstnanců	19 005	24 763	32 972
	od 1000 do 4 999 zaměstnanců	21 113	27 647	36 836
	od 5 000 zaměstnanců	21 120	27 285	36 246
2015	méně než 10 zaměstnanců	9 683	15 493	21 268
	od 10 do 49 zaměstnanců	15 803	22 630	29 752
	od 50 do 249 zaměstnanců	18 043	23 954	31 335
	od 250 do 999 zaměstnanců	20 025	25 990	34 933
	od 1000 do 4 999 zaměstnanců	22 065	28 887	38 227
	od 5 000 zaměstnanců	21 693	28 419	37 577
2016	méně než 10 zaměstnanců	10 003	15 949	22 286
	od 10 do 49 zaměstnanců	16 698	23 645	30 826
	od 50 do 249 zaměstnanců	19 043	25 235	33 138
	od 250 do 999 zaměstnanců	21 086	27 112	36 506
	od 1000 do 4 999 zaměstnanců	23 325	30 071	39 332
	od 5 000 zaměstnanců	22 910	29 636	38 641
2017	méně než 10 zaměstnanců	12 941	16 696	24 019
	od 10 do 49 zaměstnanců	18 239	25 028	32 481
	od 50 do 249 zaměstnanců	20 664	27 050	35 639
	od 250 do 999 zaměstnanců	22 844	29 179	38 473
	od 1000 do 4 999 zaměstnanců	25 240	32 118	42 794
	od 5 000 zaměstnanců	24 949	32 355	43 814
2018	méně než 10 zaměstnanců	14 412	18 120	24 758
	od 10 do 49 zaměstnanců	19 983	27 442	36 036
	od 50 do 249 zaměstnanců	22 729	29 553	38 126
	od 250 do 999 zaměstnanců	24 847	31 496	41 645
	od 1000 do 4 999 zaměstnanců	27 540	35 281	47 265
	od 5 000 zaměstnanců	27 430	35 923	48 498

Zdroj: vlastní výpočet

Tabulka č. 6 ukazuje vypočtené hodnoty testového kritéria chí-kvadrát pro čtyřparametrové a tříparametrové lognormální křivky. Z této tabulky vidíme, že ve většině případů jsou hodnoty testového kritéria chí-kvadrát pro čtyřparametrové křivky o něco menší než pro křivky se třemi parametry. Ve většině případů však tento rozdíl není příliš zásadní. Hodnota kritéria chí-kvadrát testu je stejná pro čtyřparametrové lognormální křivky a tříparametrové lognormální křivky pouze pro šest z třiceti modelů mzdových rozdělení.

Lze tedy konstatovat, že čtyřparametrové lognormální křivky poskytovaly mírně přesnější modely rozdělení mezd než tříparametrové lognormální křivky.

Rozlišujeme $k = 10$ intervalů v případě všech mzdových rozdělení. V případě čtyřparametrového lognormálního rozdělení odhadujeme $p = 3$ parametry a v případě tříparametrového lognormálního rozdělení $p = 2$ parametry. Počátek čtyřparametrických a tříparametrických lognormálních křivek představuje minimální

mzda v daném roce, tedy tento parametr považujeme za známý. Za předpokladu, že platí nulová hypotéza o předpokládaném lognormálním rozdělení, má testové kritérium chí-kvadrát asymptoticky chí-kvadrát rozdělení o $v = k - p - 1$ stupních volnosti. Odtud $v = 6$ v případě čtyřparametrického lognormálního rozdělení a $v = 7$ v případě tříparametrického lognormálního rozdělení. Tabulka č. 7 obsahuje kritické hodnoty pro chí-kvadrát testy při hladině významnosti $\alpha = 0,05$ (0,01 nebo 0,10).

Z tabulek č. 6 a 7 je zřejmé, že test vede k odmítnutí nulové hypotézy předpokládající čtyřparametrické nebo tříparametrické lognormálního rozdělení hrubé měsíční mzdy ve všech třiceti případech mzdových rozdělení a na všech uvažovaných hladinách významnosti (5 %, 1 % a 10 %). Ve všech případech tedy vede test k přijetí alternativní hypotézy, že rozdělení hrubé měsíční mzdy je jiné, než předpokládá nulová hypotéza.

Tabulka č. 4: Odhady parametrů čtyřparametrických lognormálních křivek

Rok	Velikost jednotky	$\hat{\theta}$	Odhad parametrů		
			$\hat{\tau}$	$\hat{\mu}$	$\hat{\sigma}$
2014	méně než 10 zam.	8 500	252 109 171	-10,538881	0,866471
	od 10 do 49 zam.	8 500	1 945 849 366	-11,904105	0,605534
	od 50 do 249 zam.	8 500	5 937 914 943	-12,911655	0,595708
	od 250 do 999 zam.	8 500	22 372 565 586	-14,134463	0,605886
	od 1 000 do 4 999 zam.	8 500	30 923 935 022	-14,294882	0,581146
	od 5 000 zam.	8 500	98 945 504 568	-15,476995	0,578248
2015	méně než 10 zam.	9 200	241 484 561	-10,555067	0,965431
	od 10 do 49 zam.	9 200	1 898 823 320	-11,859251	0,630811
	od 50 do 249 zam.	9 200	5 609 449 752	-12,848465	0,601470
	od 250 do 999 zam.	9 200	59 014 840 613	-15,072512	0,633039
	od 1 000 do 4 999 zam.	9 200	23 129 037 957	-13,976636	0,575665
	od 5 000 zam.	9 200	20 263 998 342	-13,868474	0,577748
2016	méně než 10 zam.	9 900	230 978 158	-10,550185	1,062661
	od 10 do 49 zam.	9 900	2 084 191 626	-11,929221	0,623184
	od 50 do 249 zam.	9 900	6 537 123 265	-12,962867	0,616272
	od 250 do 999 zam.	9 900	15 486 387 102	-13,709881	0,645742
	od 1 000 do 4 999 zam.	9 900	31 820 624 721	-14,271378	0,560186
	od 5 000 zam.	9 900	22 138 140 437	-13,930369	0,557278
2017	méně než 10 zam.	11 000	1 406 155 571	-12,416549	1,225533
	od 10 do 49 zam.	11 000	2 824 251 278	-12,212725	0,631819
	od 50 do 249 zam.	11 000	9 733 682 516	-13,315383	0,635494
	od 250 do 999 zam.	11 000	57 164 360 790	-14,961164	0,612218
	od 1 000 do 4 999 zam.	11 000	70 874 611 882	-15,026309	0,606628
	od 5 000 zam.	11 000	83 974 032 655	-15,184714	0,636859
2018	méně než 10 zam.	12 200	1 721 507 593	-12,580327	1,114949
	od 10 do 49 zam.	12 200	3 573 818 184	-12,365081	0,662954
	od 50 do 249 zam.	12 200	9 266 487 217	-13,188130	0,595232
	od 250 do 999 zam.	12 200	23 899 457 061	-14,029485	0,626610
	od 1 000 do 4 999 zam.	12 200	56 855 677 303	-14,896147	0,626309
	od 5 000 zam.	12 200	75 318 226 791	-14,970803	0,630582

Zdroj: vlastní výpočet

Tabulka č. 5: Odhady parametrů tříparametrických lognormálních křivek

Rok	Velikost jednotky	$\hat{\theta}$	Odhad parametrů	
			$\hat{\mu}$	$\hat{\sigma}$
2014	méně než 10 zaměstnanců	8 500	8,806432	0,866440
	od 10 do 49 zaměstnanců	8 500	9,484848	0,605529
	od 50 do 249 zaměstnanců	8 500	9,592965	0,595707
	od 250 do 999 zaměstnanců	8 500	9,696637	0,605885
	od 1 000 do 4 999 zaměstnanců	8 500	9,859914	0,581146
	od 5 000 zaměstnanců	8 500	9,840840	0,578248
2015	méně než 10 zaměstnanců	9 200	8,747185	0,965396
	od 10 do 49 zaměstnanců	9 200	9,505238	0,630806
	od 50 do 249 zaměstnanců	9 200	9,599250	0,601468
	od 250 do 999 zaměstnanců	9 200	9,728543	0,633039
	od 1 000 do 4 999 zaměstnanců	9 200	9,887718	0,575664
	od 5 000 zaměstnanců	9 200	9,863637	0,577748
2016	méně než 10 zaměstnanců	9 900	8,707580	1,062620
	od 10 do 49 zaměstnanců	9 900	9,528415	0,623179
	od 50 do 249 zaměstnanců	9 900	9,637892	0,616270
	od 250 do 999 zaměstnanců	9 900	9,753345	0,645741
	od 1 000 do 4 999 zaměstnanců	9 900	9,912002	0,560185
	od 5 000 zaměstnanců	9 900	9,890198	0,557278
2017	méně než 10 zaměstnanců	11 000	8,647558	1,225524
	od 10 do 49 zaměstnanců	11 000	9,548775	0,631815
	od 50 do 249 zaměstnanců	11 000	9,683472	0,635493
	od 250 do 999 zaměstnanců	11 000	9,808032	0,612218
	od 1 000 do 4 999 zaměstnanců	11 000	9,957869	0,606628
	od 5 000 zaměstnanců	11 000	9,969059	0,636859
2018	méně než 10 zaměstnanců	12 200	8,686112	1,114938
	od 10 do 49 zaměstnanců	12 200	9,631811	0,662950
	od 50 do 249 zaměstnanců	12 200	9,761537	0,595231
	od 250 do 999 zaměstnanců	12 200	9,867635	0,626609
	od 1 000 do 4 999 zaměstnanců	12 200	10,046775	0,620002
	od 5 000 zaměstnanců	12 200	10,074184	0,630581

Zdroj: vlastní výpočet

Otázka vhodnosti dané křivky pro model rozdělení mzdy v tomto případě není běžným matematicko-statistickým problémem, ve kterém testujeme nulovou hypotézu.

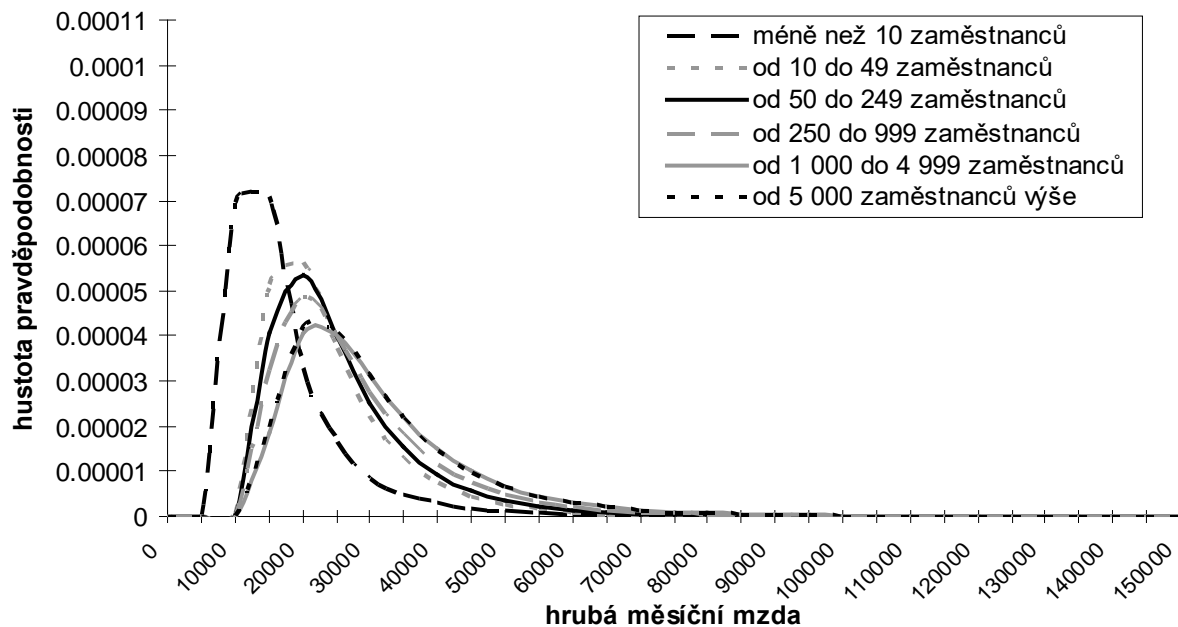
H_0 : Výběr pochází z předpokládaného teoretického lognormálního rozdělení proti alternativní hypotéze:

$$H_1: \text{non } H_0,$$

neboť při testech dobré shody se v případech mzdových rozdělení setkáváme se skutečností, že pracujeme s výběry velkého rozsahu, a proto by test téměř vždy vedl k zamítnutí nulové hypotézy o předpokládaném tvaru rozdělení. Vyplývá to nejen z faktu, že při takto velkých rozsazích výběrů je při zvolené hladině významnosti tak velká síla testu, že test odkryje sebenepatrnější odchylky skutečného mzdového rozdělení a modelu, ale rovněž ze samotného principu konstrukce testu. Malé odchylky nás ale prakticky nezajímají, proto postačí pouze přibližná shoda modelu

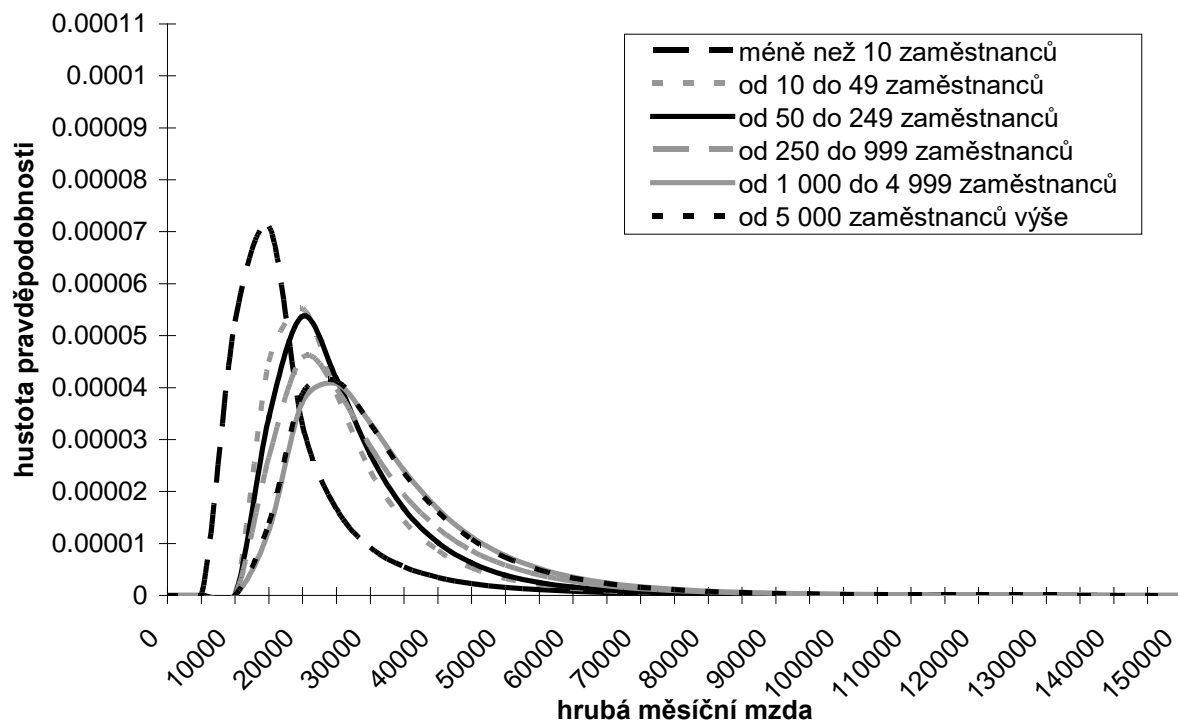
se skutečností a model (křivku) si tzv. vypůjčíme. Testové kritérium chí-kvadrát lze v tomto směru použít pouze orientačně. Při vyhodnocování vhodnosti modelu je třeba postupovat do značné míry subjektivně a opírat se o logický rozbor a zkušenost.

Obrázek č. 7: Čtyřparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2014 (hrubá měsíční mzda v Kč)



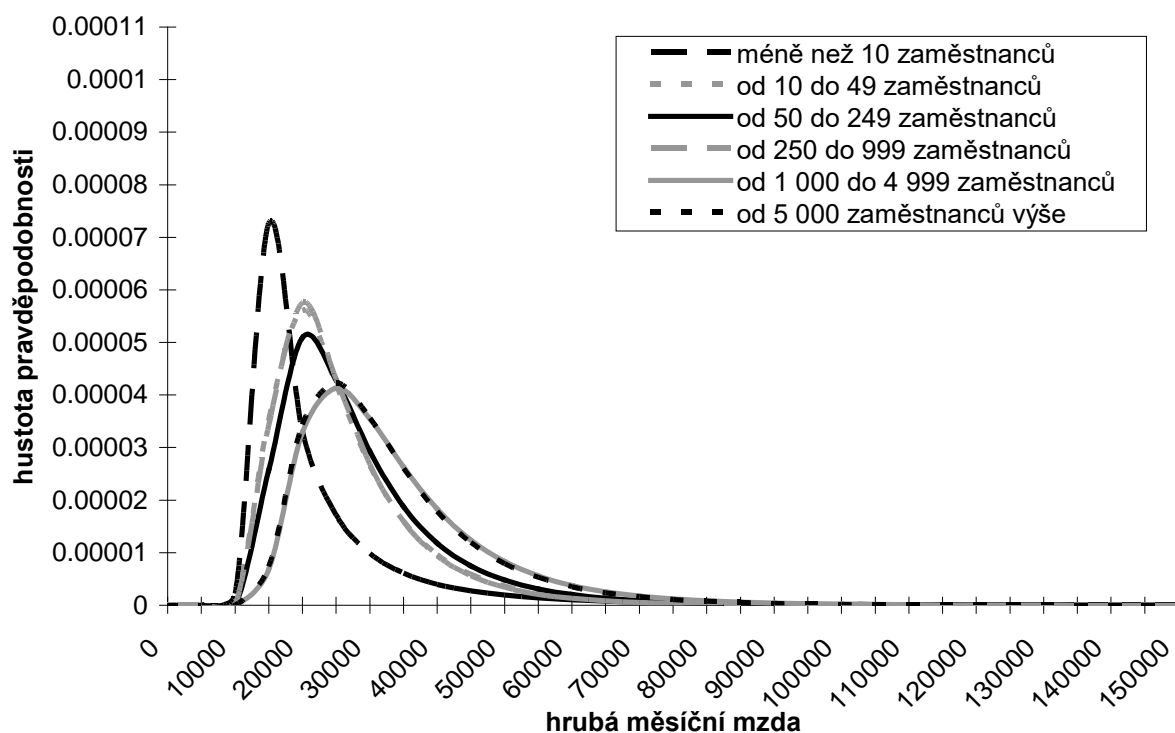
Zdroj: vlastní konstrukce

Obrázek č. 8: Čtyřparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2015 (hrubá měsíční mzda v Kč)



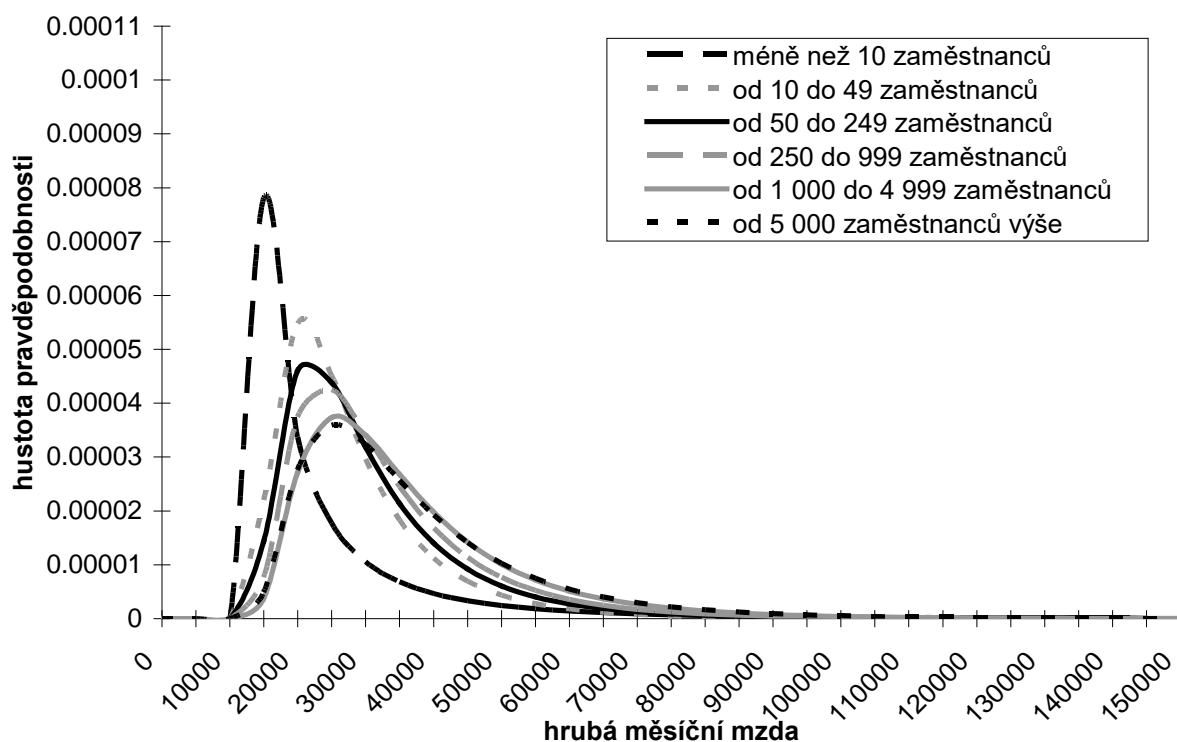
Zdroj: vlastní konstrukce

Obrázek č. 9: Čtyřparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2016 (hrubá měsíční mzda v Kč)



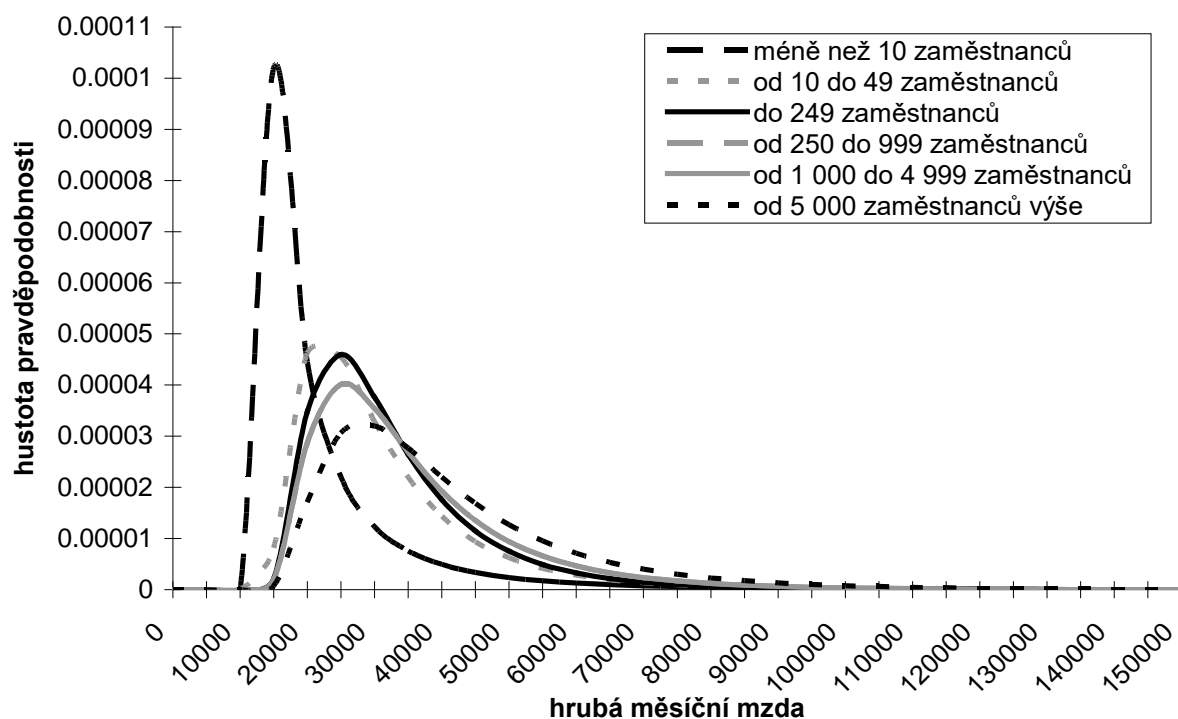
Zdroj: vlastní konstrukce

Obrázek č. 10: Čtyřparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2017 (hrubá měsíční mzda v Kč)



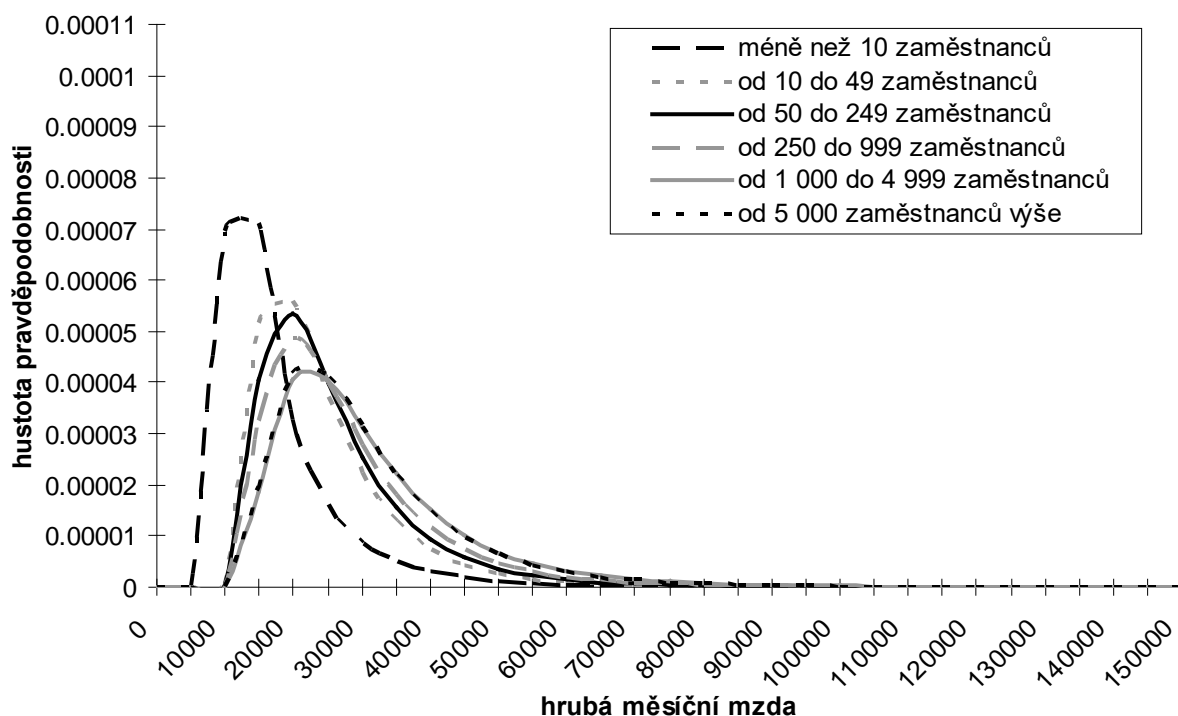
Zdroj: vlastní konstrukce

Obrázek č. 11: Čtyřparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2018 (hrubá měsíční mzda v Kč)



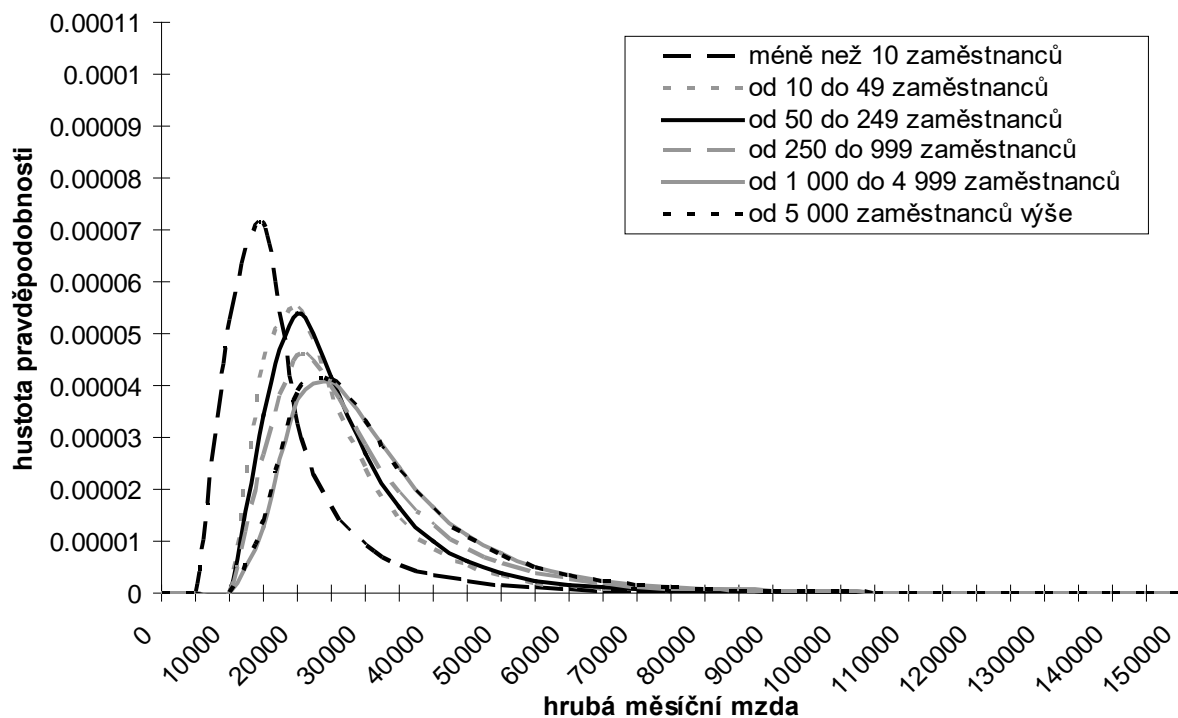
Zdroj: vlastní konstrukce

Obrázek č. 12: Tříparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2014 (hrubá měsíční mzda v Kč)



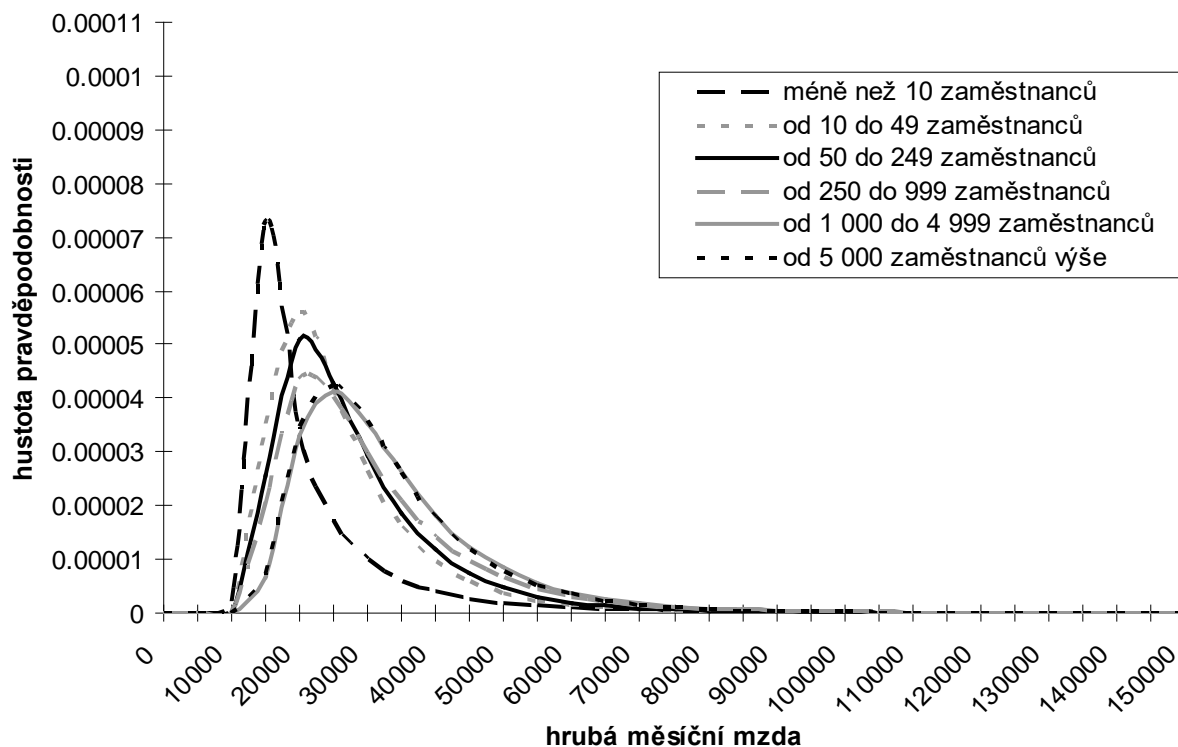
Zdroj: vlastní konstrukce

Obrázek č. 13: Tříparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2015 (hrubá měsíční mzda v Kč)



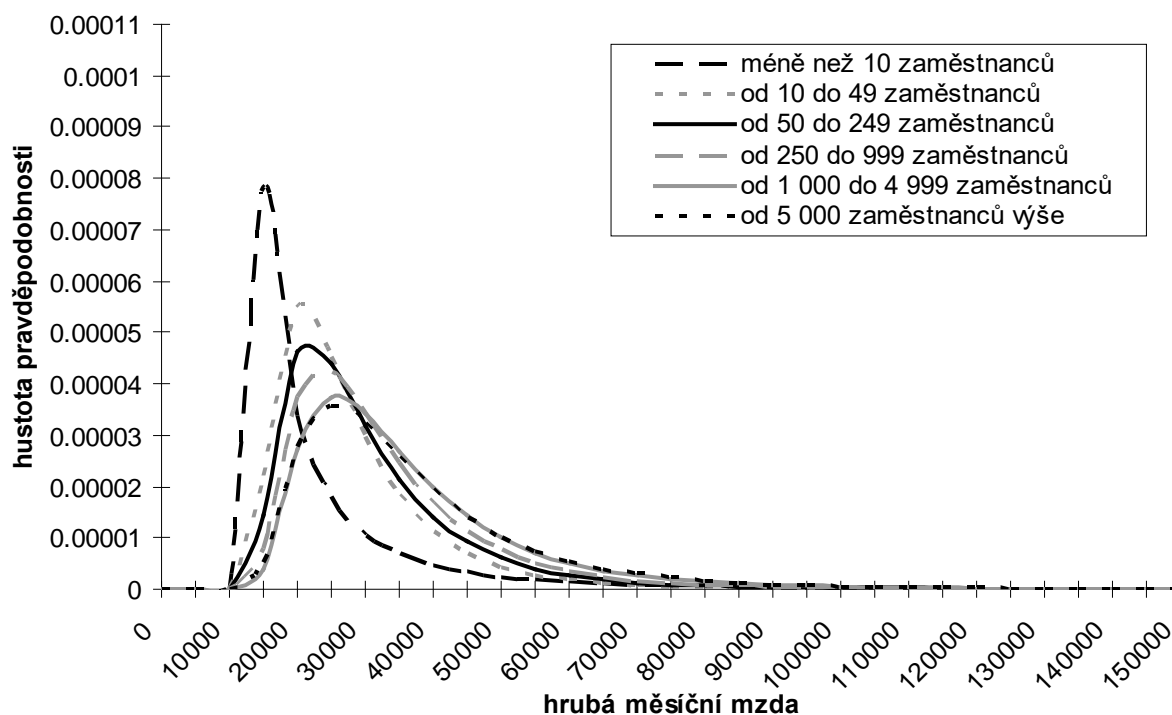
Zdroj: vlastní konstrukce

Obrázek č. 14: Tříparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2016 (hrubá měsíční mzda v Kč)



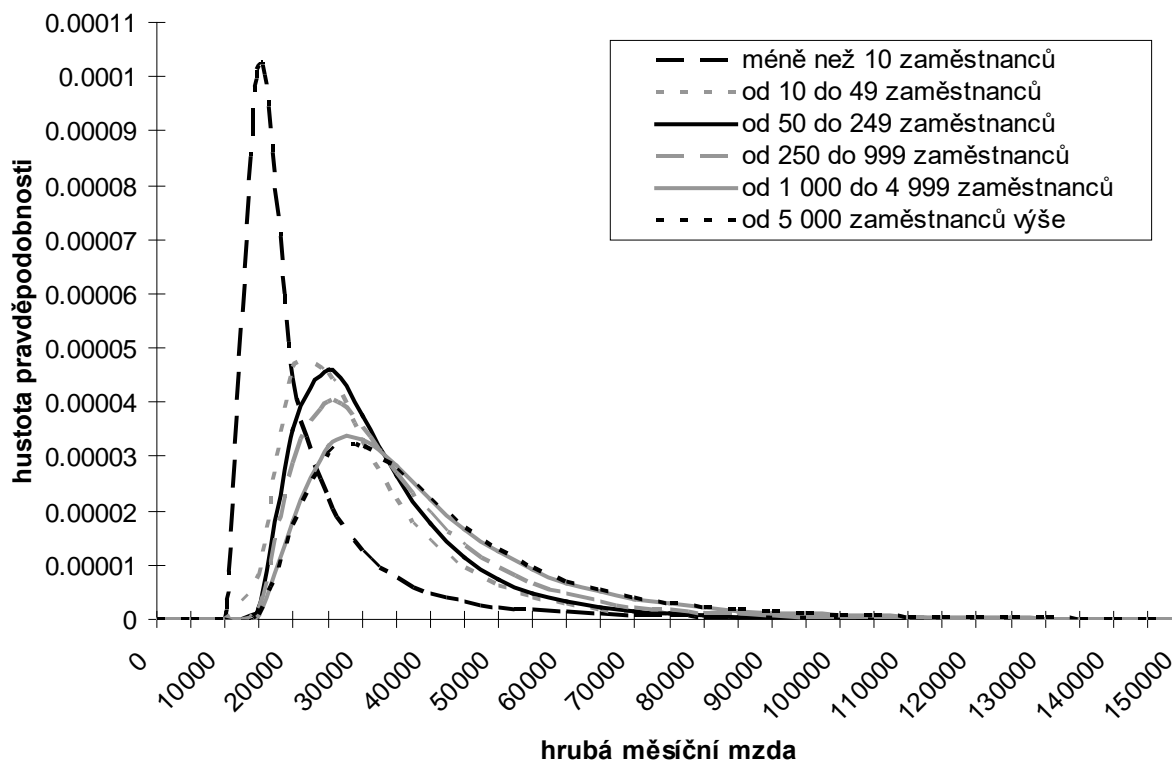
Zdroj: vlastní konstrukce

Obrázek č. 15: Tříparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2017 (hrubá měsíční mzda v Kč)



Zdroj: vlastní konstrukce

Obrázek č. 16: Tříparametrické lognormální modely rozdělení hrubé měsíční mzdy podle velikosti jednotky v roce 2018 (hrubá měsíční mzda v Kč)



Zdroj: vlastní konstrukce

Tabulka č. 6: Hodnoty testového kritéria chí-kvadrát

Rok	Velikost jednotky	Hodnoty testového kritéria	
		Čtyřparametrické rozdělení	Tříparametrické rozdělení
2014	méně než 10 zaměstnanců	52 701	52 706
	od 10 do 49 zaměstnanců	585 678	585 695
	od 50 do 249 zaměstnanců	323 095	323 095
	od 250 do 999 zaměstnanců	201 282	201 284
	od 1 000 do 4 999 zaměstnanců	361 967	361 967
	od 5 000 zaměstnanců	7 000	7 000
2015	méně než 10 zaměstnanců	79 541	79 549
	od 10 do 49 zaměstnanců	1 017 368	1 017 406
	od 50 do 249 zaměstnanců	698 039	698 056
	od 250 do 999 zaměstnanců	307 522	307 522
	od 1 000 do 4 999 zaměstnanců	1 547 992	1 548 024
	od 5 000 zaměstnanců	39 217	39 218
2016	méně než 10 zaměstnanců	125 904	125 917
	od 10 do 49 zaměstnanců	2 874 134	2 874 287
	od 50 do 249 zaměstnanců	1 626 050	1 626 090
	od 250 do 999 zaměstnanců	622 184	622 194
	od 1 000 do 4 999 zaměstnanců	2 818 638	2 818 722
	od 5 000 zaměstnanců	388 759	388 759
2017	méně než 10 zaměstnanců	99 243	99 247
	od 10 do 49 zaměstnanců	64 213 482	64 218 766
	od 50 do 249 zaměstnanců	20 721 610	20 721 977
	od 250 do 999 zaměstnanců	62 788 256	62 788 391
	od 1 000 do 4 999 zaměstnanců	20 025 793	20 025 801
	od 5 000 zaměstnanců	511 089	511 090
2018	méně než 10 zaměstnanců	25 847	25 855
	od 10 do 49 zaměstnanců	869 035	869 061
	od 50 do 249 zaměstnanců	1 144 144	1 144 153
	od 250 do 999 zaměstnanců	401 253	401 258
	od 1 000 do 4 999 zaměstnanců	83 433	83 440
	od 5 000 zaměstnanců	21 705	21 705

Zdroj: vlastní výpočet**Tabulka č. 7: Kritické obory na hladině významnosti $\alpha = 0,05$ (0,01 or 0,10)**

α	Four-parameter lognormal distribution	Three-parameter lognormal distribution
0,05	$W\alpha = \{\chi^2: \chi^2 \geq 12,592\}$	$W\alpha = \{\chi^2: \chi^2 \geq 14,067\}$
0,01	$W\alpha = \{\chi^2: \chi^2 \geq 16,812\}$	$W\alpha = \{\chi^2: \chi^2 \geq 18,475\}$
0,10	$W\alpha = \{\chi^2: \chi^2 \geq 10,645\}$	$W\alpha = \{\chi^2: \chi^2 \geq 12,017\}$

Zdroj: vlastní výpočet

4. ZÁVĚR

Na základě testového kritéria chí-kvadrát poskytly v rámci bodového odhadu parametrů s využitím kvantilové metody čtyřparametrické lognormální křivky poněkud přesnější modely rozdělení mezd než tříparametrické lognormální křivky v naprosté většině mzdových rozdělení. Čtyřparametrické a tříparametrické lognormální křivky přinesly přibližně stejně přesné výsledky pouze v šesti z třiceti mzdových rozdělení.

Tříparametrické lognormální křivky nebyly přesnější než čtyřparametrické lognormální křivky v žádném z 30 případů.

Z teorie i praktických aplikací je všeobecně zřejmé, že přidání dalšího parametru do modelu vede prakticky vždy k vyšší přesnosti modelu. V uvedeném případě čtyřparametrických lognormálních křivek se jedná o horní omezení mzdových modelů, což v tomto případě může být kladným přínosem. Obecně, přidáním dalšího parametru do modelu se však rovněž jedná vždy o složitější model než původní model bez přidaného parametru. Jde tedy o kompromis, zda upřednostníme vyšší přesnost modelu nebo jeho jednodušší verzi. Otázkou tedy zůstává, do jaké míry musí přidání dalšího parametru do modelu zvýšit přesnost tohoto modelu, abychom upřednostnili složitější model. V daném případě mzdových rozdělení nejsou čtyřparametrické lognormální křivky markantně přesnější než křivky tříparametrické, tudíž pro modely mzdových rozdělení postačí zůstat u tříparametrického rozdělení.

Význam lognormálního rozdělení od dvouparametrické až po čtyřparametrickou verzi jako modelu pro výběrová rozdělení nelze zpochybňovat. K tomuto účelu je třeba vždy vybrat některou z metod bodového odhadu parametrů, v rámci této analýzy byla pozornost věnována kvantilové metodě. Výsledky této analýzy lze využít v ekonomické oblasti především při modelování mzdových a příjmových rozdělení nebo jako modely výše škod v rámci neživotního pojištění. Kromě ekonomické oblasti jsou získané výsledky použitelné v celé řadě dalších oblastí počínaje astronomií, přes techniku, přírodní vědy až po sociologii a další obory, kde lognormální rozdělení nachází uplatnění. Charakteristickými vlastnostmi procesu vystihovaného lognormálním modelem je postupné účinkování vzájemně závislých faktorů, tendence k vývoji v geometrické posloupnosti a přerůstání náhodné variability ve variabilitu systematickou, tj. diferenciaci.

Poděkování

Tento příspěvek byl zpracován za podpory prostředků dlouhodobého koncepčního rozvoje vědy a výzkumu číslo IP400040 na Fakultě informatiky a statistiky Vysoké školy ekonomické v Praze.

LITERATURA

- [1] JOHNSON, N. L. – KOTZ, S. – BALAKRISHNAN, N.: Continuous Univariate Distributions. Sec. Ed. New York: Wiley-Interscience, 1994.
- [2] LAMBERT, J. A.: Estimation of Parameters in the Four-Parameter Lognormal Distribution. In: Australian Journal of Statistics, 1970, č. 1, s. 33 – 43.
- [3] McDONALD, J. B. – MANTRALA, A.: The Distribution of Personal Income: Revisited. In: Journal of Applied Econometrics, 1995, č. 2, 201 – 204.
- [4] MORRISON, J.: The Lognormal Distribution in Quality Control. In: Journal of the Royal Statistical Society: Series C, 1958, č. 3, 160 – 172.
- [5] SAVING, T. R.: The Four-Parameter Lognormal, Diseconomies of Scale and the Size Distribution of Manufacturing Establishments. In: International Economic Review, 1965, č. 1, 105 – 114.
- [6] TAMMET, H. – KULMALA, M.: Performance of Four-Parameter Analytical Models of Atmospheric Aerosol Particle Size Distribution. In: Journal of Aerosol Science, 2014, č. 77, 145 – 157.

[7] WAGNER, L. E. – DING, D.: Representing Aggregate Size Distributions As Modified Lognormal Distributions. In: Transactions of the American Society of Agricultural and Biological Engineers, 1994, č. 3, 815 – 821.

RESUMÉ

Tento článek se zabývá konstrukcí modelů rozdělení mezd pomocí čtyřparametrových a tříparametrových lognormálních křivek. Hlavním cílem výzkumu je porovnat přesnost pomocí obou typů lognormálního rozdělení jako modelů rozdělení mezd. Hrubá měsíční nominální mzda je hlavní proměnnou výzkumu. Výsledky z hlediska přesnosti obou typů lognormálních křivek jsou přibližně stejné pouze v několika případech. Rozdíly však nejsou kritické. Kvantilová metoda odhadu parametrů je použita k získání odhadů parametrů v obou případech. Počátek lognormálních křivek představuje minimální mzda v daném roce v obou případech pro čtyřparametrové a tříparametrové lognormální rozdělení. Testové kritérium chí-kvadrát je použito k vyhodnocení přesnosti získaných modelů. Hlavní vědecká hypotéza tedy spočívá v tvrzení, že pomocí kvantilové metody odhadu parametrů vedou čtyřparametrové lognormální křivky k přesnějším modelům než tříparametrové lognormální křivky.

V rámci bodového odhadu parametrů pomocí kvantilové metody poskytovaly čtyřparametrové lognormální křivky přesnější modely rozdělení mezd než tříparametrové lognormální křivky v převážné většině mzdových rozdělení. Čtyřparametrové a tříparametrové lognormální křivky přinesly stejnou přesnost ve výsledcích pouze v šesti z třiceti distribucí mezd. Tříparametrické lognormální křivky nebyly v žádném z případů přesnější než čtyřparametrové lognormální křivky. Tvar modelového rozdělení mezd společně s rostoucí úrovní mzdových rozdělení se výrazně mění s rostoucí velikostí firmy až na 1 000 zaměstnanců. U společností nad 1 000 zaměstnanců se tvar a úroveň rozdělení mezd už výrazně nemění.

RESUME

This paper deals with the construction of wage distribution models using four-parameter and three-parameter lognormal curves. The main objective of the research is to compare the accuracy using both types of lognormal distribution as wage distribution models. The main research variable is the gross monthly nominal wage.. The results in terms of the accuracy of both types of lognormal curves are approximately identical only in a few cases. However, the differences are not critical. The quantile method of parameter estimation is used to obtain the parameter estimations in both cases. The beginning of lognormal curves is the minimum wage in the given year in both cases, four-parameter and three-parameter lognormal distributions. The chi-square testing criterion is used to evaluate the accuracy of the models obtained. Thus, the main scientific hypothesis is based on the statement that using the quantile method of parameter estimation, four-parameter lognormal curves result in more accurate models than three-parameter lognormal curves.

Within the point parameter estimation using quantile method, four-parameter lognormal curves provided more accurate models of wage distributions than three-parameter lognormal curves in the vast majority of wage distributions. Four-parameter and three-parameter lognormal curves produced the same accuracy in the results only in six from thirty wage distributions. Three-parameter lognormal curves were not more accurate than four-parameter lognormal curves in either case. The shape of model wage distributions together with the growing level of wage distributions are changing significantly with the growing size of the company, up to

1,000 employees. For companies with more than 1,000 employees, the shape and level of wage distributions do not change significantly.

PROFESIJNÝ ŽIVOTOPIS

Doc. Ing. Diana Bílková, Dr., vyštudovala štatistiku na Fakultě informatiky a statistiky Vysoké školy ekonomické v Praze (1992) a od tohoto roku pedagogicky působí na jej katedře statistiky a pravděpodobnosti. Po získání titulu Dr. v odbore štatistika (1996) bola menovaná docentkou pre odbor štatistika v roku 2013. Vo vedecko-výskumnej práci sa zameriava najmä na oblasť teórie pravdepodobnosti a matematické štatistiky, obzvlášť na modelovanie, analýzu vývoja a predikciu mzdových a príjmových rozdelení. Je autorkou či spoluautorkou viac než stovky odborných článkov a statí v Českej republike i v zahraničí, vedeckých štúdií, skript, recenzií a oponentských posudkov a jednej učebnice. Má viac ako sedem desiatok citačných ohlasov v publikáciách iných autorov v Českej republike i v zahraničí (z toho cca tretinu predstavujú citácie sledované vo Web of Science). Podieľala sa na riešení osem vedeckých projektov, z nich vo dvoch prípadoch bola riešiteľkou. Je členkou České statistické společnosti a členkou organizačného výboru konferencie „International Days of Statistics and Economics” sledované v „Conference Proceedings Citation Index”. Bola vedúca takmer päťdesiatich úspešne obhájených diplomových prác.

KONTAKT

diana.bilkova@vse.cz

Informatívny článok/Informative article

Michal PÁLEŠ

**Katedra matematiky a aktuárstva, Fakulta hospodárskej informatiky
Ekonomickej univerzity v Bratislave**

MOŽNOSTI RIEŠENIA PROBLEMATIKY BIG DATA V JAZYKU R

ONE OF THE BIG DATA VIEWS IN R LANGUAGE

ABSTRAKT

Možno tvrdiť, že údaje sú v súčasnosti jednoznačne jedným z najhodnotnejších aktív v akejkoľvek sfére podnikania. V poisťovniach dennodenne pribúda množstvo údajov o klientoch, zmluvách, poisťných udalostiach, plneniach, majetku, záväzkoch a o iných skutočnostiach, ktoré sa opätovne využívajú v interakcii, v správach a reportoch pre manažment poisťovne a regulátorne orgány. Cieľom tohto príspevku je uviesť čitateľa do problematiky Big Data v aktuárstve a využitia v tomto kontexte jazyk R a Apache Hadoop.

ABSTRACT

It can be argued that data are clearly one of the most valuable assets in any sphere of business. A considerable amount of insurance data about clients, contracts, insurance events, claims, assets, liabilities and other events that are re-used in interactions, news and reports for the insurance company management and the regulatory authorities are increasing daily. The aim of this paper is to introduce the reader to the issue of Big Data in actuarial science and the use of R language and Apache Hadoop in this context.

KLÚČOVÉ SLOVÁ

big data, Apache Hadoop, jazyk R, aktuárstvo

KEY WORDS

big data, Apache Hadoop, R language, actuarial science

1. ÚVOD

V súčasnosti zažíva používanie Big Data značný rozmach v rôznych odvetviach. Žijeme v dobe, keď je táto technológia masívne používaná spoločnosťami na každodennej báze a často si ich použitie ani neuvedomujeme. Použitie Big Data nie je len doménou technologických firiem, ale vo veľkom ich využívajú aj banky, štátne inštitúcie či poisťovne. V čase, keď sú tieto dáta považované za štvrtý výrobný faktor, ich spoločnosti používajú viac ako kedykoľvek predtým. Získavajú tým nielen prehľad o internom prostredí firmy, ale najmä o svojich zákazníkoch, ich správani a aktuálnych trendoch. Správne použitie týchto dát znamená konkurenčnú výhodu a môže mať vplyv na budúcnosť firmy ako takej. Aby sa tieto veľké a neustále meniace objemy dát dokázali efektívne analyzovať, sú potrebné nové prístupy a platformy. Veda, ktorá sa zaoberá získavaním pridanej hodnoty z týchto dát, sa nazýva Data Science. Ide o jedno z najperspektívnejších odvetví v súčasnosti, ktoré sa neustále rozvíja veľkou rýchlosťou. Zaoberá sa analýzou, čistením, spracovaním dát a získavaním nových poznatkov, ktoré spoločnostiam umožňujú prijímať rýchlejšie a lepšie rozhodnutia. Poisťný trh a spoločnosti na ňom pôsobiace si tieto

skutočnosti veľmi dobre uvedomujú, a preto hľadajú stále nové možnosti ako skvalitniť existujúce a už implementované produkty. Pri aktuálnom trende možno niektorých aktúrov v poisťovniach považovať aj za dátových vedcov. Ich úlohou je neustále sledovať vývoj prebiehajúceho stavu a v prípade zmien v dostatočnom predstihu informovať a prispôbiť sa danej situácii. S novými možnosťami, ktoré Big Data prinášajú sa otvárajú nové možnosti aj aktúrom. Nové zdroje údajov, rastúci objem dát či meniace sa správanie zákazníkov – toto všetko sú výzvy, ktorým aktúar čelí. Dôležitým faktorom pri konkurenčnom súboji je používanie Big Data a spôsob ich spracovania. Existuje viacero možností, ako poskytnúť aktúrom prístup do týchto „veľkých dát“ a tým umožniť ich použitie v aktuárskej praxi. Jednu z takýchto možností predstavuje aj integrácia jazyka R a populárnej Big Data platformy Apache Hadoop, ktorá je jedným z najpopulárnejších open source nástrojov používaných na spracovanie Big Data.

2. BIG DATA A AKTUÁRSTVO

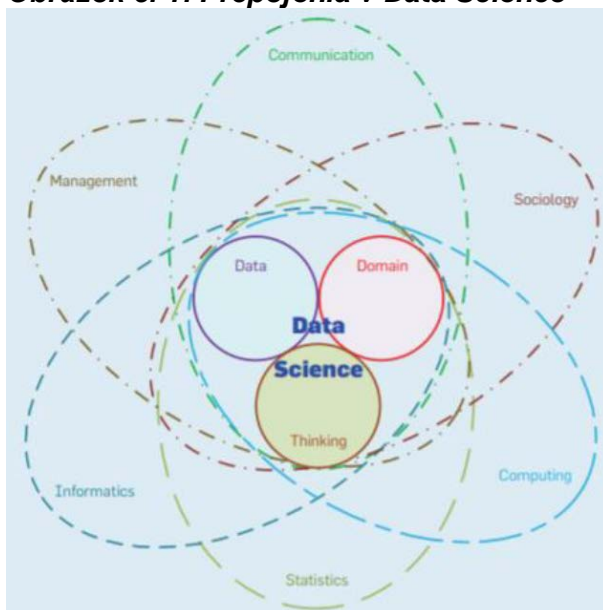
Slovné spojenie Data Science sa v posledných rokoch stáva čoraz populárnejším. Tento pojem už nie je predmetom diskusií výhradne odbornej verejnosti v oblasti informačných technológií, získal si popularitu aj v oddeleniach marketingu, predaja a je úspešným nástrojom pri prijímaní rozhodnutí vrcholového manažmentu. V doslovnom preklade ide o takzvanú dátovú vedu – účelom tejto počítačovej vedy je extrahovať zmysluplné poznatky z dát a následne ich ďalej efektívne využívať tak, aby nám priniesli určitú hodnotu. Avšak vnímať Data Science výlučne ako počítačovú vedu by nebolo celistvé. Odborníci na túto problematiku sa zhodujú, že ide o multidisciplinárny vedný odbor, ktorý využíva relevantné poznatky z oblastí informatiky, štatistiky, komunikácie, sociológie a manažmentu na analýzu dát. Tieto poznatky vznikajú analýzou dát, často hovoríme o Big Data, pričom dáta pochádzajú z rôznych informačných zdrojov a platforiem. Ide o dáta štruktúrované aj neštruktúrované, formalizované aj neformalizované a pochádzajúce z rôznych zdrojov – z firemných databáz, zo senzorov, z internetu či prieskumov. Na základe uvedomenia si týchto skutočností si môžeme utvoriť predstavu o komplexnosti a náročnosti tejto vednej disciplíny. Rozmach Data Science priamo súvisí s rozmachom Big Data, keďže spoločnosti si začali uvedomovať, že síce disponujú obrovským množstvom dát, no bez ich efektívnej analýzy a využitia nemajú žiadnu pridanú hodnotu pre spoločnosť. Uvedomenie si tejto skutočnosti viedlo spoločnosti k rozhodnutiu investovať značné finančné prostriedky práve na získanie tejto pridanej hodnoty. O dátach sa v súčasnej dobe hovorí ako o štvrtom výrobnom faktore popri tradičných troch – práci, pôde a kapitáli. Spoločnosti praktizujú tzv. Data Driven Decision Making – čo predstavuje rozhodovanie sa na základe údajov, teda rozhodnutia ktoré máme podložené dátami sú výhodnejšie, ako spoliehanie sa na intuíciu. Napríklad, marketingový manažér sa nebude spoliehať na svoje skúsenosti a odhad pri výbere typu reklamy, ale využije poznatky zistené na základe analýzy dát ako zákazníci reagujú na daný typ reklamy. Tieto metódy umožňujú spoločnostiam lepšie porozumieť tomu, čo zákazník preferuje a očakáva. Spoločnosti sa ďalej zameriavajú na to, podľa čoho sa rozhoduje, a prispôbujú sa tomu, čo im umožňuje získať konkurenčnú výhodu. Pre Data Science pozri [1], [12], [13].

Najznámejšiu a dodnes zaužívanú metódu charakterizujúcu Big Data zaviedol v roku 2001 Douglas Laney, viceprezident analytického tímu firmy Gartner. Ide o známu teóriu „troch V“. Neskôr sa k tomuto modelu pridali ešte ďalšie. Sú to začiatkové písmená anglických pojmov:

- **Data volume** (*objem údajov*),
- **Data velocity** (*rýchlosť údajov*),
- **Data variety** (*rôznorodosť údajov*),
- **Data veracity** (*pravosť údajov*),
- **Data value** (*hodnota údajov*),
- **Data volatility** (*prchivosť údajov*),
- **Data vulnerability** (*zraniteľnosť údajov*).

Špecialista, ktorý sa zaoberá Data Science sa nazýva data scientist – dátový vedec. Tento termín sa objavuje približne od roku 2008, odkedy sa na potrebu analýzy a spracovania dát začal klásť čoraz väčší význam. Ako môžeme vidieť na obrázku č. 1, výkon tohto povolania si vyžaduje znalosť mnohých vedeckých disciplín. Zároveň je však v súčasnosti považované za jedno z najžiadanejších povolání, v USA bola táto profesia označená za najatraktívnejšiu štyri roky po sebe [5]. Data scientist by mal okrem znalostí technológií disponovať nasledujúcimi schopnosťami: poznať špecifiká daného priemyslu, objasniť získané poznatky, rozumieť aktuálnej firemnej stratégii a podieľať sa na jej tvorbe a svojou prirodzenou zvedavosťou hľadať odpovede na aktuálne otázky. Z technologického hľadiska by mal dátový vedec poznať nasledujúce oblasti: rôzne programovacie jazyky (R, Python...), technologické platformy (Apache Hadoop, non-SQL databázy,...) a sledovať aktuálne technologické trendy (Cloud Computing,...).

Obrázok č. 1: Prepojenia v Data Science



Zdroj: [2]

Z opisu Data Science a profilu samotného dátového vedca vieme, že medzi jeho hlavné schopnosti a znalosti určite patrí znalosť kvantitatívnych metód, tvorba predikčných modelov, programovanie a hľadanie súvislostí v dátach a následné odporúčania a prezentácie jeho znalostí. V oblasti poisťovníctva je s oblasťou dátovej vedy spojená aj práca aktuára (pozri aj [8]). Najmä s rozmachom Big Data tieto dve profesie budú čoraz viac prepojené – potenciál Big Data v oblasti poisťovníctva a aktuárstva rozhodne existuje. Napriek tomu, stále tu môžeme pozorovať rozdiely – hlavné rozdiely medzi týmito povolaniami sú uvedené na obrázku č. 2.

Obrázok č. 2: Actuarial Science vs. Data Science

	Actuarial Science	Data Science
Basics	Mathematical Basics	
Data	Small Data	Small and Big Data
	Structured & static Data	Unstructured & dynamic Data
	Internal Data	External Data
Mathematics & Statistics	Probability Theory	Computational Statistics
	Life and Non-Life Insurance Mathematics	Algorithm
	Quantitative Risk Management	Information Theory
Computer science		Machine Learning & Visualisations
		Numeric Optimization
		Data Management
Programming languages	SAS, S Plus, R	Python, R
	SQL	SQL
	Excel / VBA	Julia, Spark, Scala
Domains/ business knowledge	Reserving, Pricing	
	ERM, ALM, Solvency	
	Accounting, Economics, Law	

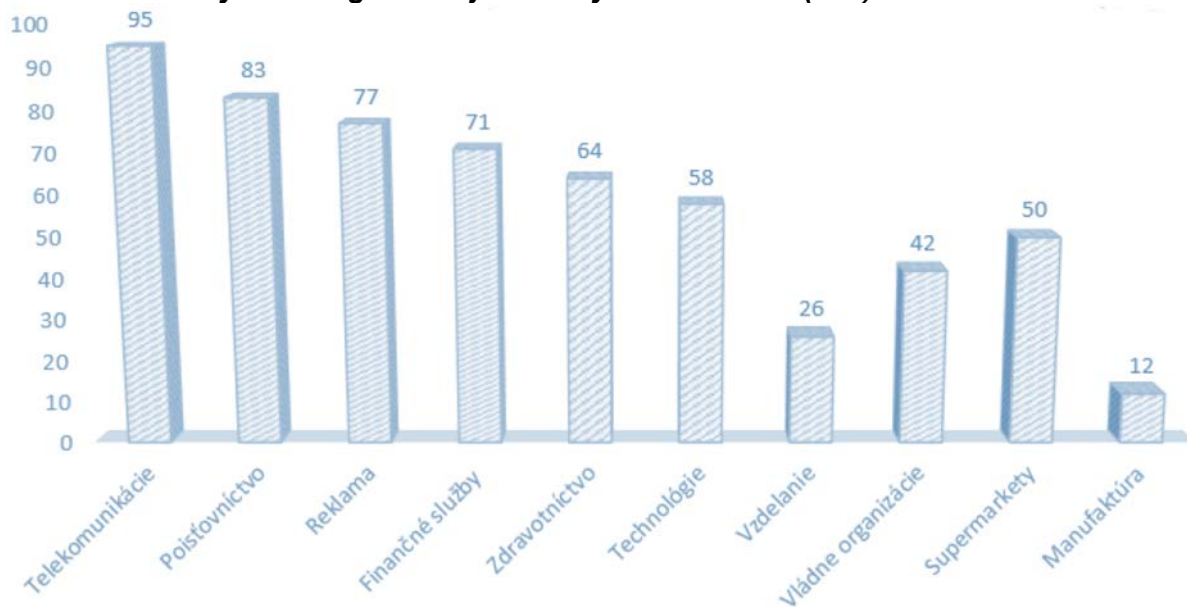
Zdroj: Swiss Association of Actuaries

Dátový vedec je viac programátorsky zameraný a medzi jeho znalosti patria aj rôzne no-SQL databázy, Hadoop či práca s neštruktúrovanými dátami. Aktuár pri svojej práci analyzuje všetky dostupné a relevantné dáta – historické, ale aj real-time (v reálnom čase). Následne pomocou rozličných techník získava požadované výsledky. Po získaní relevantných výsledkov ich aktuár prezentuje a podieľa sa na tvorbe stratégie spoločnosti. Pri svojej práci aktuári používajú rôzne matematické a štatistické programy alebo programovacie jazyky. Vďaka rýchlosti rozvoja technológií v poslednom desaťročí prešlo aj toto odvetvie významnými zmenami a nové zdroje dát si vyžadujú efektívnejšie metódy spracovania, čo predstavuje výzvu aj pre aktuárov. Z tohto dôvodu sa v tejto profesii kladie čoraz väčší dôraz na to, aby aktuár ovládal potrebné informačné technológie ako napríklad princípy databáz, programovacie jazyky či efektívne algoritmy na analýzu dát. V súčasnosti je hlavným cieľom aktuárov vytvárať presnejšie a spoľahlivejšie predikcie a modely. Dôraz sa kladie najmä na nájdanie rovnováhy medzi finančným úspechom a akceptovateľným rizikom. Dosiachnutie tejto rovnováhy predstavuje pre poisťovne ideálny stav z hľadiska ich dlhodobej stability (pozri aj [3]).

Aktuári oproti dátovým vedcom majú tú výhodu, že pracujú v oblasti, ktorá im je dôverne známa. Takisto, práca aktuárov je regulovaná legislatívou, nariadeniami a na svoju činnosť musia mať zodpovedajúce vzdelanie. Aktuár môže byť vo svojej profesii aj dobrým dátovým vedcom, no opačne to až tak neplatí. Vďaka aktuálnemu trendu rastu Big Data a rastúcej potrebe spoločností hľadať súvislosti v dátach a odhaľovať nové skutočnosti, tieto dve profesie konvergujú. Môžeme predpokladať, že aktuári v priebehu rokov budú vykonávať najmä úlohu dátových vedcov.

Poisťovníctvo ako odvetvie zohráva dôležitú úlohu v živote ľudí a firiem, pretože sa priamo zaoberá bezpečnosťou jednotlivcov a ich majetkov, pričom princíp fungovania tohto odvetvia je založený na existencii rizika. Ďalším jeho špecifikom je fakt, že poisťovníctvo nevyrába ani nespravuje žiaden fyzický produkt. V takomto prípade sú dáta pre poisťovne jedným z ich najhodnotnejších aktív. Vo všeobecnosti môžeme konštatovať, že celý princíp fungovania poisťovníctva je založený na predikcii rôznych rizík. Vysoko kompetitívny charakter tohto odvetvia a neustále zmeny správania zákazníkov tvoria predpoklad neustálych investícií do nových, efektívnych a presných spôsobov, ako predpovedať toto správanie a minimalizovať ich riziko. Všetky tieto skutočnosti naznačujú a vytvárajú predpoklady na použitie Big Data v poisťovníctve. Ako máme možnosť vidieť na obrázku č. 3, toto odvetvie si už v súčasnosti vo veľkom osvojilo používanie Big Data na svoju činnosť.

Obrázok č. 3: Využitie Big Data v jednotlivých oblastiach (v %)



Zdroj: [7], podľa Dresner Advisory Services

Jednou z najpoužívanejších techník pri práci aktuára je prediktívne modelovanie a analyzovanie. Ide o metódu, ktorá umožňuje určiť, čo sa udeje v budúcnosti na základe znalosti historických údajov. Následne sa vytvárajú modely založené na rôznych vzťahoch medzi jednotlivými premennými, ktoré boli zistené v týchto údajoch. Prediktívne modelovanie je, ako sme už uviedli, kľúčovým nástrojom Data Science a Big Data, pričom poisťovníctvo predstavuje odvetvie, ktoré tieto techniky dokázali veľmi rýchlo adaptovať. Táto metóda sa používa v aktuárskych analýzach najmä v nasledujúcich oblastiach:

- **posúdenie a hodnotenie rizík** – risk assesment; jedným z najdôležitejších oblastí pre poisťovne je stanovenie výšky poistného. Používa sa v životnom

a neživotnom poistení, napríklad pri havarijnom, zdravotnom poistení a poistení domácností. Poistovne dokážu využívať telematiku, IoT (Internet of Things) zariadenia či iné smart zariadenia na sledovanie zákazníkov a tým predpovedať dané riziko. Pomocou prediktívneho modelovania dokážu poistovne identifikovať pravdepodobnosť havárie, odcudzenia automobilu alebo inú poistnú udalosť pomocou kombinácie behaviorálnych dát a iných externých faktorov (ako napríklad počasie, stav vozovky či charakteristiky obytnej štvrte). V životnom poistení zase fitness náramky či smart hodinky umožňujú sledovať správanie a zvyky zákazníkov, čo umožňuje analýzu rizika na základe ich návykov;

- **odhalenie poistných podvodov** – fraud detection; dátová analýza spolu s kombináciou prediktívneho modelovania zlepšujú a umožňujú odhaliť kriminálnu aktivitu. Sledovaním premenných v každom poistnom plnení a ich následným porovnaním s poistnými podvodmi poistovne dokážu vytvoriť profil podozrivých nárokov na poistné plnenia. Pri odhalení zhody sa takýto prípad automaticky označí na ďalšie, podrobnejšie preskúmanie;
- **informácie o zákazníkoch** – customer insights; získavanie komplexného pochopenia správania, zvykov a potrieb zákazníkov z rôznych zdrojov je veľmi dôležité, aby poistovne mohli predvídať budúce správanie, správne produkty a identifikovať správne segmentácie zákazníkov. Informácie získané z call centier, e-mailov, sociálnych sietí, užívateľských fór či správanie zákazníkov na stránkach spoločností umožňuje poistovateľom vytvorenie unikátneho profilu jednotlivého zákazníka. Získanie prehľadu o zákazníkoch pomocou Big Data poskytuje predpovede nielen o tom, kedy zákazník pravdepodobne odíde ku konkurencii, ale formuje politiku zákazníka. To môže napomôcť poistovateľom rozvíjať dôverné vzťahy a integrovať zákazníkov správnym spôsobom a s presnými informáciami;
- **marketing** – poistovne môžu sledovať správanie spotrebiteľov v rôznych formách a vytvárať ciele reklamy, zbierať informácie pomocou cookies a iných mechanizmov či budovať modely dopytu, aby sa mohli zamerať na tých, u ktorých je pravdepodobnejšie, že si kúpia daný produkt. Tieto aktivity môžu pomôcť k zníženiu nákladov na reklamu, čo vedie aj k zníženiu celkových nákladov, a marketingové náklady prerozdeliť na iné účely;
- **vývoj produktov** – prediktívna analýza umožňuje nájsť nové oblasti pôsobenia a navrhnúť nové produkty pripravené na mieru klientov či zlepšiť už existujúce. Poistovne môžu potom ponúknuť lepšie produkty založené na analýze predchádzajúcich údajov o poistení, zdravotných záznamoch, spôsobe riadenia vozidla a životného štýlu.
- **automatizácia procesov** – poistovatelia používajú automatizáciu najmä pri jednoduchých a opakujúcich sa úkonoch, ktoré nevyžadovali ľudskú interakciu ako napríklad kontrola povinných polí či formátu dát. S nárastom Big Data je možné automatizovať aj komplexnejšie a zložitejšie úkony, ako upisovanie produktov, posúdenie majetku, overovanie nároku na plnenie, či odhalenie podvodov. Posunom k inteligentnejšej automatizácii spoločnosti môžu ušetriť obrovské množstvo času a finančných prostriedkov pomocou strojového učenia (Machine Learning);
- **data driven decision making** – prediktívne modelovanie dokáže napodobniť ľudské rozhodovanie na vytváranie nových pravidiel, ktoré sú lepšie ako tie používané predtým. Tým pádom sú získané výsledky mapované rýchlejšie a s väčšou presnosťou. Toto modelovanie prináša nesporné výhody, ale aj

určité obmedzenia – modelovanie ľudského správania môže spôsobiť, že v algoritmoch ostanú zakomponované ľudské predsudky.

Spomínané techniky je možno využiť v životnom aj neživotnom poistení. Rozvoj Big Data umožňuje nielen použitie týchto techník, ale prináša poisťovníam aj viac údajov, čo napomáha skvalitňovať ich produkty a mať lepší prehľad o zákazníkoch. V oblasti životného poistenia patria medzi tradičné zdroje dát najmä interné záznamy poisťovní, kde patrí napríklad vek poistenca, pohlavie, údaje o zdravotnom stave, či sociálnom statuse a iné. Nové, rozvíjajúce sa dáta pre životné poistenie zahŕňajú verejné záznamy, kriminálnu históriu, demografické údaje, genetické informácie, elektronické lekárske záznamy, históriu predpisov liekov, životný štýl a dáta z nosičov, ako sú napríklad smart hodinky a fitness náramky. Niektoré z týchto údajov sa používajú pred uzatvorením poistnej zmluvy, iné počas a po uzavretí ako určitá forma kontroly. Dáta zo sociálnych sietí ako Facebook, Instagram či Snapchat môžu odhaliť napríklad fajčenie alebo užívanie alkoholu, či iné dôležité informácie o zdravotnom stave poistenca.

Priekopníkom v tejto oblasti je poisťovňa John Hancock, ktorá ako prvá zaviedla zľavu na poistnom pre používateľov fitness náramkov Fitbit. Ide o náramok, ktorý je prepojený s internetom a odosiela údaje o fyzickej aktivite zákazníka. Náramok svojim klientom zadarmo zabezpečí poisťovňa. Spoločnosť uvádza, že klienti zapojení do tohto programu môžu získať zľavu na životnom poistení až do výšky 15 percent. Dáta sú pravidelne odosielané cez internet a používatelia tohto programu dostávajú body za svoju fyzickú aktivitu. Následne môžu túto aktivitu sledovať a získané body premeniť na rôzne zľavy a iné odmeny. Podľa prieskumu spoločnosti Accenture, tretina poisťovní v Severnej Amerike má vo svojom portfóliu podobný produkt, založený na sledovaní fyzickej aktivity spotrebiteľa [7].

Nové zdroje dát sa objavujú aj v neživotnom prostredí – pri poistení majetku alebo úrazovom poistení. K tradičným zdrojom dát v tejto oblasti patrili napríklad údaje o vodičovi, počet rokov skúseností, typ vozidla, počet najazdených kilometrov ročne (havarijné poistenie), typ konštrukcie budovy, geografická lokalita, druh zabezpečenia budov proti ohňu a vlámaniu (poistenie majetku), či údaje o oblasti podnikania. Nové zdroje dát v tejto oblasti sú:

- dáta o počasí, kriminalite, hustote obyvateľstva, hustote premávky, informácie zo sčítania obyvateľstva pre všetky druhy poistenia,
- telematické zariadenia v automobiloch, ktoré uľahčujú získať detailné informácie o správaní vodiča – tieto údaje sa používajú pri poistení automobilov pre firmy a jednotlivcov,
- mobilné dáta, dáta zo zariadení IoT a iných technológií, napríklad z inteligentných domov, najmä pri poistení majetku a podnikania,
- dáta zo smart zariadení, mobilov a fitness náramkov, elektronické zdravotné záznamy, záznamy o činnosti na internete, napríklad história hľadaných výrazov pri úrazovom poistení. Do tejto kategórie patria aj údaje o genofonde zákazníka a genetických predispozíciách na základe analýzy DNA, čo umožňuje genetické profilovanie.

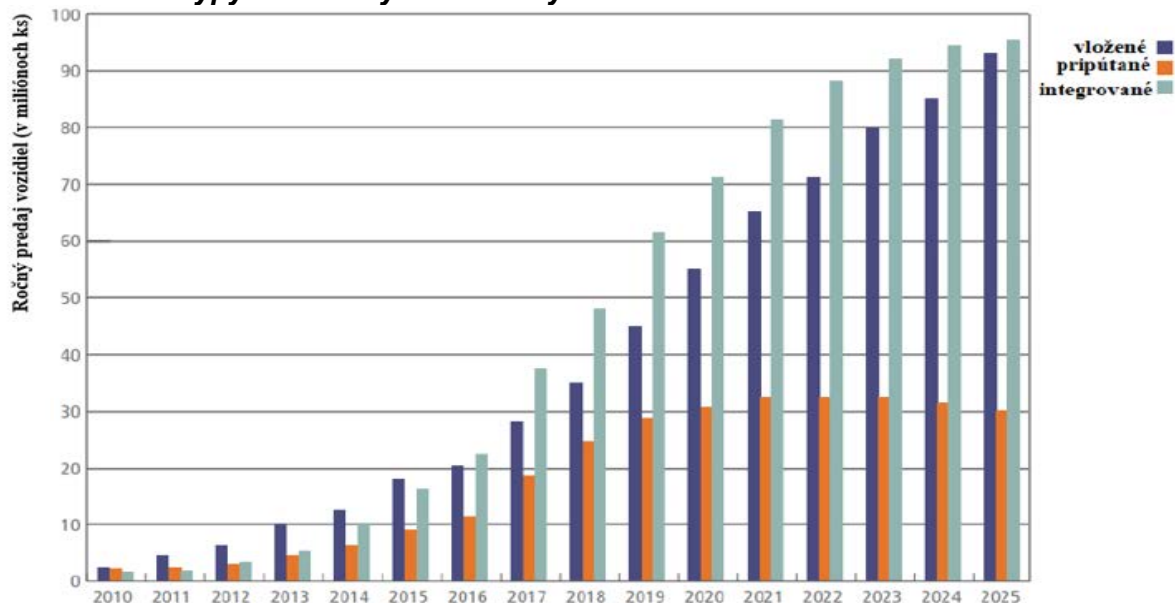
Ďalším rozmáhajúcim sa trendom v oblasti neživotného poistenia motorových vozidiel je tzv. Usage Based Insurance (UBI), známe aj ako Pay As you Drive (PAYD) alebo Pay How you Drive (PHYD), čiže poistenie založené na používaní alebo „platiť ako jazdíš“. Základnou myšlienkou UBI je, že správanie vodiča je

monitorované priamo počas toho, kedy dotyčná osoba jazdí, čo umožňuje poisťovateľom získať komplexnú predstavu o spôsobe jazdy a tým poskytovať individuálny prístup ku klientom. UBI je založené na telematike – ide o vedu, ktorá sa zaoberá diaľkovým prenosom a spracovaním dát. Najčastejšie používanou v praxi je dopravná telematika, pri ktorej ide o prenos údajov o vozidle do vzdialeného počítača, pričom sa využíva technológia GPS. Medzi najčastejšie zbierané údaje patrí počet kilometrov, poloha, rýchlosť, indikácia rýchleho zrýchlenia a spomalenia, použitie smeroviek, spôsob prechodu zákrut, použitie airbagov a iné. Úroveň zozbieraných údajov vo všeobecnosti závisí od typu použitej technológie a ochoty poistencov odosielať osobné údaje. Poisťovňa následne tieto údaje analyzuje a určí výšku poistného. Napríklad vodič často jazdiaci dlhé trasy, vo vysokých rýchlostiach bude mať vyššie poistné ako vodič jazdiaci krátke trasy pri nízkych rýchlostiach. Spôsob zberu týchto údajov sa neustále vyvíja – v začiatkoch (približne od roku 2005) sa do automobilov montovali prídavné zariadenia, ekvivalenty čiernych skriniek v leteckom priemysle, následne sa používali zariadenia na princípe USB, ktoré sa vložili do auta a v súčasnosti prevládajú mobilné aplikácie. Tento princíp v dnešnej dobe už používajú mnohé poisťovne, ako napríklad Allstate, Metromile, Root Insurance (USA), AIOI (Japonsko), Generali (EÚ), Wubi (Česká republika) či Allianz (Nemecko). Použitie tohto typu poistenia prináša poistencom viacero benefitov, najmä tým, čo ročne najazdia menej kilometrov – namiesto fixnej platby platia len za prejdené kilometre. Americká spoločnosť Metromile na svojich stránkach uvádza úsporu až 611 dolárov pri prejení 5 000 míľ (8 046 km) ročne a vo všeobecnosti dokážu vodiči s nízkym počtom kilometrov ušetriť až 50 percent za rok. [4]

Implementácia a využitie telematiky prináša do tohto segmentu poistenia veľa výhod. Poisťovne získavajú komplexný obraz o spôsobe jazdy, správaní sa zákazníkov a reálnom využití vozidiel a tieto údaje môžu využiť aj jednotlivé podniky a to prehľadom o využití firemných vozidiel a správania sa zamestnancov k nim. Zákazníkov zase láka vidina úspory na poistnom – okrem príležitostných vodičov získavajú benefit aj mladí vodiči, pretože pri klasickom spôsobe poistenia často platia vyššie poistné z dôvodu ich nízkeho veku a skúseností. Asociácia GSMA združujúca mobilných operátorov rozlišuje tri spôsoby implementácie v tejto oblasti:

- **vložené (zabudované)** – ide o inštaláciu prídavného zariadenia do motorového vozidla. Tieto zariadenia poskytujú poisťovne a zber údajov sa uskutočňuje pomocou týchto mechanizmov. V súčasnosti sú to malé skrinky, ktoré sú namontované fixne do vozidla alebo zapojené do palubnej dosky;
- **pripútané** – vodič vozidla používa mobilný telefón na odosielanie informácií pomocou mobilných dát, wifi či Bluetooth. Nevýhodou tohto typu je že je závislé od mobilného zariadenia, čo môže spôsobiť vyššie poplatky a takisto stratu dát v prípade výpadku mobilnej siete či zariadenia;
- **integrované** – aplikácie sú priamo integrované do automobilov. Ide o inteligentné automobily, ktoré obsahujú internetovú komunikáciu, pričom odpadá potreba používať mobilný telefón používateľa (napríklad systémy ako Apple CarPlay). GSMA predpokladá, že práve integrované riešenia budú v budúcnosti dominovať (pozri aj obrázok č. 4).

Obrázok č. 4: Typy telematiky v motorových vozidlách



Zdroj: [6]

Použitie telematiky so sebou takisto prináša aj určité nevýhody, plynúce najmä z nutnosti zdieľania dát s poisťovateľom. V prípade inštalácie zariadenia alebo použitia mobilnej aplikácie na zber týchto údajov, poisťovňa dostane všetky relevantné informácie, čo veľa zákazníkov považuje za narušenie ich súkromia. Keďže sa odosielajú údaje o polohe, spôsobe jazdy a rýchlosti, analýza týchto dát nemusí priniesť výhody, práve naopak – pre určitých zákazníkov môže viesť k zvýšeniu poistného (napríklad v prípade agresívneho štýlu jazdy) alebo intenzívneho používania motorového vozidla. Vo všeobecnosti sa PHYD poistenie viac oplatí vodičom s nízkym počtom najjazdených kilometrov ročne.

Používanie nových dátových zdrojov v neživotnom poistení nielenže zlepšuje segmentáciu zákazníkov, ale otvára aj nové možnosti poisťovniam, keď vďaka možnosti spracovania dát možno v reálnom čase dosiahnuť presnejšie výsledky. Až do rozvoja technológie Big Data sa aktuári pri výpočtoch spoliehali najmä na historické dáta. Klimatické zmeny sa v súčasnosti dejú rýchlejšie ako kedykoľvek predtým a globálna klíma má odlišný charakter ako pred rokmi. Tieto zmeny predstavujú nové riziko pre poisťovne, a preto je dôležité ich identifikovať a odhadnúť aktuálne trendy. Pomocou Big Data môžu poisťovne tieto skutočnosti odhaliť vďaka použitiu veľkého množstva senzorov, ktoré zachytávajú informácie o počasí, barometrickom tlaku, teplote, zmeny v prúdeniach vzduchu a podobne.

Dôležitým aspektom pri použití dát z nových zdrojov je ich kvalita. O kvalite týchto údajov pojednáva aj smernica EÚ s názvom Solventnosť II. Pre viac informácií o tejto problematike pozri [10]. Použitie Big Data v aktuárstve a poisťovníctve nesporne prináša veľa výhod a spoločnosti získavajú lepší prehľad o produktoch či zákazníkoch – avšak treba brať do úvahy aj etický aspekt použitia týchto údajov. Aktuári majú povinnosť vyplývajúcu z ich profesionalizmu dodržiavať povest' poistno-matickej profesie a nesú zodpovednosť voči verejnosti v novovznikajúcej oblasti Big Data. Dôležitou súčasťou je dodržiavanie regulácií a zákona. V mnohých situáciách majú aktuári jedinečný prehľad o výsledkoch a dôsledkoch použitíach týchto dát a musia byť ochotní s schopní tieto poznatky kľúčovým zainteresovaným

stranám (napríklad orgánom dohľadu, audítorom, verejnosti, či exekutíve) vysvetliť. Vysvetlenie svojich záverov je jedným z kľúčových atribútov poisťných matematikov. Odborný posudok aktúarov je rozhodujúci pri využívaní Big Data v poisťovniach. Posudok sa týka nielen pridanej hodnoty, ktorú Big Data prinášajú, ale aj etickosti ich využitia, pričom treba dodržiavať všetky regulácie tohto odvetvia a príslušné zákony jednotlivých krajín – napríklad v krajinách Európskej únie ochranu osobných údajov všeobecné nariadenie GDPR (General Data Protection Regulation). Toto všeobecne záväzné nariadenie vstúpilo do platnosti v roku 2018, pričom povinnosť dodržiavať ho majú všetky spoločnosti a organizácie spracúvajúce osobné údaje a pôsobiace v krajinách EÚ28, poisťovne nevynímajúc. Spoločnosti, ktoré nesplnia podmienky tohto nariadenia alebo nezabezpečia bezpečnosť osobných údajov, môžu dostať pokuty až do výšky niekoľko miliónov eur. Nové dáta ako údaje z fitness, zo sociálnych sietí, z automobilov a iných zdrojov sa dajú ľahko zneužiť. Odhalenie zneužitia údajov môže viesť k podniknutiu právnych krokov proti poisťovniam, pokutám, zákazu činnosti a poškodeniu dobrého mena spoločnosti. Je preto na individuálnom posúdení aktúara, vedenia spoločnosti a iných zúčastnených strán, ako a do akej miery budú tieto nové poznatky používať.

3. JAZYK R A APACHE HADOOP

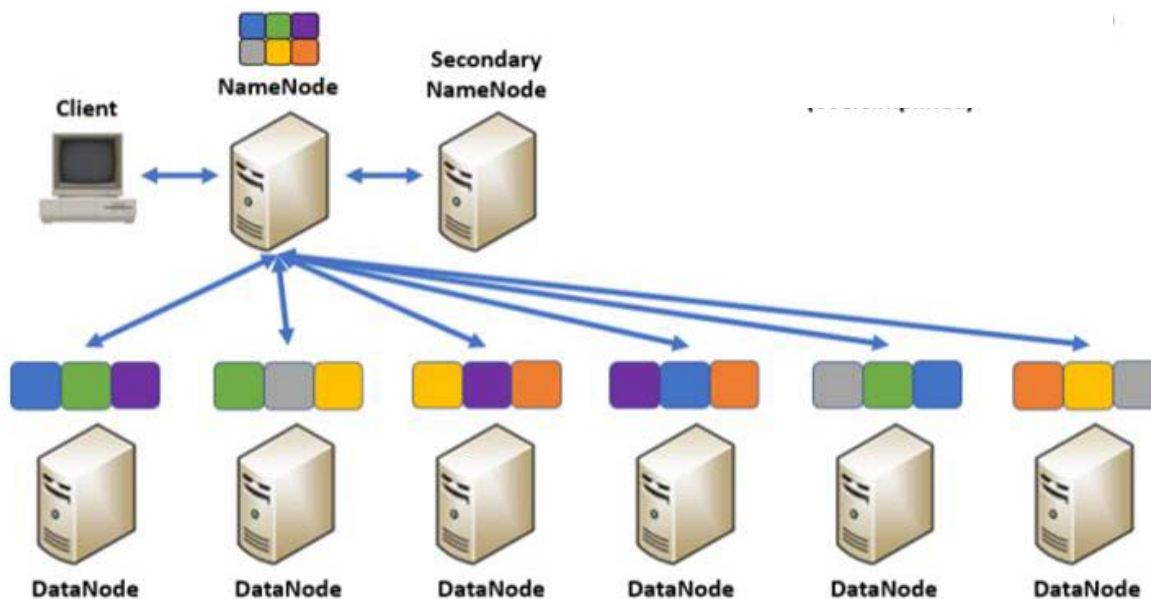
Jazyk R umožňuje používanie širokej škály štatistických a grafických techník, ako napríklad regresná analýza, analýza časových radov, štatistické testy, zhľukovanie a i. Všetky tieto skutočnosti z neho robia ideálnu voľbu pre Data Science, analýzu Big Data a Machine Learning. Nakoľko aktúar vo svojej praxi používa spomenuté techniky, nie je prekvapujúce, že R je populárnym nástrojom v tejto oblasti. Pre komplexné informácie o jazyku R a jeho využití v aktuárstve pozri [9].

Hadoop predstavuje open source projekt od spoločnosti Apache. Ide o softvérový framework, ktorý dokáže spracovať rozsiahle súbory v rámci počítačového klastra pomocou distribuovaného spracovania dát. Jednou z kľúčových charakteristík Hadoopu je jeho škálovateľnosť – dokáže pôsobiť na jednom serveri, ale aj na tisícoch počítačov. V súčasnosti ide o dominantnú platformu na spracovanie Big Data. Obsahuje viacero komponentov, ktoré zabezpečujú fungovanie klastra. Medzi základné komponenty patria:

- **HDFS** (Hadoop Distributed File System) – distribuovaný súborový systém používaný na spracovanie Big data naprieč klastrom. Je určený pre komoditný hardvér a je vysoko odolným voči jeho zlyhaniu. Dáta sú v ňom uložené v blokoch, ktoré sú následne replikované. Je napísaný v jazyku Java.
- **MapReduce** – programový model určený na spracovanie Big Data. Je navrhnutý tak, aby poskytoval spracované výsledky v prijateľnom čase. Pôvodne bol vyvinutý spoločnosťou Google.
- **YARN** (Yet Another Resource Negotiator) – doslova ďalší vyjednávač zdrojov. Jeho úlohou je efektívna alokácia zdrojov klastra na vykonávanie jednotlivých úloh. Apache Hadoop je založený na master/slave architektúre. Klastor pozostáva z viacerých serverov/počítačov, ktoré nazývame nody, respektíve uzly. Zvyčajne obsahuje aspoň jednu NameNode – hlavný server, ktorý riadi a rozdeľuje úlohy ostatným, podriadeným serverom, ktoré nazývame DataNodes. Týchto serverov môže byť v klastri ľubovoľný počet, pričom sú všetky riadené NameNode. Dáta uložené v HDFS naprieč klastrom sú replikované – to koľko krát majú byť replikované, určuje replikačný faktor, ktorý zabezpečí, že v prípade zlyhania jednej z DataNodes nedochádza k strate

údajov. O tom, kde sú jednotlivé bloky uložené má prehľad NameNode. Ak zlyhá jedna z DataNode, nastane automatický replikácia dát medzi ostatné, stále dostupné uzly a tým pádom sa zabezpečí dodržanie replikačného faktora. Informácie o tom, kde sú dáta v HDFS uložené má k dispozícii NameNode v svojej pamäti RAM. Aby sa predišlo jej zlyhaniu, klaster často obsahuje aj sekundárnu NameNode. Tá prevezme úlohu riadenia klastra v prípade výpadku primárnej NameNode a zabezpečí chod klastra. Pozri aj obrázok č. 5.

Obrázok č. 5: Zjednodušená architektúra Hadoop klastra



Zdroj: [11]

V závislosti od replikačného faktora dokáže klaster fungovať aj pri výpadku viacerých Nodes naraz. Po výmene chybných DataNodes a ich nahradení novými sú na tieto servery dáta znova automaticky replikované. To, že niektorá z Nodes zlyhá je očakávané, nie je dôležité či nastane, ale kedy. Tento princíp fungovania klastra nám nielen eliminuje hrozbu straty dát uložených v HDFS a potrebu záloh, ale aj ponúka značnú výpočtovú silu na spracovanie týchto dát. Existuje viacero distribúcií Hadoopu, najväčšími hráčmi na trhu sú spoločnosti Cloudera, Hortonworks a MapR. Všetky uvedené ponúkajú svoje vlastné distribúcie a pokročilé nástroje na manažovanie Hadoop klastrov. Najväčším podielom na súčasnom trhu disponuje Cloudera, ktorá niekedy býva označovaná za prvú komerčnú Hadoop spoločnosť. Ponúka vlastný produkt pod názvom Cloudera's Distribution Including Apache Hadoop (CDH), ktorý okrem už spomínaných komponentov používa množstvo iných nástrojov, ktoré pomáhajú k efektívnemu spracovaniu veľkých dát ako Apache Hive, HBase, Impala, a iné. Dôležitým prvkom CDH je Cloudera Manager – centralizovaný správca celého klastra, v ktorom je možné spravovať, monitorovať a nastavovať jednotlivé služby Apache Hadoop.

Tabuľka č. 1: Knižnice (packages) kolekcie RHadoop

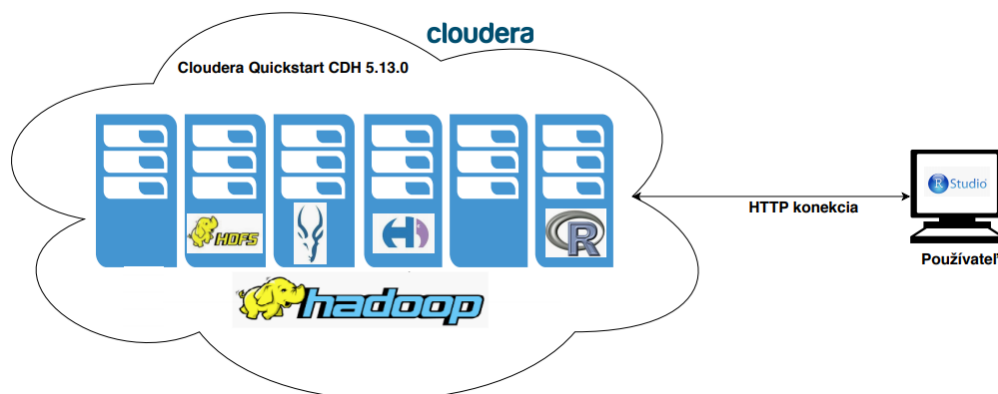
Knižnica	Opis
<i>Rhdfs</i>	Poskytuje základnú konektivitu medzi R a HDFS. Používatelia môžu čítať, zapisovať a modifikovať súbory z HDFS priamo z R. Tento balíček stačí nainštalovať na nóde, na ktorej beží samotný R server.
<i>Rhbase</i>	Poskytuje základnú konektivitu medzi R a databázou HBASE s použitím Apache Thrift servera. Používatelia môžu čítať, zapisovať a modifikovať tabuľky v databáze priamo z R. Tento balíček stačí nainštalovať na nóde, na ktorej beží samotný R server.
<i>Plymr</i>	Umožňuje používateľovi vykonávať bežné dátové operácie na dátach uložených priamo v Hadoope. Vychádza z populárneho balíčka plyr. Tento balíček je nutné nainštalovať na každú nódu v klastrí.
<i>rmr2</i>	Umožňuje vykonávanie analýz s použitím Hadoop MapReduce funkcionality. Tento balíček je nutné nainštalovať na každú nódu v klastrí.
<i>Ravro</i>	Poskytuje možnosť čítania a zápisu súborov typu avro pomocou R z HDFS. Tento balíček stačí nainštalovať na nóde, na ktorej beží samotný R server.

Zdroj: [7]

Ako platformu na spracovanie Big Data v [7] bol zvolený produkt Cloudera Quickstart. Ide o funkčnú simuláciu skutočného Hadoop klastra na jednom počítači. Obsahuje všetky potrebné služby a služby funkcionality skutočného, distribuovaného Hadoop klastra. Ako pracovné prostredie jazyka R bol použitý R Studio server, ktorý je nainštalovaný priamo na serveri, kde beží Cloudera Quickstart. Pre použitie tejto metódy sme sa rozhodli z dôvodu uľahčenia komunikácie Hadoopu a R. Prepojenie lokálneho R Studia so vzdialeným Hadoop klastrom nemá praktický význam a vyžadovalo by si zložité nastavenia komunikácie služieb – použitím R Studio server túto potrebu eliminujeme. Výhodou je aj jeho dostupnosť zo vzdialených počítačov cez URL prakticky z ľubovoľného počítača bez potreby inštalácie jazyka R na lokálnom počítači používateľa. Tento server sa vizuálne nijako nelíši od prostredia R Studio.

Na realizáciu konektie medzi R a Hadoopom sa používa kolekcia knižníc zvaná RHadoop. Skladá sa z viacerých balíčkov, ktoré umožňujú manipuláciu a analýzu dát uložených v Hadoop, ktoré uvádza tabuľka č. 1. Návrh architektúry riešenia zachytáva obrázok č. 6.

Obrázok č. 6: Architektúra možného prepojenia Hadoop a R



Zdroj: [7]

V súvislosti s Big Data a aktuárstvom, pochopiteľne, existuje viacero možností spracovania dát uložených v prostredí Hadoop. Okrem použitia externých nástrojov ako R a ich spojením s technológiou Big Data. Hadoop ponúka svoje vlastné, integrované riešenie na prístup a prácu s jeho ekosystémom – Hadoop User Experience (ďalej len HUE). Ide o open-source webový UI pre Hadoop, ktorý podporuje prechádzanie, vizualizáciu či dopytovanie dát v klastri. Tento komponent je súčasťou všetkých hlavných distribúcií Hadoopu, prípadne je možné ho dodatočne na klaster nainštalovať. Autentifikácia používateľa prebieha priamo na úrovni HUE, za pridelovanie prístupov je zodpovedný systémový administrátor. V prípade, že používateľ nemá vytvorený priečinok na HDFS, po registrácii v HUE sa priečinok automaticky vytvorí. Ak používateľ existuje, HUE sa zosynchronizuje s už existujúcim priečinkom a bude ho registrovať ako domovský. Po prihlásení do HUE sa spustí krátky tutoriál o jeho základných funkciách.

4. ZÁVER

Oba predstavené prístupy umožňujú spracovanie Big Data podľa potrieb používateľa. Každý z prístupov má svoje výhody aj nevýhody, pričom aj R aj HUE obsahujú veľké množstvo funkcií, ktoré aktuár pri svojej práci dokáže využiť. Hodnotenie oboch prístupov je uvedené v tabuľke č. 2 (1 – najlepšie, 5 – najhoršie). Po vyhodnotení všetkých kritérií máme možnosť vidieť, že obe technológie majú približne rovnaké hodnotenie, každá z nich vyniká v niečom inom. Môžeme konštatovať, že obe prinášajú široké možnosti spracovania Big Data a ďalej je už na stratégií jednotlivých spoločností, resp. rozhodnutí aktuárov, ktoré budú používať. Je logické, že aktuári so silnou znalosťou jazyka R budú preferovať spojenie jazyka R a Hadoopu na svoje analýzy. Pre prácu s HDFS a inými službami je však potrebná aspoň základná znalosť princípov technológií Big Data, preto aktuár pre lepšie pochopenie prepojenia technológií by mal absolvovať rôzne školenia ohľadom Hadoopu. Znalosť systému Linux, resp. aspoň jeho základných príkazov bude takisto veľkou výhodou. Výhodou R je, že aktuár môže použiť všetky funkcie a balíčky, ktoré R ponúka a následne ich aplikovať na Big Data.

Tabuľka č. 2: Komparácia prostredí R a HUE

	R	HUE
Užívateľské prostredie	2	1
Rýchlosť spracovania	2	2
Náročnosť nastavenia a prepojenia s Hadoop	3	1
Komplexnosť riešenia	2	3
Integrácia s inými technológiami	1	2
Variabilita možností spracovania	1	3
Riešenia pre Big Data	3	1
Možnosti rozšírenia funkcionality	1	3

Zdroj: [7]

HUE môžeme považovať za jednoduchšie prostredie z hľadiska nastavení potrebných na zabezpečenie konektivity. Ide o nástroj, ktorý prichádza s väčšinou Hadoop distribúcií ako integrovaný. HUE bolo vytvorené na priame spracovanie Big Data používateľmi, čo znamená, že ponúka rozsiahlejšie a robustnejšie nástroje ako R. Z hľadiska výkonnosti má HUE takisto výhodu oproti R – využívaním spojenej výpočtovej sily počítačového klastra môžeme dosiahnuť vynikajúce výsledky v rýchlosti spracovania dát aj v reálnom čase. Prostredie HUE je veľmi intuitívne a za krátky čas si ho osvoja aj noví používatelia. Ťažiskom práce s touto službou je znalosť SQL. Aktuár si následne môže vybrať, ktorú z Hadoop technológií použije pri svojich požiadavkách. Užívateľsky príjemný front-end služby prináša aj rozsiahle možnosti prechádzania súborov na HDFS a vizualizácie dát. Súborov môžeme jednoducho, na pár kliknutí nahráť, resp. stiahnuť aj z lokálneho počítača. Práca s dátami je takisto jednoduchá a intuitívna – vrátane vytvárania nových tabuliek z databázy. HUE je nepochybne nástrojom, ktorý hrá hlavnú úlohu v spracovaní Big Data v spoločnostiach, aktuárstvo nie je výnimkou. Aktuári teda nájdu v HUE široké spektrum možností, ktoré môžu použiť pri svojej práci.

Ktorú technológiu bude poisťovňa resp. oddelenie aktuárstva používať, závisí od ich preferencie či internej firemnej stratégie. Obe ponúkajú komplexné možnosti práce s Big Data a závisí od rozhodnutia kompetentných osôb, ktorý zo spôsobov použijú v každodennej aktuárskej praxi. Tento príspevok bol spracovaný najmä na základe snahy prezentovať závery získané pri tvorbe práce [7], ktorá bola ocenená Slovenskou spoločnosťou aktuárov. Práca obsahuje zdrojové kódy v R a obsahuje manuál na prepojenie R a Apache Hadoop. Samozrejme, dôležité je aj finančné hľadisko zavádzania daných platforiem v organizáciách, ktorému sa tu špeciálne nevenujeme.

Príspevok bol vytvorený v rámci projektu VEGA č. 1/0120/18 – Moderné nástroje riadenia rizika v interných modeloch poisťovní v kontexte direktívy Solvency II a VEGA č. 1/0647/19 – Moderné nástroje na riadenie a modelovanie rizík v neživotnom poistení.

LITERATÚRA

- [1] BERKLEY. What is Data science [online]. UC Berkley School of Information, 2018. [online]. [cit. 26. 1. 2020]. Dostupné na: <<https://datascience.berkeley.edu>>
- [2] CAO, L.: Data Science: Challenges and Directions. In Communications of the ACM 2017, č. 8, s. 59 – 68. ISSN 000-10782. [online]. [cit. 26. 1. 2020]. Dostupné na: <<https://doi.org>>
- [3] CIPRA, T.: Riziko ve financích a pojišťovnictví: Basel III a Solvency II. Praha: Ekopress, 2015. ISBN 978-80-87865-24-8.
- [4] FINANCIAL WEB. Pros and Cons of Pay as You Drive Insurance. 2016. [online]. [cit. 26. 1. 2020]. Dostupné na: <<https://www.finweb.com/>>
- [5] GLASDOOR. 50 Best Jobs in America. 2018. [online]. [cit. 26. 1. 2020]. Dostupné na: <<https://www.glassdoor.com>>
- [6] GSMA. 2025 Every Car Connected. 2012. [online]. [cit. 26.1.2020]. Dostupné na: <<https://www.gsma.com>>
- [7] MASÁR, J.: Problematika Big Data v súčasnej aktuárskej praxi. Diplomová práca: Ekonomická univerzita v Bratislave, 2019. 69 s.
- [8] PÁLEŠ, M.: Aktuárstvo v režime Solventnosť II (S riešenými príkladmi v jazyku R). Bratislava: Vydavateľstvo EKONÓM, 2016. 189 s. ISBN 978-80-225-4288-3.
- [9] PÁLEŠ, M.: Jazyk R pre aktuárov. Bratislava: Letra Edu, 2019. 349 s. ISBN 978-80-89962-26-6.

- [10] PÁLEŠ, M. Kvalita údajov a jej význam pre aktúarov. In: Slovenská štatistika a demografia, 2019, č. 1, s. 38 – 53.
- [11] SCHNEIDER, J.: Hadoop Essentials. Bratislava: Datavard, s.r.o., 2018, (interný dokument).
- [12] PIERSON, L.: Data Science for Dummies. Hoboken: John Wiley & Sons, Inc., 2017. 365 s. ISBN 978-1-119-32763-9.
- [13] PROVOST, F. – FAWCETT, T.: Data Science And Its Relationship to Big Data and Data-Driven Decision Making. 2013. [online]. [cit. 26. 1. 2020]. Dostupné na: <<https://www.liebertpub.com>>.

RESUMÉ

Informácie hrajú v dnešnom svete veľmi dôležitú rolu. Žijeme v dobe keď sa informácie stali najdôležitejšou konkurenčnou výhodou. V minulosti ľudia neboli obklopení toľkými informáciami ako dnes, a preto dnešná spoločnosť čelí problémom s kvalitou dát, nie s jej nedostatkom. Informácie ovplyvňujú každého z nás v našom rozhodovaní, a čím sú kvalitnejšie, tým nám umožňujú lepšie sa rozhodovať. Kvalitné dáta, z ktorých získame potrebné relevantné informácie, nám umožňujú znižovať straty, zvyšovať ziskovosť, analyzovať rôzne oblasti, jednoducho povedané, vedia nám zefektívniť fungovanie spoločnosti. Big Data priamo súvisia s vedným odborom Data Science. Existujú isté spoločné prieniky medzi povoláním dátových vedcov a aktúarov a v súčasnosti aj niektorí aktuári plnia v poisťovniach úlohu dátových vedcov. Na základe tohto predpokladu možno konštatovať potrebu použitia Big Data v tejto profesii. V súčasnosti už možno uviesť rôzne aplikačné príklady a spôsoby používania Big Data v poisťovníctve. Práca aktuára vyžaduje znalosť mnohých softvérových nástrojov, ktoré im pomáhajú dosahovať želané výsledky. Populárnym sa najmä v posledných rokoch stal jazyk R, ktorý predstavuje výkonnú platformu na spracovanie, analýzu a vizualizáciu dát. Popri jazyku R jeden z najpopulárnejších open source nástrojov používaných na spracovanie Big Data je Apache Hadoop. Možno realizovať porovnanie jazyka R a Hadoop User Experience – ich možností, ktoré pri spracovaní veľkých dát ponúkajú, architektúru a princípy ich fungovania, ako aj dôvody ich popularity. Na základe týchto skúseností možno zvoliť tú platformu, ktorá je ideálnym nástrojom pre prácu aktúarov.

RESUME

Information plays a very important role in today's world. We are living in an era when information has become the most important competitive advantage. In the past, people were not surrounded by as much information as today, and today's society is facing data quality problems, not their deficiency. The information affects all of us in our decision making, and the better it is, the better decisions it enable us to make. High-quality data, from which we obtain the necessary relevant information, provide us with reducing losses, increasing profitability, and analyzing various areas and to put it simply, they can improve the functioning of the society. Big Data is directly related to Data Science. There are some common intersections between the profession of data scientists and actuarians, and at present some actuarians also act as data scientists in insurance companies. The necessity of using Big Data in this profession can be concluded on the basis of this assumption. Various application examples and ways of using Big Data in the insurance industry can already be mentioned. Actuarial work requires knowledge of various software tools helping them to achieve the desired results. Especially in recent years, the R language has become popular as a powerful platform for data processing, analysis and visualization. In addition to the R

language, one of the most popular open source tools used in the processing of Big Data is the Apache Hadoop. Consequently, a comparison of the R language and the Hadoop User Experience can be carried out- their possibilities offered in the processing of big data, architecture and principles of their functioning, as well as reasons for their popularity. These experiences enable the selection of the platform that is the ideal work tool of actuarians.

PROFESIJNÝ ŽIVOTOPIS

Ing. Michal Páleš, PhD., pôsobí ako vedúci Katedry matematiky a aktuárstva Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave. V rámci pedagogickej činnosti vyučuje predmety matematika, teória pravdepodobnosti, softvérové aplikácie pre aktúarov, teória rizika v poistení, úvod do aktuárstva a vybrané kapitoly z matematiky. Vo svojej vedeckej práci sa orientuje na aktuársku vedu, využitie kvantitatívnych metód v ekonómii a softvérovú podporu riadenia rizík (najmä jazyk R). Je autorom viacerých medzinárodne ocenených vysokoškolských učebníc, vedeckých monografií a článkov z oblasti aktuárstva.

KONTAKT

pales.euba@gmail.com

Informatívny článok/Informative article

Alžbeta RIDZOŇOVÁ
Štatistický úrad Slovenskej republiky

ASYMETRIE V ŠTATISTIKE ZAHRANIČNÉHO OBCHODU

ASYMMETRIES IN FOREIGN TRADE STATISTICS

1. ÚVOD

Nariadenie (ES) č. 223/2009 Európskeho parlamentu a Rady z 11. marca 2009 o európskej štatistike¹ je základom pre rozvoj, tvorbu a šírenie európskych štatistík. Keďže tieto štatistiky majú zodpovedať požiadavkám užívateľov vrátane požiadaviek na kvalitu, musia byť tvorené v súlade s kritériami kvality stanovenými v uvedenom nariadení.

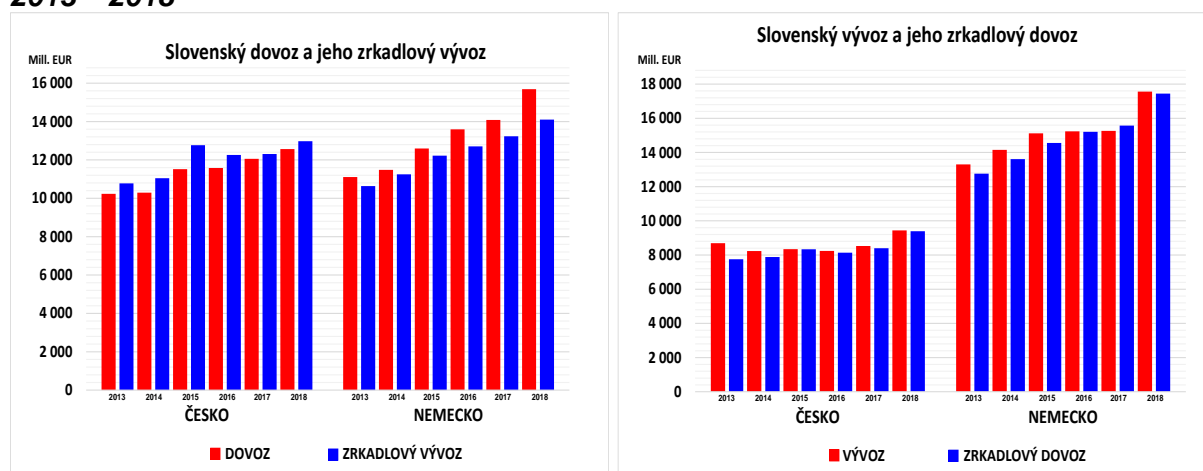
Porovnateľnosť je jedným z kritérií, ktoré hodnotí mieru vplyvu rozdielnosti v aplikovaných štatistických koncepciách a nástrojoch a postupoch merania, ak sa porovnávajú štatistiky medzi geografickými oblasťami alebo odvetvami alebo ak sa porovnávajú v čase. Porovnateľnosť štatistík medzi geografickými oblasťami sa dá merať na základe asymetrií, ak príslušné oblasti, napr. členské štáty, zisťujú a zostavujú ukazovatele o tokoch smerujúcich proti sebe.

Štatistika zahraničného obchodu je jednou z tých štatistík, kde sa v značnej miere na monitorovanie a hodnotenie kvality využívajú asymetrie.

2. PRÍČINY ASYMETRIÍ V ŠTATISTIKE ZAHRANIČNÉHO OBCHODU

Štatistika zahraničného obchodu danej krajiny sa prezentuje ako „dovoz krajiny A z krajiny B“ resp. „vývoz krajiny A do krajiny B“. Užívateľ preto opodstatnene očakáva, že opačné toky, t. j. „vývoz krajiny B do krajiny A“ resp. „dovoz krajiny B z krajiny A“, sa zhodujú. Tento teoretický koncept však nefunguje v praxi, ako to ukazujú grafy.

Graf č. 1 a č. 2: Zahraničný obchod Slovenska s Českom a Nemeckom v rokoch 2013 – 2018



Zdroj údajov: Eurostat, EasyComext (vybrané 15.1.2020)

¹ Úradný vestník EÚ L 87 z 31.3.2009, s. 164.

Dôvody rozdielov pri porovnávaní „zrkadlových“ tokov v štatistike zahraničného obchodu sú do určitej miery už zabudované v konceptoch a definíciách, na ktorých je táto štatistika založená. Jej základný princíp je zaznamenať všetky tovary, ktoré opúšťajú alebo vstupujú na ekonomické územie zostavujúcej krajiny. Colné vyhlásenia a Intrastat hlásenia sú štandardnými zdrojmi informácií, ktoré sa riadia fyzickým pohybom tovaru. Sú predkladané vývozcami a dovozcami v čase, keď tovar opúšťa krajinu alebo do nej vstupuje. Zatiaľ čo vývozca vykazuje údaje o vývoze na základe predpokladov, dovozca poskytuje informácie na základe faktov. Zaznamenané asymetrie v štatistike zahraničného obchodu sú dôkazom toho, že predpoklady vývozcu, t. j. informácie známe v čase vývozu, sa môžu líšiť od reality, ktorá nastane, keď tovar opustí krajinu vývozu.

Metodické dôvody ako definícia sledovaného obdobia, princíp určenia hodnoty vrátane použitého výmenného kurzu, obchodný systém, definícia partnerskej krajiny atď. môžu vysvetliť asymetrie najmä v obchode s tretími krajinami, keď sa porovnávajú údaje o obchode daného členského štátu s údajmi zostavovanými krajinou, ktorá nie je členom EÚ.

Aj keď je zostavovanie štatistiky obchodu s tovarom medzi členskými štátmi v rámci EÚ harmonizované, asymetrie zaznamenané v tejto štatistike sú užívateľmi považované za vážny problém. Navyše, okrem príčin uvedených vyššie môžu byť asymetrie spôsobené aj transakciami typu „kvázi tranzit“ a „trojstranný obchod“. Kým trojstranný obchod môže byť pozorovaný aj v obchode s tretími krajinami, kvázi tranzit je špecifikom obchodu medzi členskými štátmi ako dôsledok obchodovania s tovarom v rámci jednotného trhu.

Keďže pri vykazovaní vývozu a dovozu sú požadované komplexné informácie, môžu sa vyskytnúť aj chyby. Najčastejšími chybami sú najmä chybné zatriedenie tovaru podľa kombinovanej nomenklatúry alebo nesprávne vykázaná partnerská krajina.

Príčiny rozdielov pri zrkadlovom porovnávaní údajov v štatistike zahraničného obchodu sú podrobne opísané v *Compilers guide on European statistics on international trade in goods – 2017 edition*.

3. ASYMETRIE V OBCHODE MEDZI ČLENSKÝMI ŠTÁTMI – AKO SA RIEŠILI

Asymetrie v štatistike zahraničného obchodu, a to najmä v obchode medzi členskými štátmi, sa pravidelne monitorujú, preskúmavajú, a ak je to možné, údaje sa revidujú.

Eurostat v spolupráci so zostavovateľmi štatistiky zahraničného obchodu v členských štátoch organizoval niekoľko kôl „zrkadlových“ porovnávaní údajov o vzájomnom obchode dvoch členských štátov. Identifikované boli významné asymetrie na najpodrobnejšej komoditnej úrovni, t. j. na úrovni položiek kombinovanej nomenklatúry a tieto sa analyzovali. Úlohou Eurostatu bolo vybrať asymetrie na analýzu tak, aby pracovné zaťaženie členských štátov bolo vyvážené. Cieľom týchto analýz bolo identifikovať príčiny asymetrií a tam, kde to bolo možné, revidovať štatistické výsledky.

Tieto analýzy viedli nielen k zlepšeniu kvality štatistiky obchodu medzi členskými štátmi, ale problémy identifikované počas realizácie analýz viedli k novelizácii Intrastat legislatívy. Špecifické ustanovenie umožňujúce výmenu dôverných údajov o obchode medzi členskými štátmi bolo zavedené Nariadením (ES) č. 638/2004 Európskeho parlamentu a Rady z 31. marca 2004 o štatistike spoločenstva o obchodovaní s tovarom medzi členskými štátmi a o zrušení nariadenia Rady (EHS) č. 3330/91.

SIMSTAT projekt bol míľnikom v zrkadlových porovnávaní údajov o obchode medzi členskými štátmi. Mikroúdaje o vývoze v rámci EÚ, t. j. údaje umožňujúce identifikovať dovozcu, vymieňané medzi príslušným štatistickými orgánmi boli použité pri porovnávaní v roku 2015. Mikroúdaje, hoci simulované na základe údajov zo súhrnných výkazov DPH, umožnili zostavovateľom štatistiky zahraničného obchodu členského štátu, ktorý tieto mikroúdaje dostal, identifikovať potenciálnych dovozcov pri tých asymetriách, kde je vývoz odosielajúceho členského štátu vyšší ako zodpovedajúci dovoz.

Členské štáty realizovali aj bilaterálne zrkadlové porovnávaní. Štatistický úrad SR spolupracuje veľmi úzko s Českým štatistickým úradom tiež v oblasti štatistiky zahraničného obchodu. Táto spolupráca sa začala okamžite po rozdelení Československa, keď obe krajiny museli začať používať štandardné colné pravidlá vo vzájomnom obchode. Zrkadlové porovnávaní štatistík o dovoze a vývoze tovaru oboch krajín boli nástrojom umožňujúcim oboj orgánom, colným aj štatistickým, validovať svoje štatistické výsledky.

4. ASYMETRIE V OBCHODE MEDZI ČLENSKÝMI ŠTÁTMI – AKO SA BUDÚ RIEŠIŤ

Intrastat systém sa od svojho zavedenia v roku 1993 považuje za jedno z najviac zaťažujúcich štatistických zisťovaní. Na vyhodnotenie možností, ako znížiť zaťaženie vykazujúcich vývozcov a dovozcov a zároveň zachovať kvalitu štatistiky obchodu medzi členskými štátmi, bol spustený ESS VIP projekt REDESIGN. Okrem potenciálneho zníženia spravodajskej záťaže boli jednotlivé možnosti preskúmané z pohľadu nákladov a prínosov implementácie, ako aj nutných legislatívnych a metodických zmien.

ESS VIP projekt SIMSTAT, ktorého hlavným cieľom bolo otestovať realizovateľnosť výmeny mikroúdajov o vývoze v rámci EÚ, dopĺňoval projekt REDESIGN. Tieto vymieňané mikroúdaje obsahujúce identifikačné číslo pre DPH dovozcu môžu byť použité na zostavenie štatistiky dovozu v rámci EÚ v partnerskom členskom štáte.

Oba projekty boli úspešne ukončené v prvej polovici roka 2016. Výsledkom projektov bolo, že Eurostat a členské štáty začali implementovať „modernizovaný Intrastat systém“, ktorého právna infraštruktúra – základné nariadenie o európskych podnikových štatistikách tzv. FRIBS a jeho vykonávacie a delegované akty sa budú v štatistike zahraničného obchodu uplatňovať od 1. januára 2022.

Modernizovaný Intrastat systém zavedie nové údajové prvky vo vykazovaní vývozu tovaru v rámci EÚ: identifikačné číslo pre DPH obchodného partnera v členskom štáte dovozu a krajinu pôvodu. Takto vykázané a zozbierané mikroúdaje

o vývoze tovaru v rámci EÚ spolu s ďalšími údajmi o vývoze tovaru v rámci EÚ z colných vyhlásení a iných zdrojov údajov pre špecifické tovary a pohyby budú prostredníctvom Eurostatu zaslané dovážajúcemu členskému štátu.

Dovážajúci členský štát môže použiť všetky prijaté mikroúdaje alebo len ich časť na zostavenie svojej štatistiky dovozu tovaru v rámci EÚ. Je na rozhodnutí členského štátu, kedy a aká časť zbieraných údajov o dovoze tovaru v rámci EÚ bude nahradená vymieňanými mikroúdajmi.

Rozhodnutie, prečo sa budú vymieňať údaje o vývoze a nie o dovoze, vychádzalo zo skutočnosti, že vývoz je koncentrovanejší ako dovoz, ktorý realizuje oveľa viac hospodárskych subjektov, a teda aj počet spravodajských jednotiek je výrazne vyšší. Okrem toho vývozcovia poznajú svoju produkciu, a preto sú schopní presnejšie a s menším úsilím zatriediť tovar podľa kombinovanej nomenklatúry (spravodajské jednotky považujú určovanie kódu kombinovanej nomenklatúry za zaťažujúce).

Bez ohľadu na to, či dovážajúci členský štát použije vymieňané mikroúdaje priamo na zostavenie svojej štatistiky dovozu tovaru v rámci EÚ alebo nie, tieto údaje budú doplnkovým zdrojom informácií o tovarovej štruktúre dovozu v rámci EÚ. Ak budú pri porovnávaní vymieňaných mikroúdajov so zbieranými údajmi identifikované asymetrie alebo chybné zatriedenia tovaru, tie bude možné vyriešiť oveľa ľahšie. Okrem toho dovážajúci členský štát môže vymieňané mikroúdaje použiť na odvodenie tovarovej štruktúry chýbajúcich dovozov v rámci EÚ, teda dovozov nevykázaných, pretože je dovozca oslobodený od spravodajskej povinnosti alebo si túto povinnosť nespĺnil.

5. ZÁVER

Zníženie asymetrií v štatistike obchodu medzi členskými štátmi je nevyhnutným predpokladom na nahradenie zbieraných údajov o dovoze tovaru v rámci EÚ vymieňanými údajmi. Aj keď bude možné pozorovať výrazné zlepšenie, asymetrie vychádzajúce z používanej metodiky nebudú úplne odstránené.

Mgr. Alžbeta RIDZOŇOVÁ

Autorka je riaditeľkou odboru štatistiky zahraničného obchodu sekcie makroekonomických štatistík Štatistického úradu SR.

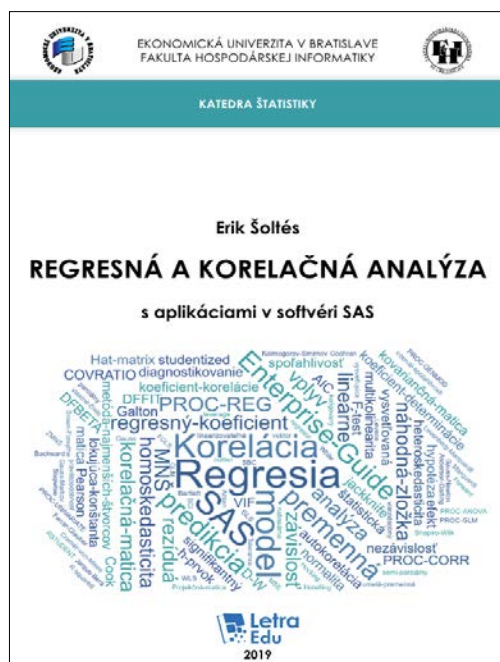
Recenzia publikácie/Review of Publication

Erik Šoltés
**REGRESNÁ A KORELAČNÁ ANALÝZA
S APLIKÁCIAMI V SOFTVÉRI SAS**

Erik Šoltés
**REGRESSION AND CORRELATION ANALYSIS
WITH APPLICATIONS IN SAS SOFTWARE**

Bratislava: LetraEdu, 2019. 238 s.
ISBN 978-80-89962-38-9 (kniha)
ISBN 978-80-89962-39-6 (online)

Metódy a postupy regresnej a korelačnej analýzy patria medzi najčastejšie využívané štatistické techniky v rôznych vedných oblastiach. Pretože umožňujú opísať komplikované vzťahy medzi rôznymi premennými, majú široké možnosti uplatnenia nielen v ekonomických, ale aj sociálnych analýzach, pričom svoje dôležité miesto majú aj v prírodných vedách. V súčasnej dobe, keď sú analytici aj vedci konfrontovaní s veľkým množstvom údajov (čo sa týka veľkosti súborov, ako aj množstva analyzovaných premenných), je využitie štatistického softvéru takmer nevyhnutným predpokladom aplikácie všetkých štatistických metód, teda aj regresnej a korelačnej analýzy. Z uvedených dôvodov treba vysokoškolskú učebnicu *Regresná a korelačná analýza s aplikáciami v softvéri SAS*, ktorej autorom je Erik Šoltés z Katedry štatistiky Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave, považovať za veľký prínos pre prípravu a vzdelávanie odborníkov v oblasti kvantitatívnych metód.



Učebnica je prioritne určená pre študentov Fakulty hospodárskej informatiky, predovšetkým pre študijné programy *Štatistické metódy v ekonómii* a *Aktuárstvo*, ktoré sú súčasťou 2. stupňa vysokoškolského vzdelávania. Publikácia má však potenciál byť učebnou pomôckou aj na iných vysokých školách ekonomického (prípadne spoločenského) zamerania, pričom cenné poznatky a návody na riešenia praktických úloh v nej môžu nájsť aj analytici a vedeckí pracovníci, ktorí skúmajú vzájomné vzťahy medzi kvantitatívnymi premennými v ekonómii a v iných vedných disciplínach.

Za pridanú hodnotu učebnice možno považovať aj to, že na vzorových príkladoch uvádza riešenia čiastkových úloh pre komplexnú regresnú a korelačnú analýzu v module

Enterprise Guide od spoločnosti SAS. Štatistický softvér SAS patrí medzi svetových lídrov v oblasti analytiky. Na Slovensku je takisto pomerne rozšírený, a to najmä na analytických oddeleniach bánk, poisťovní, telekomunikačných spoločností, ale aj

v ďalších spoločnostiach korporátnej sféry, čo zvyšuje možnosti uplatnenia tejto publikácie.

Učebnica poskytuje podrobný opis metód a postupov na komplexný rozbor lineárneho regresného modelu. V prvých piatich kapitolách sa okrem definovania základných pojmov venuje pozornosť odhadu parametrov klasického lineárneho regresného modelu, indukčným úsudkom o parametroch modelu, hodnoteniu vplyvu vysvetľujúcich premenných na cieľovú premennú, výberu vysvetľujúcich premenných do regresného modelu, posúdeniu kvality regresného modelu, diagnostikovaníu vplyvných pozorovaní, analýze multikolinearity, overeníu predpokladov klasického lineárneho regresného modelu a predikcii hodnôt vysvetľovanej premennej. Piata kapitola sa navyše zameriava na riešenie problému nesplnenia predpokladov o náhodnej zložke regresného modelu. Šiesta kapitola sa venuje odhadu parametrov nelineárnych regresných modelov. Poznatky prezentované v každej kapitole majú ucelenú logickú štruktúru, sú podložené argumentami a doplnené potrebnými súvislosťami.

Metódy a postupy regresnej a korelačnej analýzy sú v učebnici bezprostredne po ich vysvetlení aplikované na vzorových úlohách. Tieto aplikácie sú realizované prostredníctvom modulu Enterprise Guide softvéru SAS, ktorý je medzi používateľmi veľmi obľúbený, aj vďaka tomu, že nevyžaduje znalosť SAS programovacieho jazyka. Na základe práce používateľa v interaktívnych oknách sa SAS kód generuje automaticky na pozadí. V učebnici je ukázané ako treba nastaviť interaktívne okná pri riešení vzorových úloh. Tieto riešenia na pozadí využívajú najmä procedúru regresnej analýzy (PROC REG) a procedúru korelačnej analýzy (PROC CORR), avšak riešenia niektorých úloh sú realizované aj s využitím analýzy rozptylu (PROC ANOVA), zovšeobecnených lineárnych modelov (PROC GENMOD) alebo nelineárnych regresných modelov (PROC NLIN). V učebnici sa popri tom ukazuje, ako treba v uvedených procedúrach upraviť SAS kód tak, aby bolo možné riešiť aj špecifické úlohy, pri ktorých sa obyčajne vyžaduje zásah do programovacieho kódu.

Učebnica obsahuje dve prílohy. V prílohe A sú základy maticového počtu, ktorý je matematickým podkladom prezentovaných metód. Príloha B poskytuje vybrané štatistické tabuľky, ktoré sú potrebné pri riešení praktických úloh. Čitatelia určite ocenia aj to, že autor publikácie prepája poznatky z rôznych častí regresnej a korelačnej analýzy a poukazuje na vzťahy medzi rôznymi metódami a postupmi (nielen pri vysvetľovaní metód, ale aj pri riešení vzorových úloh), čo umocňuje didaktickú hodnotu tejto učebnej pomôcky. Prezentované súvislosti pomáhajú vytvoriť potrebný nadhľad a schopnosť orientácie v problematike. Dôraz sa kladie aj na správnu interpretáciu výsledkov vzorových úloh.

Knihu možno jednoznačne považovať za vydarenú vysokoškolskú učebnicu, ktorá využíva vhodné didaktické postupy na vysvetlenie metód a postupov regresnej a korelačnej analýzy, pričom záujemcovia si môžu poznatky ešte doplniť z publikácií, na ktoré sa autor odvoláva. Je komplexnou učebnou pomôckou, pretože prináša aj návody na riešenia praktických úloh v renomovanom štatistickom softvéri, ktorý je aj v slovenskej praxi pomerne rozšírený.

RNDr. Eva KOTLEBOVÁ, PhD.

Autorka je zástupkyňa vedúcej Katedry štatistiky na Fakulte hospodárskej informatiky Ekonomickej univerzity v Bratislave.

Recenzia publikácie/Review of Publication

Věra Kuchařová a kol.
**ČESKÁ RODINA NA POČÁTKU 21. STOLETÍ
ŽIVOTNÍ PODMÍNKY, VZTAHY A POTŘEBY**

Věra Kuchařová a kol.
**CZECH FAMILY AT THE BEGINNING OF THE 21ST CENTURY
LIVING CONDITIONS, RELATIONSHIPS AND NEEDS**
Praha: Sociologické nakladatelství SLON, 2019. 299 s.
ISBN 978-80-7419-275-3

Predmetom mnohých odborných štúdií sú rodinné, partnerské a reprodukčné vzťahy a ich zmeny a aktuálne trendy. Skúmanie demografického, reprodukčného a rodinného správania je venovaná zvýšená pozornosť, pretože tieto javy významne ovplyvňujú mnohé stránky sociálno-ekonomického a kultúrneho života spoločnosti. Vytvára sa pestrá mozaika foriem rodinného súžitia, ktoré sú výsledkom rozmanitých životných volieb a preferencií. V diskusiách rezonuje otázka znižujúcej sa stability rodín, ktorú so sebou prinášajú premenlivé životné a rodinné dráhy.

V roku 2019 vydalo Sociologické nakladateľstvo (SLON) v Prahe v koedícii s Výzkumným ústavem práce a sociálních věcí, v. v. i., dlho očakávanú publikáciu Věry Kuchařovej a kol. zameranú na zmapovanie súčasnej českej rodiny a zachytenie hlavných rysov a zásadných zmien, ktorými prešla od roku 2000.



Publikácia bola vytvorená v rámci projektu Ministerství práce a sociálních věcí ČR a nadväzuje na *Národní zprávu o rodině* (2004) a zároveň je voľným prepracovaním *Zprávy o rodině* (2017), ktorá tvorí výstup projektovej činnosti zamestnancov Výzkumného ústavu práce a sociálních věcí, v. v. i.. Publikácia je odpoveďou na vládou ČR definovanú potrebu vytvoriť súhrnnú štúdiu o „stave rodiny“ a je reakciou na dlhodobý deficit komplexnej analýzy životných podmienok a života rodín v ČR, ako aj na absenciu koncepcie rodinnej politiky založenej na takejto analýze. Iba na okraj uvádzame, že Koncepcia štátnej rodinnej politiky bola v SR spracovaná v roku 1996 a neskôr priebežne aktualizovaná.

Cieľom publikácie je vytvoriť portrét súčasnej českej rodiny, s dôrazom na trvalé črty a zásadnejšie zmeny vzťahujúce sa na rodiny v kontexte meniacich sa spoločensko-ekonomických a kultúrnych podmienok. Snahou autoriek bolo popísať súčasný stav v nadväznosti na minulé roky a zároveň upozorniť aj na charakteristiky rodín, ktoré sa ocitli v určitom vnútornom či sociálnom ohrození.

Kniha sa v prvých kapitolách zaoberá najmä premenou rodín a (nových) rizík pre rodiny, partnerstvom a rodičovstvom. Prvé kapitoly predkladajú široký záber na demografický kontext života rodín a autorky kladú dôraz na demografické trendy a fázovanie rodinného cyklu. Podrobne je rozobraná najmä problematika partnerstva a rodičovstva, teda tém ústredných pre založenie a fungovanie rodiny. Publikácia zachytáva celú škálu zmien najmä v demografickom kontexte (počet obyvateľov v ČR, veková štruktúra, sobášnosť a rozvodovosť, pôrodnosť a plodnosť, potratovosť, rodiny a domácnosti ich štruktúra, partnerstvo a jeho vznik a rozpad, rodičovstvo a plánovanie detí, bezdetnosť a otázky asistovanej reprodukcie) pôsobiach, ale najmä opisujúcich a umožňujúcich pochopiť trendy týkajúce sa zmien rodín, partnerstiev a rodičovstva. Premeny rodiny so sebou prinášajú určité riziká, ktoré môžu byť ešte nepoznané. Mierne sa zvyšuje intenzita sobášnosti, ale stále je pomerne nízka vzhľadom na zvyšujúci sa vek vstupu do prvého manželstva. Naďalej pretrváva vysoká intenzita rozvodovosti. Postupne sa zvyšuje priemerný vek matiek pri narodení prvého dieťaťa. Dlhodobou sa znižuje počet umelých prerušení tehotenstva ale zvyšuje sa počet samovoľných potratov. Rastie podiel detí narodených mimo manželstva a každá piata rodina v ČR je neúplná. Nezosobášené súžitie sa nepovažuje už za manželstvo na skúšku, ale je chápané ako trvalý vzťah. Medzi kľúčové faktory ovplyvňujúce založenie rodiny autorky zaraďujú samostatné bývanie a finančné zabezpečenie. Psychická nepripravenosť na rolu rodiča, ako aj neexistencia vhodného partnera sú pre mladých ľudí ďalšími dôvodmi odkladania reprodukcie. Ideálom väčšiny obyvateľov je stále model rodiny s dvoma deťmi. K uvedenej téme bola v roku 2014 vydaná v Čechách publikácia *Rodičovské dráhy. Dvacet let vývoja české porodnosti v sociologické perspektivě* (2014).

Ako sme už uviedli, finančné a bytové podmienky ovplyvňujú založenie rodiny ako aj jej neskoršie fungovanie. Horší príjem a materiálnu situáciu majú zväčša rodiny s vyšším počtom detí, nepracujúcimi rodičmi, rodičmi samoživiteľmi, rodičmi s nízkou vzdelanostnou úrovňou a nezosobášené páry. Príjmová chudoba zostáva na nízkej úrovni, cílenosť sociálnych dávok sa znižuje. Na druhej strane rastie význam netestovaných dávok pri redukcii chudoby, predovšetkým rodičovského príspevku ako významného zdroja príjmu. Náklady na bývanie pre väčšinu rodín s deťmi znamenajú veľkú alebo určitú záťaž. Náklady na vzdelanie zo všetkých nákladov predstavovali najmenší podiel. Čo sa týka výdavkov na pravidelné a nárazové voľnočasové aktivity, patria medzi najmenej dostupné. Zásadný problém bývania nie je celkový nedostatok bytov v ČR, ale ich náročná dostupnosť pre niektoré skupiny obyvateľov. Vyplýva to z poklesu počtu nájomných obecných a štátnych bytov a zároveň z rastúceho nájomného v porovnaní s nákladmi spojenými s vlastným bývaním. Záver kapitoly nezabúda ani na závažný problém a to priestorové vylúčenie niektorých skupín obyvateľstva a s tým súvisiace viacpočetné bariéry obmedzujúce riešenia.

Za dôležité autorky knihy považujú aj vzdelanie detí, ktoré je spoločnou úlohou spoločnosti, rodiny a nerodinne starostlivosti o deti. S tým súvisia zistenia ohľadom nárastu počtu materských škôl a detí, ktoré ich navštevujú. Na druhej strane sa systém zariadení dennej starostlivosti a výchovy malých detí nezlepšuje z hľadiska kvalitatívnych parametrov. Autorky konštatujú, že Česko je stále krajinou, kde veľmi silno závisí dosiahnuté vzdelanie od rodinného prostredia. Kapitola nezabúda ani na voľnočasové a záujmové aktivity, ako kľúčové pre individuálny rozvoj jedinca

a spôsob prevencie pred sociálnym vylúčením, avšak pre mnoho rodín sú príliš finančne nákladné a zlyháva ich podpora v rámci opatrení štátnej, lokálnej alebo regionálnej rodinnej politiky.

Autorky konštatujú, že rodina si zachováva vysokú hodnotu, ale podobný význam má aj profesia či zamestnanie. V niektorých etapách rodinného cyklu však zvládnutie rodinných a pracovných povinností naráža na subjektívne a objektívne bariéry, pričom však zladenie oboch sfér je predpokladom spokojnosti jednotlivca a stability rodiny, a preto sa autorky zaoberajú v kapitole *Sladovanie pracovného a rodinného života* otázkami súvisiacimi nielen s flexibilnými formami zamestnania ale aj predstavami o zosúladowaní a realitou napr. v dobe rodičovstva a genderovej deľby práce a kapitolu uzatvárajú štátnou podporou rodín so zamestnanými rodičmi.

V súčasnosti sa odborná literatúra stále viac zaoberá témou seniorov v spoločnosti. Práve siedma kapitola sa obsahom zameriava na skutočnosť, že rozvoľnením vekových noriem, väčšou disponibilitou prostriedkov asistovanej reprodukcie a čiastočne taktiež v dôsledku zvyšujúcej miery rozvodovosti a nových partnerstiev, rastie počet a podiel detí narodených starším matkám a otcom. Kapitola sa teda nevenuje iba rodinám so seniorom, seniorským partnerstvám, ale aj rodičovstvu vo vyššom veku a prarodičovstvu. V Sociologickom nakladateľstve SLON vyšli aj ďalšie knihy *Postarat se ve stáří* (2015) a *Prarodičovství v současné české společnosti* (2018), ktoré sa uvedenej problematike venujú hlbšie.

Jedna z posledných kapitol knihy je určená rodinám so špecifickými potrebami. Kapitola nie je príliš obsiahla a v krátkosti sa venuje predovšetkým problémom a podpore rodín so zdravotne postihnutým dieťaťom a rodinám migrantov. Obsah knihy uzatvára časť zachytávajúca reflexiu potrieb rodiny v rodinnej politike, jej konceptuálne smerovanie a inštitucionálne zabezpečenie. V tejto časti sa autorkám podarilo zachytiť rôznorodé názory a očakávania verejnosti od rodinnej politiky súvisiace s hľadaním stratégií ako zvládnuť súčasné riziká, v čom by mala rodinná politika rodinám pomôcť, zachovávajúc rešpekt k záujmom a potrebám individuálnych rodín.

Zvlášť významným prínosom publikácie je zachytenie procesov v rodine označovaných ako zmena „od inštitúcie ku vzťahu“. Okrem množstva štatistických a demografických dát je snahou autoriek pozrieť sa na rodinu aj v jej vzťahovom vnútri, čo preukazuje najmä kapitola s názvom *Vzťahy v rodině a výchova dětí*. Prináša niekoľko zaujímavých postrehov týkajúcich sa výchovy, vzťahov a rodín. Vyzdvihuje teóriu vzťahového trojuholníka otec – matka – dieťa. Dotýka sa najčastejších problémov v dnešných rodinách, najmä rozpadu rodiny (rozchod a rozvod). Poukazuje aj na špecifické problémy tzv. ohrozených rodín, kde preberá a komentuje výsledky špecializovaného výskumu českej nadácie *Sirius*. Ďalej sa venuje dieťaťu v ohrození – násiliu páchaného na deťoch v blízkych vzťahoch. Vo veľmi krátkej podkapitole sa zmieňuje o náhradnej (rodinnej) starostlivosti.

Záverom možno uviesť, že text je súhrnom poznatkov významných odborníkov zaoberajúcich sa z rôznych uhlov pohľadu problematikou súvisiacou s rodinou a rodinnou politikou. Zameriavajú sa primárne na aktuálnu situáciu a portrétovanie súčasného stavu, avšak v nadväznosti na vývoj v minulých rokoch. Využitie tohto postupu umožnilo zdôrazniť najnovšie vývojové trendy v sledovaných oblastiach.

Obsahovo sa publikácia zameriava na kľúčové témy, ktoré sa priamo dotýkajú života súčasných českých rodín a v ktorých dochádza k zásadnejším zmenám a/alebo sú predmetom odborných a verejných debát.

Kniha je určená podľa vydavateľa odbornej verejnosti, predstaviteľom štátnej správy a samosprávy alebo profesionálom v oblasti služieb pre rodiny. Môže byť aj významným zdrojom informácií pre študentov spoločenskovedných odborov. Zaujímavé súhrnné informácie o živote rodín v nej nájde i širšia verejnosť.

Kniha prináša obsiahle literárne a dátové zdroje, ako aj vecný register a pre prax zhrnutie množstva štatistických dát a demografických údajov, ktoré sú výsledkami rôznych výskumov realizovaných na tému rodiny v jej rôznorodosti v posledných rokoch v Česku. Je veľmi dobrým mostíkom k získaniu základného prehľadu, ako aj podkladov o rodine a k porovnaniu dát o rodine s inými krajinami. Autorky sa mnohým témam venujú naozaj iba okrajovo s upozornením na nové a špecifické otázky v daných oblastiach, ktorým by bolo vhodné sa venovať v rozsiahlejších a hlbších odborných výskumoch a štúdiách. Na porovnanie možno napríklad využiť dve obdobné publikácie zameriavajúce sa na slovenskú rodinu: *Rodinné správanie v demografických dátach* (2012) a *Rodina na Slovensku v teórii a vo výskume* (2015).

Mgr. Vladimíra LAUKO JACKOVÁ

Autorka je doktorandkou Sociologického ústavu SAV.

PRIPRAVUJEME/COMING SOON

Monotematické číslo Slovenskej štatistiky a demografie 3/2020 zamerané na viacrozmerné štatistické metódy.

*A monothematic issue of the Slovak Statistics and Demography No. 3/2020 focused on **multivariate statistical methods**.*

Hana ŘEZANKOVÁ

ZPŮSOBY VÝBĚRU VYSVĚTLUJÍCÍCH PROMĚNNÝCH V KLASIFIKAČNÍCH STROMECH
SELECTION METHODS EXPLANATORY VARIABLES IN CLASSIFICATION TREES

Viera LABUDOVA

POUŽITIE JEDNODUCHÝCH METÓD VIACROZMERNÉHO POROVNÁVANIA PRI ANALÝZE ZADLŽENOSTI KRAJÍN
THE USE OF SIMPLE METHODS OF MULTI-DIMENSIONAL COMPARISON IN THE ANALYSIS OF COUNTRIES' DEBT

Tatiana ŠOLTÉSOVÁ, Jana KÚTIKOVÁ

VYUŽITIE REGRESNEJ ANALÝZY PRI MODELOVANÍ ÚMRTNOSTI V ŽIVOTNOM POISTENÍ
USING OF REGRESSION ANALYSIS AT MODELING OF A MORTALITY IN A LIFE INSURANCE

Iveta STANKOVIČOVÁ, Alena MOJSEJOVÁ

INDEX ĽUDSKÉHO ROZVOJA: KLASIFIKÁCIA EURÓPSKYCH KRAJÍN POMOCOVI VIACROZMERNÝCH METÓD
HUMAN DEVELOPMENT INDEX: CLASSIFICATION OF EUROPEAN COUNTRIES BY MULTIVARIATE METHODS

* * *

ONLINE VERZIA ČÍSLA 2/2020 SLOVENSKEJ ŠTATISTIKY A DEMOGRAFIE JE VEREJNE DOSTUPNÁ na internetovej stránke ssad.statistics.sk od 15. APRÍLA 2020.

THE ONLINE VERSION OF THE JOURNAL SLOVAK STATISTICS AND DEMOGRAPHY No 2 (2020) IS PUBLICLY BE AVAILABLE at the website ssad.statistics.sk from **APRIL 15, 2020**.

INFORMÁCIE PRE PRISPIEVATEĽOV

Príspevky prijímame v slovenskom, v českom a v anglickom jazyku. Musia rešpektovať odborné zameranie časopisu a jeho vedecký charakter. Zaslaný príspevok nesmie byť v recenznom konaní v inom časopise, ani uverejnený v odbornej a inej tlači.

Príspevky zasielajte v elektronickej forme vo formáte MS Word alebo Open Office, typ písma Arial, veľkosť 12, riadkovanie 1. Nad titulkom treba uviesť meno autora a jeho pracovisko.

Súčasťou príspevku je abstrakt (základný popis cieľa a spôsobu spracovania faktov v rozsahu do 100 slov), kľúčové slová (maximálne 5), resumé (stručné zhrnutie obsahu článku s dôrazom na jeho prínos a najvýznamnejšie závery v rozsahu do 500 slov), profesijný životopis (v rozsahu do 120 slov) a kontakt (e-mailová adresa autora). Názov článku, abstrakt, kľúčové slová a resumé poskytne autor aj v anglickom jazyku. Zoznam použitej literatúry v abecednom poradí s úplnými bibliografickými údajmi sa uvádza na konci článku. Odkazy na literatúru sa uvádzajú v texte číslami v hranatých zátvorkách. Poznámky s poradovým číslom sú umiestnené pod čiarou na príslušnej strane textu, ku ktorému sa vzťahujú. Podrobnejšie pokyny nájdete autori na ssad.statistics.sk.

Maximálny rozsah vedeckých článkov je 15 normostrán, informatívnych článkov 6 normostrán, recenzie, rozhovory a informácie publikujeme v rozsahu maximálne 3 normostrany. Tabuľky, mapy, grafy a obrázky musia mať názov a uvedený zdroj údajov; odporúčame, aby kopírovali šírku textu. Skratky sa používajú len minimálne, pri prvom použití je potrebné skratku v zátvorke rozpísať. Redakcia zabezpečuje jazykovú úpravu textu.

Príspevky sú recenzované. Oponentské konanie je obojstranne anonymné. Konečné rozhodnutie o publikovaní článku vydáva redakčná rada.

Redakcia si vyhradzuje právo zverejniť články schválené redakčnou radou v tlačenej a elektronickej podobe na ssad.statistics.sk.

INFORMATION FOR AUTHORS

Articles are accepted in Slovak, Czech and English languages and must comply with the journal's professional specialisation and scientific nature as well. The submitted articles should not be reviewed by another journal and should not have already been published in any specialised or other press.

Please submit your articles in electronic form, in MS Word or Open Office format, Arial font, size 12 and typed in single spacing. The author's name and workplace should be indicated above the title.

Articles should contain an abstract (general description of the objective and the processing methods used up to 100 words), key words (max. 5), resume (brief summary of the article's content emphasizing its contribution and the most important conclusions up to 500 words), curriculum vitae of the author (no more than 120 words) and the author's contact (e-mail address). The author should submit the article's title, abstract, key words and resume in English language. List of the literature used with full bibliographic data should be given in alphabetical order at the end of an article. Bibliographic citations should be given in square brackets. References are indicated by numbers in a text in square brackets. Footnotes should be numbered in the order of the corresponding page of a text. Authors can find more details at the website ssad.statistics.sk.

Maximum scope of a scientific article is up to 15 standard pages, informative articles should be up to 6 standard pages in length, reviews, discussions and information not more than 3 standard pages. Tables, maps, graphs and pictures should have a title and the data source indicated, it is also advised to copy the width of a text. Abbreviations should be used only rarely and should be appropriately explained in parentheses when first used. Language text revisions are provided by the editorial office.

Articles are reviewed. The opponent procedure is mutually anonymous. The final decision on the article's publication is made by the editorial board.

The editorial office reserves the right to publish articles approved by the editorial board in printed and electronic form at the website ssad.statistics.sk.

je jediný recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov. Propagujeme miesto a význam slovenskej štatistiky v Európskom štatistickom systéme, spoluprácu Eurostatu a národných štatistických úradov pri harmonizácii zisťovaní a multidimenzionálny rozmer štatistiky. Podporujeme rozvoj štatistickej teórie a jej prepojenie s praxou. Naším cieľom je prispievať k využiteľnosti štatistických výstupov v rôznych oblastiach a k zvyšovaniu ich kvality a efektivity.

Publikujeme analytické články, prognózy, názory, diskusné príspevky, recenzie, rozhovory, informácie a oznamy z rôznych oblastí štatistiky (národné účty, produkčné štatistiky, sociálne štatistiky, štatistika životného prostredia a pod.) a demografie (demografická štatistika, teoreticko-metodologické východiská demografie, historická demografia a pod.), vrátane sčítania obyvateľov, domov a bytov ako neodmysliteľnej súčasti demografickej štatistiky.

Vydáva:

Štatistický úrad SR

Identifikačné číslo vydavateľa:

IČO 00166197

Vychádza:

Štyrikrát ročne

Dátum vydania:

15. apríl 2020

Tlač:

Reprografické stredisko
Štatistického úradu SR

Predplatné:

20 € (na rok)

5 € (za jeden výtlačok)

Objednávky prijíma:

Informačný servis
Štatistického úradu SR
Tel.: +4212/502 36 339
+4212/502 36 335
E-mail: info@statistics.sk

is the only scientific reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures. Our aim is to promote the position and importance of Slovak statistics in the European Statistical System, cooperation between the Eurostat and the national statistical offices in the field of survey harmonisation and the multidimensional character of statistics as well. We support the development of statistical theory and its connection with practice. We aim to contribute to the utility of statistical outputs in various fields and to the improvement of quality and efficiency.

We publish analytic articles, prognoses, views, discussion contributions, reviews, discussions, information and announcements from various statistical fields (national accounts, production statistics, social statistics, environmental statistics etc.) and demography (demographic statistics, theoretical and methodological bases of demography, historical demography etc.) including the population and housing census as an essential part of demographic statistics.

Issued by:

Statistical Office of the SR

Company registration number:

00166197

Published:

Four times a year

Date of issue:

15th April 2020

Press:

Reprographic centre of the
Statistical Office of the SR

Subscription:

€20 (per year)

€5 (for one copy)

Orders are to be addressed to:

Information Service of the
Statistical Office of the SR
Tel.: +4212/502 36 339
+4212/502 36 335
E-mail: info@statistics.sk