

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

Evidenčné číslo: 103006/D/2022/36086129769085956

MODELOVANIE ÚMRTNOSTI
NA SLOVENSKU Z POHLADU AKTUÁRA

Dizertačná práca

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

Katedra matematiky a aktuárstva

MODELOVANIE ÚMRTNOSTI
NA SLOVENSKU Z POHLADU AKTUÁRA

Dizertačná práca

Študijný program: 3.2 Kvantitatívne metódy v ekonómii

Študijný odbor: 8 Ekonómia a manažment

Školiace pracovisko: Katedra matematiky a aktuárstva

Školiteľ: doc. Mgr. Tatiana Šoltésová, PhD.

Bratislava 2022

Ing. Jana Kútiková

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Ing. Jana Kútiková
Študijný program: kvantitatívne metódy v ekonómii (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: 8. - ekonómia a manažment
Typ záverečnej práce: Dizertačná záverečná práca
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Modelovanie úmrtnosti na Slovensku z pohľadu aktuára

Cieľ: Cieľom dizertačnej práce je sumarizovať metodológiu deterministického a stochastického modelovania úmrtnosti populácie, ktoré vyžadujú pochopenie spoločných črt úmrtnostného správania populácie. Naštudované metódy a postupy využiť v analýze úmrtnosti populácie na Slovensku podľa veku a pohlavia a poukázať na ďalší vývoj úmrtnosti. Cieľom práce je tiež porovnať kvalitatívne stránky jednotlivých spôsobov modelovania úmrtnosti s dôrazom na využitie v aktuárstve a vytvoriť postup odhadovania budúcich mier úmrtnosti pre potreby minimalizovania rizika úmrtnosti a dlhovekosti v životnom poistení.

Anotácia:

1. Charakterizovať úmrtnosť vo vzťahu k aktuárstvu.
2. Analyzovať úmrtnostné správanie populácie Slovenskej republiky a opísať problematiku vplyvu rovnosti pohlaví na poisťovníctvo s uvedením vlastného stanoviska.
3. Opísať deterministické modely úmrtnosti, tzv. zákony úmrtnosti s rôznym počtom parametrov a opísať ich vhodnosť pre jednotlivé vekové kategórie populácie.
4. Opísať stochastické modely úmrtnosti (Lee Carterov model, Cairns-Blake-Dowdov model a ich modifikácie).
5. Aplikovať vybrané modely na úmrtnostné správanie populácie Slovenska a posúdiť ich vhodnosť vo vybranom softvéri.
6. Prognózovať úmrtnosť populácie Slovenska na krátke časové obdobie a testovať kvalitu prognózovania modelov.
7. Poukázať na dôležitosť odhadu mier úmrtnosti pre vybrané druhy životného poistenia z pohľadu minimalizácie rizika úmrtnosti a dlhovekosti.

Školiteľ: prof. RNDr. Ľudovít Pinda, CSc.
Školiteľ: doc. Mgr. Tatiana Šoltésová, PhD.
Katedra: KMA FHI - Katedra matematiky a aktuárstva FHI
Vedúci katedry: doc. Ing. Michal Páleš, PhD.

Dátum zadania: 06.11.2020

Dátum schválenia: 20.10.2020

prof. RNDr. Katarína Sakálová, CSc.
predseda odborovej komisie

Pod'akovanie

Touto cestou si dovoľujem poďakovať školiteľke doc. Mgr. Tatiane Šoltésovej, PhD. za profesionálne vedenie, odbornú pomoc, užitočné rady a pripomienky, za ochotný a ústretový prístup počas celého doktorandského štúdia. Moje poďakovanie patrí aj celej rodine za veľkú podporu, pomoc a povzbudenie počas celého štúdia.

Abstrakt

KÚTIKOVÁ, Jana: *Modelovanie úmrtnosti na Slovensku z pohľadu aktuára* – Ekonomická univerzita v Bratislave. Fakulta hospodárskej informatiky; Katedra matematiky a aktuárstva. – Školiteľ záverečnej práce: doc. Mgr. Tatiana Šoltésová, PhD. – Bratislava: FHI EU, 2022, 157 strán.

Cieľom dizertačnej práce je nahliadnuť do problematiky modelovania úmrtnosti pomocou deterministických a stochastických modelov, ich aplikovateľnosti na úmrtnostné správanie populácie Slovenska a tiež na ich využitie v oblasti aktuárstva. Predmetom modelovania sú vekovo špecifické miery úmrtnosti mužov a žien v rozpätí od 0 do 100 rokov v rokoch od vzniku Slovenskej republiky (1993) až po súčasnosť vypočítané pomocou údajov o úmrtnosti získaných zo Štatistického úradu Slovenskej republiky. Práca je členená na štyri kapitoly. Začiatok práce je venovaný základným pojmom aktuárskej bázy so zameraním na úmrtnosť. Okrem teoretického vymedzenia úmrtnosti a jej ukazovateľov, táto kapitola obsahuje aj vývoj úmrtnosti medzi rokmi 1993 až 2020 a predstavenie legislatívnej úpravy poisťovníctva v Európskej únii. Druhá kapitola opisuje ciele dizertačnej práce. Tretia kapitola je zameraná na metodiku dizertačnej práce, ktorá pozostáva z opisu vybraných deterministických a stochastických modelov, tiež použitých metód a prostriedkov na ich modelovanie a porovnávanie. Na modelovanie sme použili metódy nelineárnej regresie a metódu najmenších štvorcov za pomoci singulárneho rozkladu matice a na prognózovanie sme využili ARIMA modely. Štvrtá a zároveň posledná kapitola pozostáva z aplikácií vybraných modelov na populáciu Slovenska a ich vzájomného porovnania. Na porovnanie kvality modelov sme využili štandardnú odchýlku rezíduí RMSE. V závere práce analyzujeme, aký dopad má výber konkrétneho modelu na poisťné sadzby, ale aj vplyv zavedenia rodovej rovnosti a efekt zhoršenej úmrtnosti v dôsledku koronavírusu na dočasné poistenie na úmrtie a dočasný predľhotný dôchodok.

Kľúčové slová: životné poistenie, úmrtnosť, parametrické modely úmrtnosti, nelineárna regresia, metóda najmenších štvorcov, SAS Enterprise Guide

Abstract

KÚTIKOVÁ, Jana: *Mortality modelling in Slovakia from the point of view of the actuary* – University of Economics in Bratislava. Faculty of Economic Informatics; Department of Mathematics and Actuarial Science – Thesis supervisor: doc. Mgr. Tatiana Šoltésová, PhD. – Bratislava: FHI EU, 2022, 157 pages.

The aim of this dissertation thesis is to look into the problematic of mortality modelling using deterministic and stochastic models, their applicability to the mortality behaviour of the population of Slovakia and their use in the field of actuarial science. The subjects of modelling are age-specific mortality rates for men and women in the range from 0 to 100 years from the establishment of the Slovak Republic (1993) till now, calculated using mortality data obtained from the Statistical Office of the Slovak Republic. The thesis is divided into four chapters. The beginning of the thesis is dedicated to the basic terms related to the actuarial basis with a focus on mortality. In addition to the theoretical definition of mortality and its indicators, this chapter also includes the development of mortality between 1993 and 2020 and the introduction of insurance legislation in the European Union. The second chapter describes the aim of the dissertation thesis. The third chapter is focused on the methodology of the dissertation thesis, which consists of a description of selected deterministic and stochastic models, as well as the methods and tools used for their modelling and comparison. We used nonlinear regression and least squares methods with singular value decomposition for modelling, and we implement ARIMA models for the forecasting. The fourth and at the same time the last chapter consists of applications of selected models to the population of Slovakia and their comparison. We used the root mean square error RMSE to compare the quality of the models. At the end of the dissertation thesis, we analyse the impact of the model selection on insurance rates, but also the effects of gender directive and the results of impaired coronavirus mortality on the temporary term life insurance and temporary annuity insurance.

Key words: life insurance, mortality, parametric mortality models, nonlinear regression, least squares method, SAS Enterprise Guide

Obsah

Zoznam skratiek.....	9
Úvod.....	11
1 Súčasný stav riešenej problematiky doma a v zahraničí.....	15
1.1 Aktuárska báza	15
1.1.1. Úroková miera	16
1.1.2. Úmrtnosť.....	17
1.1.3. Náklady.....	18
1.1.4. Inflácia	18
1.1.5. Zrušenie alebo odkup poisťky	19
1.1.6. Investičný výnos	20
1.2 Úmrtnosť a jej ukazovatele	20
1.2.1 Analýza úmrtnosti na Slovensku	23
1.3 Legislatívna úprava poisťovníctva Európskej Únie.....	26
2 Cieľ práce.....	33
3 Metodika práce a metódy skúmania	35
3.1 Modely úmrtnosti.....	35
3.1.1 Deterministické modely – zákony úmrtnosti	36
3.1.1.1 Gompertzov - Makehamov zákon úmrtnosti.....	38
3.1.1.2 Kannistov - Makehamov zákon úmrtnosti	39
3.1.1.3 Heligmanov - Pollardov zákon úmrtnosti	40
3.1.1.4 CoDe zákon úmrtnosti.....	41
3.1.2 Stochastické modely	43
3.1.2.1 Lee-Carter trieda modelov	45
3.1.2.2 CBD-Perks trieda modelov	47
3.1.2.3 Ďalšie stochastické modely	48
3.2 Metódy odhadu parametrov deterministických modelov úmrtnosti	49
3.2.1 Gaussova - Newtonova metóda	50
3.2.2 Gradientná metóda.....	51
3.2.3 Levenbergova - Marquardtova metóda	52
3.3 Metódy odhadu parametrov stochastických modelov úmrtnosti	54
3.3.1 Metóda najmenších štvorcov	54

3.3.2	Prognózovanie – ARIMA modely	61
3.4	Meranie kvality modelov a presnosti ich prognóz	65
3.5	Niektoré základné druhy poisťných produktov	67
3.6	Softvér SAS Enterprise Guide.....	68
4	Výsledky práce a diskusia.....	83
4.1	Deterministické modely úmrtnosti	85
4.1.1	Gompertzov – Makehamov model úmrtnosti.....	85
4.1.2	Kannistov – Makehamov model úmrtnosti	90
4.1.3	Heligmanov - Pollardov model úmrtnosti	95
4.1.4	CoDe model úmrtnosti.....	103
4.2	Stochastické modely úmrtnosti	111
4.2.1	Lee-Carterov model úmrtnosti.....	111
4.2.2	Cairns-Blake-Dowdov model úmrtnosti.....	119
4.3	Porovnanie modelov úmrtnosti na základe RMSE	127
4.4	Analýza vplyvov na stanovenie poisťných sadzieb	131
4.4.1	Analýza vplyvu výberu modelu.....	132
4.4.2	Analýza vplyvu gender smernice	134
4.4.3	Analýza vplyvu koronavírusu.....	138
	Záver	143
	Zoznam obrázkov, tabuliek a grafov	147
	Zoznam použitej literatúry	151
	Prílohy.....	159

Zoznam skratiek

<i>AIC</i>	Akaikovo informačné kritérium
<i>ALS</i>	alternatívna metóda najmenších štvorcov
<i>AR</i>	autoregresný člen ARIMA modelu
<i>ARIMA</i>	autoregresný integrovaný kľzavý priemer
<i>BIC</i>	Bayesovské informačné kritérium
<i>CBD</i>	Cairns-Blake-Dowdow model
<i>CD</i>	CoDe model
<i>EÚ</i>	Európska únia
<i>gender smernica</i>	smernica Rady 2004/113/ES z 13. decembra 2004 o vykonávaní zásady rovnakého zaobchádzania medzi mužmi a ženami v prístupe k tovaru a službám a k ich poskytovaníu
<i>GM</i>	Gompertzov-Makehamov model
<i>GMM</i>	general model measurement
<i>HMD</i>	Human mortality database
<i>HP</i>	Heligmanov-Pollardov model
<i>I</i>	integračný člen ARIMA modelu
<i>IASB</i>	Rada pre medzinárodné štandardy
<i>IFRS</i>	Medzinárodný štandard finančného výkazníctva
<i>KM</i>	Kannistov-Makehamov model
<i>KPSS test</i>	Kwiatkowski-Phillips-Schmidt-Shin test
<i>LC</i>	Lee-Carterov model
<i>MA</i>	člen kľzavých priemerov ARIMA modelu
<i>MSE</i>	priemerný súčet štvorcov rezíduí
<i>OLAP</i>	Online Analytical Processing
<i>OLS</i>	obyčajná metóda najmenších štvorcov
<i>OXERA</i>	Oxfordská asociácia ekonomického výskumu
<i>PAA</i>	Premium allocation approach
<i>PLS</i>	parciálna metóda najmenších štvorcov
<i>RMSE</i>	štandardná odchýlka rezíduí
<i>SAS EG</i>	Statistical Analysis Software Enterprise Guide

<i>SFCR</i>	Správa o solventnosti a finančnom stave poisťovne
<i>SLASPO</i>	Slovenská asociácia poisťovní
<i>SVD</i>	Singulárny rozklad matice
<i>ŠÚ SR</i>	Štatistický úrad Slovenskej republiky
<i>VACF</i>	výberová autokorelačná funkcia
<i>VFA</i>	Variable fee approach
<i>VIACF</i>	inverzná výberová autokorelačná funkcia
<i>VPACF</i>	parciálna výberová autokorelačná funkcia
<i>WLS</i>	vážená metóda najmenších štvorcov

Úvod

Úmrtnosť obyvateľstva sa neustále vyvíja. Kvalita života sa každým rokom zvyšuje, čo sa samozrejme odráža na vývoji úmrtnosti, ktorá sa za posledné desaťročia výrazne zlepšila. Je to okrem iného spôsobené technickým pokrokom v medicíne, sociológii, farmácii a iných vedách, ktoré priamo alebo nepriamo súvisia so smrťou. Vekové zloženie úmrtnosti obyvateľstva má vo väčšine krajín Európskej únie podobný priebeh. Dojčenská úmrtnosť je pomerne vysoká, následne ostro poklesne v skorom detstve, postupne začína prudko narastať v priebehu dospievania a neskôr už len exponenciálne rastie, pričom miera nárastu v starobe sa postupne znižuje. Parametrické modely úmrtnosti sa pokúšajú toto úmrtnostné správanie zachytiť pomocou matematického modelu s vymedzeným počtom parametrov. Prvé pokusy o modelovanie úmrtnosti boli založené na deterministických scenároch. To viedlo k vytvoreniu zákonov úmrtnosti, ktoré opisujú úmrtnosť v pevnom časovom okamihu, a tým podceňujú zlepšenie úmrtnosti, pretože skutočná úmrtnosť je stochastická. Vybranými deterministickými modelmi úmrtnosti sa zaoberajú napríklad Gavrilova a Gavrilov (2011) či Forfar (2014), ktorí sa okrem iného venovali porovnaniu týchto modelov. S problematikou využitia deterministických modelov úmrtnosti vo vyšších vekoch sa postupne zaoberali Perks (1932), Beard (1959), Coale a Kisker (1990), Kannisto (1994), Thatcher a kol. (1998), Gampe (2010) a napokon Saikia a Borah (2014). O aplikáciu zákonov úmrtnosti na údaje o úmrtnosti konkrétnej populácie sa pokúsil napríklad Debón a kol. (2005), ktorý sa zameral na španielsku populáciu alebo Ibrahim (2017), ktorý použil Heligmanov-Pollardov model na populáciu Malajzie. Bardoutsos, de Beer a Janssen (2016, 2018) v niekoľkých publikáciách predstavili najnovší parametrický model s rôznymi variáciami, ktorý nazývajú CoDe model a význam tohto modelu spočíva v rozlišovaní efektov zhutňovania a oneskorenia úmrtnostného správania ľudskej populácie. Ďalšie informácie o deterministických modeloch úmrtnosti sa môže čitateľ dozvedieť v publikačných zdrojoch: Cipra (2006), Šoltésová (2019), Dicson a kol. (2009), Pitacco (2016), Promislow (2015), Remund a kol. (2018) a Rotar (2014). Objavilo sa aj pár pokusov o vytvorenie stochastického modelu z deterministických zákonov úmrtnosti. Úsilie McNowna a Rogersa (1989) bolo úspešné a podarilo sa im rozšíriť Heligman-Pollardov deterministický model úmrtnosti (Heligman a Pollard, 1980) o stochastické predpoklady, čím vytvorili jeden

zo stochastických modelov úmrtnosti. Stochastické modely sa od deterministických odlišujú tým, že berú do úvahy vývoj úmrtnosti v čase. Existuje niekoľko stochastických modelov úmrtnosti, ale všeobecne ich môžeme rozdeliť do troch tried. Prvou je trieda modelov Lee-Cartera, ktorej hlavným predstaviteľom je Lee-Carterov model. Výhodou tohto modelu je jeho jednoduchosť, vďaka ktorej sa stal veľmi populárnym. Používa sa na modelovanie a predpovedanie úmrtnosti v mnohých krajinách, ako sú USA (Lee a Carter, 1992), Čína (Jiang, 1995), Japonsko (Wilmoth, 1996), India (Yadav et al., 2012), škandinávské krajiny (Kossi et al., 2006) a mnoho ďalších. Druhou triedou modelov sú P-spliny, ktoré sa bežne používajú vo Veľkej Británii (Currie, 2006), ale v iných častiach sveta už toľko nie. Posledná je trieda modelov CBD-Perks s hlavným zástupcom, a to Cairns-Blake-Dowdovým modelom. Analýzu kvalitatívnych a kvantitatívnych charakteristík všetkých troch tried modelov vykonali Cairns a kol. (2011). Porovnaniu modelov Lee-Cartera a Cairns-Blake-Dowda sa venoval napríklad Nocito (2015), ktorý aplikoval tieto dva modely na taliansku populáciu. Detailnú analýzu CBD modelu vykonali napríklad Chan a kol. (2014) a dospeli k záveru, že parametre CBD modelu sú schopné reprezentovať úmrtnostné správanie populácie v spojitosti so zlepšovaním úmrtnosti do budúcnosti, parametre odhadnuté v predchádzajúcich rokoch sú invariantné voči novým údajom a taktiež sú interpretovateľné. Prepojeniu stochastických modelov úmrtnosti a aktuárstva sa venoval napríklad Koller (2012). Parametrické modely môžu vo všeobecnosti nájsť využitie pri konštrukcii úmrtnostných tabuliek, na graduáciu špecifických mier úmrtnosti alebo na redukciu štatistických chýb, prípadne na porovnávanie úmrtnosti jednotlivých krajín a na prognózovanie očakávaného vývoja úmrtnosti.

Úmrtnosť je významným predpokladom aktuárskej bázy. Dobrým zdrojom informácií o aktuárskej báze je napríklad publikácia Sakálová (2006). Dramatický pokles úmrtnosti predstavuje veľmi vážne finančné riziká pre životné poisťovne, pre ktoré je znalosť úmrtnostného správania obyvateľstva v závislosti od veku kľúčovým faktorom mnohých poisťno-matematických výpočtov, ako je výpočet poisťného a technických rezerv. Čím presnejšie je možné odhadnúť pravdepodobnosť úmrtia poistencov, tým nižšie je riziko nesprávneho ocenenia poisťnej zmluvy (resp. nevhodného uloženia finančných prostriedkov). Poisťovne pri svojich výpočtoch zohľadňujú dlhodobé pozorovania, ktorých základom sú úmrtnostné tabuľky. Tieto úmrtnostné tabuľky si poisťovňa vytvára na základe informácií o vývoji úmrtnosti vo vlastnom portfóliu poisťných zmlúv, alebo tiež môže použiť

údaje o úmrtnostnom správaní populácie v danej krajine, ktoré sú dostupné na Štatistickom úrade príslušnej krajiny. Informácie o úmrtnostných tabuľkách môžeme nájsť napríklad v prácach Sekerovej a Bilíkovej (2005), Špírkovej a Urbaníkovej (2012). Sivašová (2008) vo svojej monografii opisuje pravdepodobnostné modely úmrtnosti, spôsoby konštrukcie úmrtnostnej tabuľky a metódy ich graduácie. Rovnako sa touto problematikou konštrukcie a vyrovňovania úmrtnostných tabuliek zaoberá Cipra (2006) a Mészáros (2016).

Dizertačná práca sa venuje problematike modelovania úmrtnosti pomocou deterministických a stochastických modelov, ich aplikovateľnosti na úmrtnostné správanie populácie Slovenska a tiež ich využitiu v oblasti aktuárstva. Predmetom modelovania sú vekovo špecifické miery úmrtnosti, prípadne pravdepodobnosti úmrtia mužov a žien v rozpätí od 0 do 100 rokov v čase od vzniku Slovenskej republiky (1993) až po súčasnosť. Na dosiahnutie cieľov sme použili predovšetkým analytický softvér SAS Enterprise Guide. Ako pomocné nástroje nám poslúžili aj softvérové rozhranie R a Microsoft Excel.

Dizertačná práca je obsahovo členená na teoretickú a praktickú časť. Teoretická časť pozostáva z troch kapitol. Prvá kapitola sa venuje základným pojmom aktuárskej bázy so zameraním na úmrtnosť, ktorá je predmetom tejto záverečnej práce. V tejto časti priblížime okrem teoretického vymedzenia úmrtnosti aj jej základné ukazovatele a faktory, ktoré ju ovplyvňujú. Analyzujeme tiež vývoj úmrtnosti v období od vzniku Slovenskej republiky (1993) až po súčasnosť s ohľadom na koronavírus, ktorý mal za následok prudký nárast úmrtnosti v posledných rokoch. V závere kapitoly sa pozrieme na legislatívnu úpravu poisťovníctva v Európskej únii, ktorá má vplyv aj na stanovenie predpokladu úmrtnosti pri aktuárskych výpočtoch (Solvency II, gender smernica, IFRS 17). Druhá kapitola slúži na vymedzenie hlavných a vedľajších cieľov dizertačnej práce. Tretia kapitola je zameraná na metodiku dizertačnej práce. V rámci nej sa zoznámime so širokým spektrom deterministických a stochastických modelov, od jednoduchších po zložitejšie a od najstarších po najmladšie. Našou základnou úlohou je odhadnúť parametre vybraných modelov, pričom využijeme analytický softvér SAS Enterprise Guide. V tejto kapitole sa preto venujeme aj metódam odhadu parametrov deterministických a stochastických modelov, ktoré sú v softvéri dostupné. Pri deterministických modeloch využívame na odhad parametrov metódy nelineárnej regresie, ktorými sú Gaussova-Newtonova metóda, gradientná metóda a v neposlednom rade Levenbergova-Marquardtova metóda. Pri stochastických modeloch aplikujeme metódu najmenších štvorcov za pomoci

singulárneho rozkladu matice a Habermanovho-Russolillovho prístupu k odhadu parametrov. Jednou zo základných črt stochastického modelovania je prognózovanie. Existuje niekoľko metód predikcie, jednou z najpopulárnejších je ARIMA modelovanie (autoregresný integrovaný kľzavý priemer). Ďalej sa v tejto kapitole zaoberáme mierami kvality modelov a presnosti ich prognózovania – priemerný súčet štvorcov rezíduí, štandardná odchýlka rezíduí, Akaikovo informačné kritérium a Bayesovské informačné kritérium, ktoré slúžia pre nás ako nástroj na vzájomné porovnanie modelov v praktickej časti ako aj na porovnanie prognostických schopností stochastických modelov. Na konci kapitoly predstavíme softvérové rozhranie SAS Enterprise Guide so zameraním na procedúru nelineárnej regresie a ARIMA modelovania.

Praktickej časti je venovaná posledná kapitola tejto záverečnej práce a je priamo zameraná na aplikáciu deterministických a stochastických modelov úmrtnosti, ich vzájomné porovnanie a analýzu ich dopadu, ale aj na analýzu ďalších vonkajších vplyvov (gender smernice a koronavírusu) na poisťné sadzby. Hlavným cieľom je porovnanie projekčných vlastností jednotlivých úmrtnostných modelov pre dáta Slovenskej republiky vo vekovom intervale od 0 do 100 rokov v závislosti od druhu modelu. Vychádzame z údajov o úmrtnosti obyvateľstva Slovenskej republiky v rokoch 1993 až po najaktuálnejšie získané údaje z databázy Štatistického úradu Slovenskej republiky. Praktická časť je rozdelená na niekoľko podkapitol. Najskôr sa venujeme samostatne vybraným deterministickým a stochastickým modelom. Potom sa pokúsime na základe mier presnosti nájsť model, ktorý najlepšie zodpovedá reálnym dátam úmrtnosti populácie v Slovenskej republike zvlášť pre mužov a ženy, ale aj so zanedbaním pohlavia ako faktora, ktorý vplyva na úmrtnosť. Napokon sa pokúsime analyzovať, aký dopad má výber konkrétneho modelu na poisťné sadzby, ale aj aký vplyv má zavedenie rodovej rovnosti a aký vplyv má efekt zhoršenej úmrtnosti v dôsledku koronavírusu na konci sledovaného obdobia na deterministický a stochastický model. Pri analýze budeme vychádzať zo základných a zároveň najviac kupovaných poisťných produktov, ktorými sú dočasné poistenie na úmrtie a dočasný predľhotný dôchodok. K základným produktom patrí aj zmiešané poistenie, ktoré ale nie je až tak ovplyvnené rizikom úmrtnosti, preto sme ho do našej analýzy nezaradili.

Pevne veríme, že spracovaná dizertačná práca prispeje k rozšíreniu vedomostí z oblasti modelovania úmrtnosti a zároveň poukáže na to, že štatistika a aktuárska demografia sú dôležitou súčasťou aktuárskych vied.

1 Súčasný stav riešenej problematiky doma a v zahraničí

Úmrtnosť je dôležitou súčasťou aktuárskej bázy. Pre životnú poisťovňu je rozhodujúce poznať pravdepodobnosti úmrtia populácie v rôznych vekových skupinách, pretože na základe nich vykonáva aktuárske výpočty a ich nesprávne určenie môže mať nepriaznivý vplyv na finančný stav poisťovne.

Táto kapitola je venovaná aktuárskej báze so zameraním na úmrtnosť a jej ukazovatele. Obsahuje tiež analýzu úmrtnosti populácie Slovenska od vzniku samostatnej Slovenskej republiky (1993) až po súčasnosť. Ďalej sa v kapitole zaoberáme legislatívnou úpravou poisťovníctva. Najzásadnejší vplyv v tomto smere majú pre životnú poisťovňu smernice Solvency II, gender smernica a najnovšie tiež štandard IFRS 17.

1.1 Aktuárska báza

Súbor predpokladov slúžiacich na realizovanie komplexných poistno-matematických výpočtov sa nazýva aktuárska báza. Pod komplexnými poistno-matematickými výpočtami rozumieme analýzy ziskovosti poisťovne, oceňovanie produktov, projekciu peňažných tokov, najlepší odhad technických rezerv a mnohé ďalšie, pričom aktuárska báza sa mení v súlade s cieľom, na ktorý je použitá. Vo všeobecnosti môžeme aktuársku bázu rozdeliť na dva typy, a to poistnú bázu a bázu ohodnocovania. Poistnú bázu tvoria predpoklady, na základe ktorých poisťovne určujú výšku poistného. Báza ohodnocovania predstavuje predpoklady na oceňovanie aktív a pasív životnej poisťovne. Medzi fundamentálne zložky poistnej bázy patrí úroková miera, úmrtnosť a náklady. Pri metóde peňažných tokov do aktuárskej bázy vstupujú okrem uvedených aj ďalšie predpoklady. Najčastejšie sa jedná o infláciu a mieru odkupov, alebo napríklad pri produktoch s podielom na zisku je dôležitou súčasťou poistnej bázy miera podielu poisteného na zisku životnej poisťovne.

Ak by sme sa pozreli na tieto predpoklady z iného uhla, a to podľa toho, či súvisia so samotným vývojom trhu alebo či zohľadňujú aktuálny vývoj v portfóliu životných poistení, môžeme ich rozdeliť na ekonomické (úroková miera, inflácia, investičný výnos) a neekonomické predpoklady (úmrtnosť, chorobnosť, invalidita, stornovanosť, náklady a provízie).

1.1.1. Úroková miera

K najrozhodujúcejším finančným predpokladom zaradíme úrokovú mieru. Môžeme ju zdefinovať ako priemernú návratnosť investícií poisťovne. Pre poisťovňu je nevyhnutné okrem správneho výberu finančných predpokladov aj čo najpresnejšie odhadnutie ich vývoja do budúcnosti. Nesprávne očakávanie vývoja predpokladov môže mať negatívny vplyv nielen na ziskovosť daného produktu, ale aj na finančný stav poisťovne ako celku.

Ak zvolíme hodnoty úrokovej miery príliš nízke, podhodnocujeme výšku úrokovej miery a môže nastať situácia, kedy skutočný zisk poisťovne bude vyšší ako očakávaný. V takomto prípade nazývame predpoklad o úrokovej miere pesimistický alebo inak konzervatívny. Naopak v prípade, že nastane situácia, že skutočný zisk bude nižší, ako sme očakávali, tak sme zvolili predpoklady o úrokovej miere príliš vysoké (nadhodnotili sme ich). Taký predpoklad označujeme ako optimistický a charakterizuje ho nižšia tvorba rezerv v rezervotvorných poisteniach. Realistický predpoklad sa vyskytuje niekde medzi pesimistickým a optimistickým predpokladom o úrokovej miere.

Na výber úrokovej miery pri oceňovaní produktov vplyvajú rôzne faktory. Príkladom je druh kontraktu alebo citlivosť kontraktu na úrokovú mieru. Napríklad pri produktoch s garantovanou výplatom poistného plnenia by sme mali zvoliť investície, ktoré sú bezpečné a s veľkou pravdepodobnosťou zaručujú výnos zodpovedajúci garantovanej poistnej sume. Niektoré typy produktov sú extrémne citlivé na zmeny úrokovej miery. Sú to najmä rezervotvorné produkty, medzi ktoré patrí aj zmiešané poistenie. Výber úrokovej miery závisí ďalej od doby trvania poistenia a spôsobu platenia poistného. Pri poistných zmluvách s kratšou dobou splatnosti vie poisťovňa presnejšie odhadnúť zhodnotenie svojich investícií. To isté platí aj pri produktoch s jednorazovým platením poistného, kedy poisťovňa vie presne výšku investovaného poistného a je tiež známa úroková miera, pri ktorej môže tieto prostriedky investovať. K ďalším faktorom výberu úrokovej miery, ktoré je nutné spomenúť, patria miera reinvestície v budúcnosti a názor na súčasnú hodnotu úrokovej miery štátnych dlhopisov, na základe ktorých sa poisťovňa snaží prognózovať výkyvy v úrokovej miere, a tiež frekvenciu týchto výkyvov v budúcnosti.

Výber vhodnej úrokovej miery sa podľa Sakálovej a Kotlárovej (2015) realizuje v nasledujúcich dvoch krokoch:

- a) **Výber vhodných typov aktív** – rozličné typy aktív predstavujú pre poisťovňu rôzne úrovne rizík, a taktiež rôznu výnosnosť. Životné poisťovne najčastejšie investujú do akcií podnikov, štátnych dlhopisov, hotovostných depozít a do nehnuteľností. Tieto typy aktív sa zväčša vyznačujú vysokou likviditou, a tiež by mali byť pomerne bezpečné s ohľadom na investičnú stratégiu životnej poisťovne, ktorá závisí najmä od voľných aktív, ktorými disponuje.
- b) **Určenie realistickej úrokovej miery pre typy aktív, ktoré životná poisťovňa preferuje.**

1.1.2. Úmrtnosť

Ďalším významným predpokladom aktuárskej bázy je úmrtnosť. Očakávania o úmrtnostnom správaní poisťného kmeňa úzko súvisia s typom produktu životnej poisťovne. Podhodnotenie úmrtnosti má negatívny vplyv pri dočasnom poistení na úmrtie a naopak, v prípade dôchodkového poistenia môže práve nadhodnotenie úmrtnosti viesť k vysokým stratám poisťovne. Poisťovňa pri modelovaní úmrtnosti používa tzv. úmrtnostné tabuľky, ktoré môžeme vo všeobecnosti rozdeliť na generačné a prierezové. Tvorba generačných úmrtnostných tabuliek je veľmi časovo náročná, keďže obsahujú informácie o úmrtnostnom správaní kohorty od narodenia až po smrť posledného zo skupiny obyvateľstva. V poisťovníctve sa zväčša nepoužívajú najmä kvôli časovej náročnosti. Namiesto toho poisťovne odhadujú úmrtnosť na základe vývoja vlastného poisťného kmeňa, pričom výsledky analýz upravujú na základe populačných úmrtnostných tabuliek. Pri prierezových úmrtnostných tabuľkách sa sleduje úmrtnosť danej populácie za krátke časové obdobie. Na základe mier úmrtnosti sa následne zostrojí hypotetický obraz života osôb narodených v jednom roku. Okrem týchto druhov úmrtnostných tabuliek existujú aj selekčné tabuľky, v ktorých je zaznamenané úmrtnostné správanie obyvateľstva na základe spoločných vlastností (znakov), akými sú napríklad vek, fajčenie, pohlavie, upisovacie normy, spôsob predaja poistenia alebo druh poisťného kontraktu. Vzhľadom na to, že podstata témy tejto práce je založená na predpoklade úmrtnosti, budeme sa jej podrobnejšie venovať v nasledujúcich častiach práce.

1.1.3. Náklady

Úlohou poisťovne v súvislosti s analýzou nákladov je čo najpresnejšie určenie predpokladanej výšky nákladov v budúcnosti, vychádzajúc zo skúseností z minulosti a plánovaných objemov predaja v blízkej budúcnosti, a ich následná alokácia podľa použitia. Správna alokácia nákladov vedie k optimálnym výsledkom, ktoré slúžia na nepretržité fungovanie poisťovne a efektívne využitie jej zdrojov. Náklady je potrebné alokovať na základe typu poistného produktu a rovnako podľa úkonu, na ktorý boli vynaložené. Rozlišujeme náklady súvisiace so vznikom poistenia, s predajom poistky, s jej správou a následne s likvidáciou v prípade nastatia poistnej udalosti. Náklady je potrebné rozdeliť na náklady, ktoré sú kalkulované na jednu poistnú zmluvu a na náklady priamo úmerné výške poistnej sumy prípadne zaplatenému poistnému. Pri kalkulácii poistného môžu byť vyjadrené ako pevná hodnota (fixné náklady) alebo premenlivá zložka (variabilné náklady).

Náklady poisťovne je možné posudzovať z rozličných hľadísk. Podľa toho, v akom časovom období počas trvania poistnej zmluvy vznikli, ich môžeme členiť na začiatkové, priebežné a terminálne náklady. Začiatkové náklady súvisia so vznikom a predajom poistky. Bližšie ich môžeme členiť na náklady na tvorbu a vývoj produktov, na provízie, marketingové a administratívne náklady súvisiace s uzavretím poistnej zmluvy, náklady na predaj a distribúciu. Najväčšiu časť z nich tvoria provízie. Vysoká provízia platená finančným agentom za predaj produktu vyzdvihuje relevantnosť stratégie predaja a marketingu. Môže to mať však negatívny vplyv, pretože to môže lákať agentov k vyšším predajom produktov, ktoré v konečnom dôsledku nespĺňajú predstavu klientov. Výsledkom je potom nárast počtu zrušených poistných zmlúv. Priebežné náklady vznikajú počas poistnej doby a predstavujú náklady na chod poisťovne a správu poistných zmlúv. Môžeme sem zaradiť náklady potrebné na inkaso poistného, poštové náklady, nájomné, náklady na výpočtovú techniku, režijné náklady, dane, platy, audit a iné. Terminálne náklady sú náklady spojené s likvidáciou poistných udalostí a výplatou poistnej sumy.

1.1.4. Inflácia

Pri oceňovaní produktov musí životná poisťovňa brať do úvahy aj infláciu. Inflácia sa vyznačuje úbytkom kúpnej sily peňazí a nárastom celkovej úrovne cien. Najlepším zdrojom

informácií o predpokladanom vývoji inflácie sú údaje Štatistického úradu Slovenskej republiky alebo stránka Národnej banky Slovenska. V poisťovni sa účinok inflácie odzrkadlí na prevádzkových nákladoch, ale aj na nákladoch na jednu poistnú zmluvu. Zvyšovanie nákladov sa ďalej odzrkadlí na výške poistného, preto inflácia je vo všeobecnosti nepriaznivý jav pre životnú poisťovňu. Okrem toho, že rastú náklady poisťovne, vplyvom inflácie tiež klesá reálna hodnota poistného plnenia.

Nežiadúce účinky inflácie riešia poisťovne viacerými spôsobmi. Jednou z možností je dynamizácia produktov. Dynamizácia je založená na princípe, že klient poisťovne má právo bez opätovného preverovania zdravotného stavu zmeniť výšku poistného, a tiež poistnej sumy. Ďalším spôsobom je tzv. automatická dynamizácia alebo inak indexácia, ktorá je tiež vopred dohodnutá v poistnej zmluve. Na vyrovnanie vplyvu inflácie sa na Slovensku najčastejšie využíva revalorizácia. Je založená na tom, že klient sa podieľa na zisku poisťovne.

1.1.5. Zrušenie alebo odkup poistky

Zrušenie a odkup poistky znamená predčasné ukončenie alebo stornovanie poistnej zmluvy z akéhokoľvek dôvodu zo strany poistníka. Rozdiel medzi týmito dvoma výrazmi je taký, že pri zrušení poistky nedochádza k výplate finančnej náhrady a pri odkupe vzniká nárok na výplatu odkupnej hodnoty. Odkupná hodnota býva najčastejšie aktuálna hodnota matematickej rezervy znížená o odkupný poplatok. Aby bolo možné vyplatiť odkupnú hodnotu, musí byť výška aktuálnej hodnoty rezervy vyššia ako začiatkové náklady, preto v prvých rokoch zvyčajne nevzniká nárok na jej výplatu. Ak by klienti vo väčšom množstve stornovali svoje poistné zmluvy a požadovali by vyplatenie odkupnej hodnoty, malo by to negatívny vplyv na životnú poisťovňu, preto je jej cieľom predvídať a predchádzať stornovaniu poistných zmlúv. Analýza stornovanosti v praxi spočíva najmä v pozorovaní správania vlastného poistného kmeňa, pričom vo všeobecnosti platí, že v prvých rokoch trvania poistenia je pravdepodobnosť storna najvyššia a potom v čase klesá. Vtedy je vhodné skúmať storno poistných zmlúv v kratších časových intervaloch.

K odkupom dochádza najčastejšie z dôvodu, že sa klienti dostanú do finančných problémov a nie sú schopní ďalej platiť poistné alebo aj náhle potrebujú nejakú finančnú hotovosť. Tiež je častým dôvodom aj konkurencia na trhu, keď sa klient rozhodne presunúť

poistenie do inej poisťovne, či už z dôvodu lepších podmienok, alebo z dôvodu, že daný produkt nevyhovuje úplne jeho požiadavkám.

1.1.6. Investičný výnos

Modelovanie investičného výnosu tvorí významný predpoklad pri oceňovaní produktov v životnom poistení, ktorý treba brať do úvahy najmä pri produktoch s podielom na zisku poisťovne. Tento predpoklad je ovplyvňovaný hlavne bezrizikovou úrokovou mierou. Pri modelovaní sa využívajú výnosové krivky štátnych dlhopisov, pretože patria k najbezpečnejším formám investície. Štátne dlhopisy sú aktíva s najnižším rizikom. Ďalej sa tiež používa swapová výnosová krivka, ktorej výhodou je, že v porovnaní s krivkou štátnych dlhopisov, pokryje väčšie množstvo kontraktov swapovej povahy. Použitie swapovej krivky sa odporúča aj z dôvodu vyššej výnosnosti, a tiež z dôvodu, že rozdiely medzi štátnymi dlhopismi rôznych krajín môžu byť veľké.

1.2 Úmrtnosť a jej ukazovatele

Úmrtie patrí ku kľúčovým poistným rizikám pre oblasť životného poistenia. Smrť považujeme za náhodný jav, pretože nevieme presne určiť kedy nastane. Práve kvôli tomu je častým predmetom poistenia, ktoré v prípade nastatia poistnej udalosti (úmrtia poisteného) poskytne pozostalým finančné prostriedky. Úmrtnosť sa vyznačuje vlastnosťami, ktoré môžeme opísať podľa Cipru (2006) v nasledujúcich bodoch:

- je reprezentovaná práve dvoma stavmi, ktoré označujeme ako stav živý a stav zomretý, pričom je možné jednoznačne identifikovať stav každého z poistených,
- k prechodu medzi stavmi môže dôjsť iba jedným smerom, a to zo stavu živý do stavu zomretý, pričom tento proces sa nazýva úmrtie,
- okamih, v ktorom nastane úmrtie, je náhodný a nezvratný jav, ktorý je možné opísať pomocou nástrojov teórie pravdepodobnosti.

Na úmrtnosť vplyva mnoho faktorov. Medzi hlavné faktory patrí životná úroveň, pohlavie, vierovyznanie, miera kriminality či chorobnosti, nehodovosť, vojny, samovraždy. Vo všeobecnosti ich môžeme rozdeliť do štyroch východiskových skupín, ktorými sú genetické, ekologické, sociálno-ekonomické faktory, a taktiež pokrok v medicíne.

Do genetických faktorov patrí napríklad aj fakt, že miera úmrtnosti u mužov je vyššia oproti miere úmrtnosti žien. Vzhľadom na túto skutočnosť, a tiež vzhľadom na to, že ženy sa dožívajú vyššieho veku, je logické analyzovať úmrtnosť oddelene pre jednotlivé pohlavia. Avšak na základe nariadenia Európskej únie 2004/113/ES, ktoré v skratke nazývame gender smernica, majú štáty Európskej únie zakázané brať do úvahy faktor pohlavia pri analýze úmrtnosti, a tiež aj v ostatných oblastiach. Bližšie sa gender smernici budeme ďalej venovať v kapitole 1.3, ktorá sa zaoberá legislatívnou úpravou poisťovníctva. Príkladmi ekologických faktorov sú napríklad klimatické podmienky a životné prostredie. Sociálno-ekonomické faktory môžeme ďalej rozčleniť na individuálne (životná úroveň, vzdelanie, stravovacie návyky, fyzická aktivita) a také, ktoré sú spôsobené vplyvom prostredia (ekonomická situácia krajiny, sociálne zabezpečenie). Za samostatnú skupinu považujeme pokrok v medicíne, ktorá sa za posledné desaťročie neuveriteľne posunula k lepšiemu. K zlepšeniu došlo vo viacerých oblastiach ako je dostupnosť liečiv, ale aj v objavení nových liečebných postupov pre rôzne civilizačné choroby ako sú cukrovka, rakovina a iné.

Ukazovatele úmrtnosti môžeme rozdeliť na absolútne a relatívne. Absolútne ukazovatele majú tú nevýhodu, že nie sú vhodné na porovnávanie v čase ani v priestore, a preto sa častejšie v praxi používajú relatívne ukazovatele úmrtnosti. Poznanie miery úmrtnosti poisťencov je potrebné pri všetkých druhoch produktov životného poistenia. Pravdepodobnosť nastatia poisťnej udalosti ovplyvňujú okrem miery úmrtnosti aj iné faktory, akými sú napríklad zamestnanie, záľuby, zdravotný stav, návyky a iné. Všetky tieto faktory sa berú do úvahy pri posúdení a oceňovaní poisťného rizika. Individuálne miery úmrtnosti poistených nemusia byť totožné s mierami úmrtnosti celej populácie, ktoré sú publikované centrálnymi štatistickými úradmi.

Informáciu o procese vymierania populácie zachytáva všeobecná miera úmrtnosti, ktorá je zároveň aj najsledovanejšou charakteristikou úmrtnosti. Jej hodnota nám podáva informáciu o počte zomretých osôb pripadajúcich na 1 000 obyvateľov stredného stavu vo vybranom kalendárnom roku. Všeobecná miera úmrtnosti nie je vhodná na porovnávanie, pretože neberie do úvahy vekovú štruktúru obyvateľstva. Tento nedostatok odstraňujú špecifické miery úmrtnosti, ktoré nás informujú o vývoji úmrtnosti podľa pohlavia a veku. Výsledná hodnota tejto charakteristiky definuje počet zomretých žien alebo mužov v dokončenom veku na 1 000 žien alebo mužov v tom istom veku.

Pravdepodobnosť úmrtia sa s rastúcim vekom zvyšuje, výnimkou je však úmrtnosť v prvom roku života, ktorá je vyššia (dojčenská úmrtnosť). Keďže sa jedná o akúsi anomáliu, venuje sa jej osobitá pozornosť. Ukazovateľ dojčenskej úmrtnosti sa často využíva na medzinárodné porovnávanie sociálnej vyspelosti jednotlivých populácií a zaznamenáva počet detí, ktoré zomreli v priebehu prvého roku života (365 dní) na 1 000 živo narodených detí. Ďalším ukazovateľom úmrtnosti v prvom veku života je kvocient dojčenskej úmrtnosti, ktorý vyjadruje pravdepodobnosť úmrtia novorodenca 365 dní po narodení. Počíta sa ako podiel zomretých novorodencov k celkovému počtu narodených detí. Zlepšovanie zdravotníckej starostlivosti a celkovo pokrok v medicíne má za následok, že hranica úmrtnosti narodených detí sa posúva stále viac k začiatkom života, čo viedlo k vzniku ďalších ukazovateľov zameraných na úmrtnosť v skorších obdobiach života. Zaradujeme sem ukazovateľ popôrodnej úmrtnosti, ktorý zisťuje koľko detí zomrelo počas prvých troch dní života na 1 000 živo narodených detí. Včasná novorodenecká úmrtnosť sa sústreďuje na úmrtnosť detí počas prvých siedmych dní života a je zároveň súčasťou ukazovateľa perinatálnej úmrtnosti, ktorá okrem toho berie do úvahy aj počet mŕtvo narodených detí na 1 000 živo narodených. Novorodenecká úmrtnosť sa zaoberá úmrtnosťou počas prvých 28 dní života dieťaťa, rovnako na 1 000 živo narodených detí. Ďalej si spomenieme ešte ukazovateľ ponovorodeneckej úmrtnosti, ktorá sleduje počet zomretých detí medzi 29. dňom života a jedným rokom na 1 000 živonarodených detí.

Druhou anomáliou pri analýze úmrtnosti populácie je tzv. “havarijný hrb”. Prejavuje sa nárastom úmrtnosti u mladých mužov vo veku 15 – 35 rokov. Je to z dôvodu zvýšeného rizikového správania a následného úmrtia najmä v dôsledku vonkajších príčin ako sú dopravné nehody, úrazy, samovraždy a podobne. V niektorých krajinách sa tento hrb objavuje aj pri mladých ženách. Príčina zvýšenej úmrtnosti u mladých žien však pramení skôr z rizika materskej úmrtnosti. Materská úmrtnosť je úmrtnosť spojená s tehotenstvom, pôrodom a šestonedelím. S určitosťou však vieme povedať, že vo všetkých vekových kategóriách miera úmrtnosti mužov presahuje mieru úmrtnosti žien. Takýto jav sa nazýva tiež ako mužská nadúmrtnosť alebo aj maskulinita zomretých.

Okrem spomínaných je významným ukazovateľom aj stredná dĺžka života novorodenca alebo stredná dĺžka života osoby v presnom veku x . Hovorí o priemernom počte rokov, ktorých sa ešte novorodenec (alebo osoba vo veku x) dožije. Vychádza z predpokladu, že úmrtnostné pomery populácie zostanú nezmenené, a teda vyjadruje súčasnú úmrtnostnú

situáciu. Iné pomenovanie pre strednú dĺžku života je pojem očakávaná dĺžka života alebo nádej dožitia. Častejšie ako stredná dĺžka života sa v praxi používa medián dĺžky života (pravdepodobná dĺžka života novorodenca). Interpretujeme ho ako vek, ktorého sa dožije polovica populácie. Tiež ale vychádza z predpokladu zachovania úmrtnostných pomerov obyvateľstva. Výhodou mediánu dĺžky života je, že nie je ovplyvnený extrémnymi hodnotami. Modus dĺžky života (normálna dĺžka života) vyjadruje vek, v ktorom ľudia zomierajú najčastejšie. V roku 2020 na Slovensku ľudia umierali najčastejšie vo veku 84 rokov, zatiaľ čo v roku 1993 to bolo 80 rokov, čo predstavuje nárast o 4 roky.

1.2.1 Analýza úmrtnosti na Slovensku

Štatistiky úmrtnosti v predchádzajúcich dvoch rokoch výrazne ovplyvnila pandémia koronavírusu a úmrtia spôsobené ochorením Covid-19. Celkovo na Slovensku v roku 2020 zomrelo 59 089 ľudí, z čoho bolo 30 428 mužov a 28 661 žien. Údaje o príčinách úmrtnosti populácie na Slovensku v roku 2020 sú uvedené v tabuľke č. 1. Dlhodobou najčastejšou príčinou úmrtia u oboch pohlaví boli choroby obehovej sústavy, ktoré tvoria až 46 % všetkých príčin úmrtí. Druhou najčastejšou príčinou úmrtí na Slovensku sú nádory. Majú takmer 24 % podiel na úmrtiach na Slovensku. Infekcia Covid-19 mala za následok 6,8 % úmrtí, pričom najviac sa prejavila vo vekovej skupine seniorov (vo veku nad 65 rokov). V tejto vekovej kategórii na tisíc obyvateľov zomrelo až 78 osôb, kým napríklad rok predtým (2019) to bolo 46 osôb. Miera dojcenskej úmrtnosti v roku 2020 dosiahla úroveň 5,08 ‰, čo v absolútnom vyjadrení predstavuje úmrtie 288 detí vo veku do 1 roka.

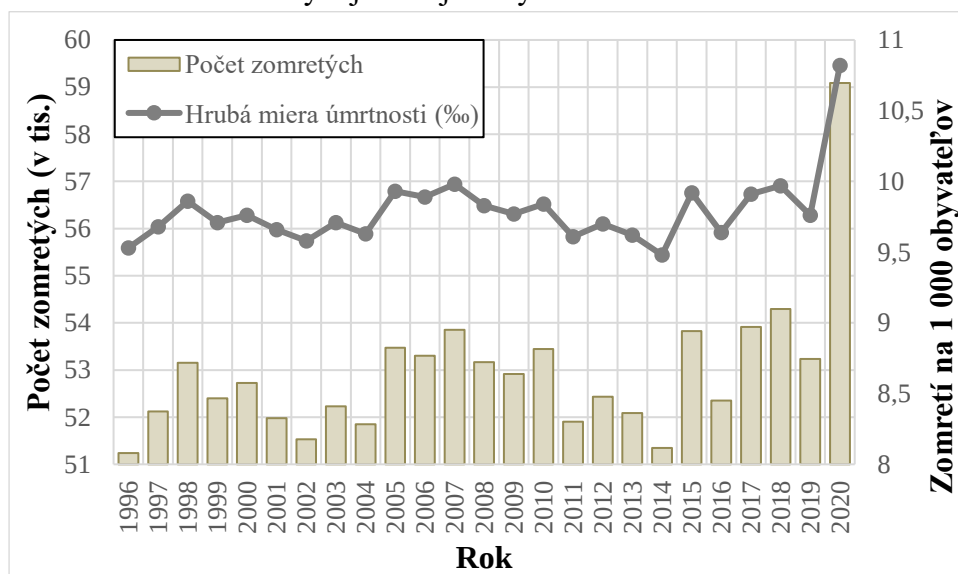
Tabuľka č. 1 - Úmrtnosť podľa príčiny na Slovensku v roku 2020

Príčiny úmrtí v SR	Podiel za rok 2020
<i>choroby obehovej sústavy</i>	46,0 %
<i>nádory</i>	23,7 %
<i>infekcia COVID-19</i>	6,8 %
<i>choroby dýchacej sústavy</i>	6,4 %
<i>choroby tráviacej sústavy</i>	4,9 %
<i>vonkajšie príčiny úmrtnosti</i>	4,1 %
<i>ostatné príčiny úmrtí</i>	8,1 %

Zdroj: vlastné spracovanie na základe údajov zo Štatistického úradu SR

Počas pozorovaného obdobia medzi rokmi 1993 až 2020 sa počet úmrtí pohyboval v intervale 51 000 – 60 000, pričom najnižší počet úmrtí bol zaznamenaný v roku 1996, kedy zomrelo 51 236 ľudí. Naopak, najvyšší počet úmrtí pripadlo na spomínaný pandemický rok 2020, ktorý prispel aj k vyššej celkovej hrubej miere úmrtnosti na Slovensku. Hrubá miera úmrtnosti sa udržiavala pod úrovňou 10 úmrtí na 1 000 obyvateľov až do roku 2020, kedy vzrástla z úrovne 9,76 zomretých na tisíc obyvateľov v roku 2019 na hodnotu 10,82 zomretých na tisíc obyvateľov. Inak povedané, v priemere za celý rok 2020 zomrelo takmer 11 osôb na každých tisíc obyvateľov krajiny. Ide o výrazný nárast hodnoty, ktorá sa zvyčajne medziročne mení najviac o niekoľko desiatín, pričom aktuálne to bolo o viac ako jednu osobu. Hrubá miera úmrtnosti v roku 2020 je podľa štatistického úradu najvyššia za posledných 70 rokov. Najnižšie hodnoty hrubej miery úmrtnosti na úrovni 9,5 ‰ boli zaznamenané v rokoch 1996 a 2014. Medziročný pokles bol zaregistrovaný v prípade dojčenskej úmrtnosti, reálne poklesla o 4 úmrtia oproti roku 2019, čo malo za následok zníženie miery dojčenskej úmrtnosti o 0,04 bodu (z 5,12 ‰ na 5,08 ‰).

Graf č. 1 - Vývoj hrubej miery úmrtnosti na Slovensku

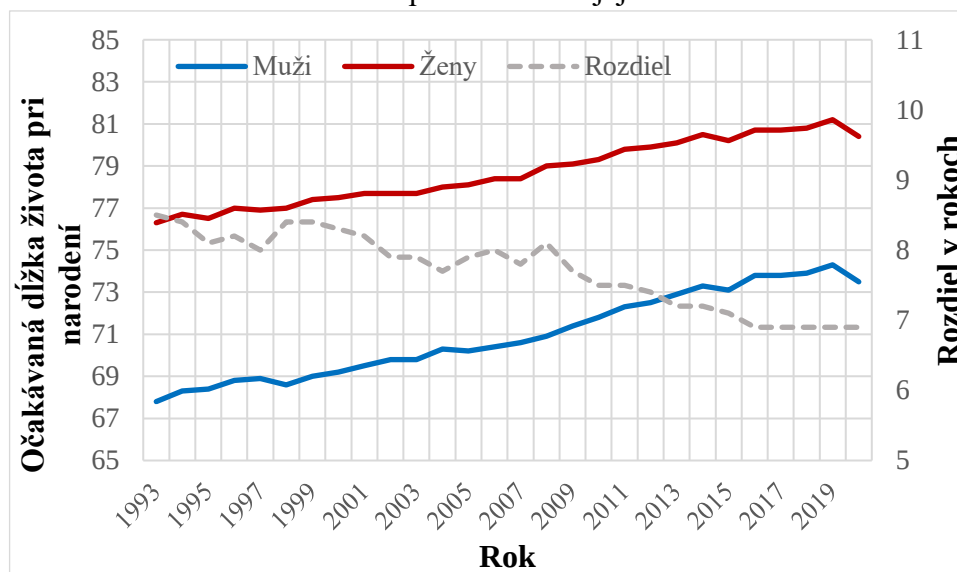


Zdroj: vlastné spracovanie na základe údajov zo Štatistického úradu SR

V sledovanom období pretrvávala nadúmrtnosť mužskej populácie viditeľná do veku 78 rokov vrátane. Nadúmrtnosť mužskej populácie je podľa Filadelfiovej (2005) spojená s vyššou chorobnosťou a horšou životosprávou mužov. Navyše muži menej dbajú na prevenciu, ľahšie podliehajú závislostiam a majú vyššie riziko samovrážd. Rozdiely

medzi úmrtnosťou mužskej a ženskej populácie sa však neustále znižujú. Zatiaľ čo v roku 1993 pripadalo na 1 000 zomretých žien až 1 200 zomretých mužov, v roku 2019 to bolo len 1 040 zomretých mužov na 1 000 zomretých žien. Tento fakt nám potvrdzuje aj syntetický ukazovateľ strednej dĺžky života pri narodení, ktorého hodnota sa z roka na rok zvyšuje. V roku 2019 bola stredná dĺžka života žien pri narodení 81,2 rokov a stredná dĺžka života mužov pri narodení 74,3 rokov. V porovnaní s predchádzajúcim rokom sa stredná dĺžka života u oboch pohlaví zvýšila o 0,4. Počas celého sledovaného obdobia sa však stredná dĺžka života žien pri narodení zvýšila o 4,9 rokov, stredná dĺžka života mužov sa zvýšila o 6,5 rokov. Na základe predbežných výsledkov sa odhaduje, že v roku 2020 došlo k poklesu očakávanej dĺžky života u oboch pohlaví o 0,8 roka oproti roku 2019. Dôležitým trendom je pokles rozdielu v očakávanej dĺžke života pri narodení medzi pohlaviami znázornený na grafe č. 2. V roku 1993 bol rozdiel medzi pohlaviami 8,5 rokov, do roku 2020 klesol na 6,9 rokov. Detailný rozbor vplyvu pandémie koronavírusu na očakávanú dĺžku života pri narodení, ale aj na ďalšie úmrtnostné charakteristiky môžeme nájsť v publikácii od Šprochu (2021).

Graf č. 2 - Očakávaná dĺžka života pri narodení a jej rozdiel medzi mužmi a ženami

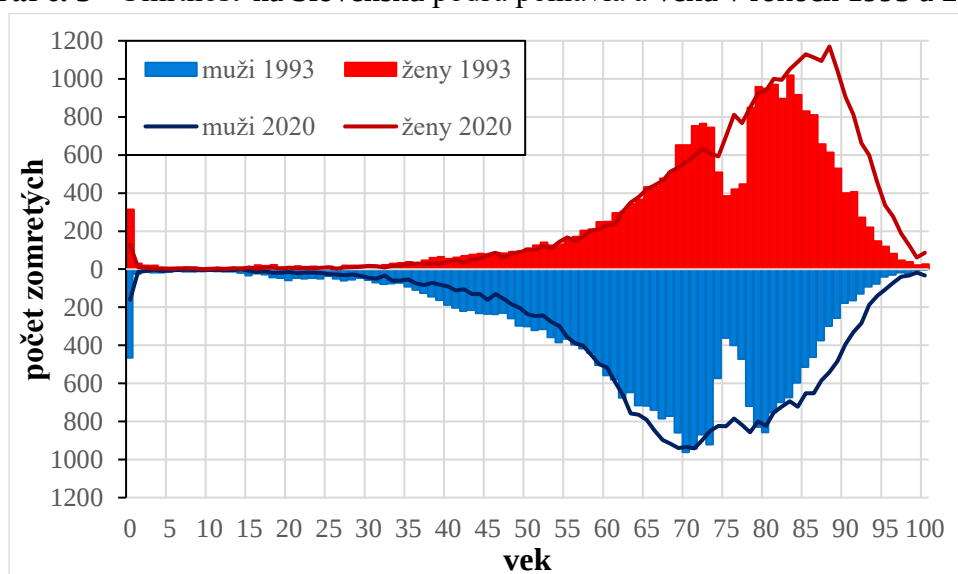


Zdroj: vlastné spracovanie na základe údajov z Eurostatu

Na grafe č. 3 je znázornená zmena trendu úmrtnosti podľa pohlavia a veku medzi rokmi 1993 a 2020. Z grafu je zrejmé, že napriek pandémie koronavírusu, úmrtnosť vo vyšších vekoch sa neustále posúva. Tento jav je následkom či už pokroku v medicíne, sociológii

a farmácii, ale aj v iných vedách, ktoré priamo alebo nepriamo súvisia s úmrtnosťou. Kvalita života na Slovensku sa každým rokom zvyšuje, čo sa samozrejme odzrkadľuje aj na vývoji úmrtnosti a takisto to má vplyv aj na znižovanie rozdielov v úmrtnosti mužov a žien. Medzi faktory, ktoré súvisia s poklesom rozdielov v úmrtnosti medzi pohlaviami patrí aj sociálna vyspelosť krajiny a stratégia Slovenska na dosiahnutie rodovej rovnosti v rôznych oblastiach života. V poslednom desaťročí sa stále viac kladie dôraz na zdravý životný štýl, na psychické zdravie a na spochybňovanie rodových stereotypov, odstránenie rodových rozdielov na trhu práce a v oblasti poskytovania starostlivosti.

Graf č. 3 - Úmrtnosť na Slovensku podľa pohlavia a veku v rokoch 1993 a 2020



Zdroj údajov: vlastné spracovanie na základe údajov zo Štatistického úradu SR

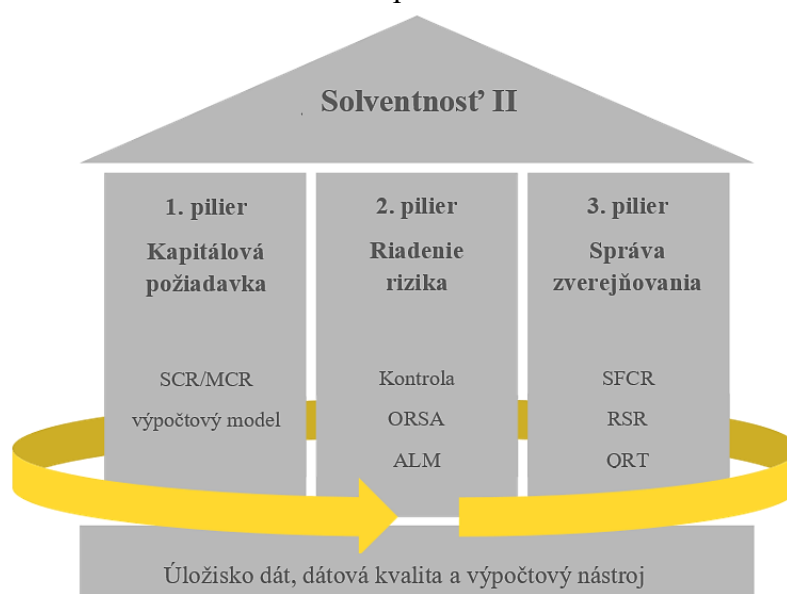
1.3 Legislatívna úprava poisťovníctva Európskej Únie

Vstupom do Európskej únie sa členské štáty zaviazali akceptovať právo únie, s čím súvisí aj rešpektovanie jej smerníc a nariadení. Viaceré z nich sú spojené aj s oblasťou poisťovníctva a boli vytvorené za účelom regulácie a dohľadu nad týmto odvetvím. Smernice týkajúce sa oblasti poisťovníctva sú verejne prístupné na webovej stránke Ministerstva financií Slovenskej republiky alebo na stránke EUR-Lex, ktorá obsahuje kompletnú legislatívu Európskej únie vo všetkých úradných jazykoch členských štátov.

Za najpodstatnejšiu a zároveň prevratnú v tejto oblasti považujeme smernicu Európskeho parlamentu a Rady 2009/138/ES z 25. novembra 2009 o začatí a vykonávaní

poistenia a zaistenia (tzv. smernica Solventnosť II). Išlo o medzinárodný projekt, ktorý bol vytvorený za účelom rekodifikácie trinástich už existujúcich smerníc, ktoré dovtedy upravovali činnosť poisťovní a zaistovní. Mechanizmus Solventnosť II sa dostal do platnosti 1. januára 2016, hoci samotný projekt na jeho vytvorenie odštartoval už v roku 2000. Poisťovne nezaznamenali výraznejšie problémy s implementáciou tejto smernice. Medzi ciele smernice Solventnosť II patrilo nastolenie nového režimu regulácie poisťného sektora, a tým zvýšenie ochrany poisťencov a ostatných príjemcov plnenia vyplývajúceho z poisťných zmlúv. Úmyslom smernice bol takisto rozvoj integrácie poisťného trhu únie za pomoci harmonizácie poisťného sektora, a tiež nárast medzinárodnej konkurencieschopnosti poisťovní. Solventnosť II je postavená na koncepte troch pilierov, ktorých stručný prehľad je znázornený na obrázku č. 1.

Obrázok č. 1 - Princíp režimu Solventnosť II



Zdroj údajov: vlastné spracovanie

Zásadnými prínosmi smernice Solvetnosť II bolo zabezpečenie medziodvetvovej jednotnosti, a tiež jednotnosti v regulácii poisťného sektora v rámci Európskej únie, zvýšenie transparentnosti poisťného sektora a orgánov dohľadu, zavedenie prístupu zameraného na riziko a jeho riadenie, nastolenie súladu s realistickým oceňovaním podľa aktuálneho trhového oceňovania aktív a pasív poisťovne a v neposlednom rade aj kladenie dôrazu na správne riadenie poisťovne s osobitným zameraním na systém riadenia rizík, vnútornej kontroly, aktuársku funkciu a vnútorný audit.

Ďalšou veľmi dôležitou a zlomovou európskou smernicou, ktorá formovala sektor poisťovníctva, je smernica Rady 2004/113/ES z 13. decembra 2004 o vykonávaní zásady rovnakého zaobchádzania medzi mužmi a ženami v prístupe k tovaru a službám a k ich poskytovaní alebo jej kratšie pomenovanie tzv. gender smernica. Ako už vyplýva aj z názvu, predmetom smernice je antidiskriminácia na základe pohlavia, teda identické zachádzanie s osobami bez ohľadu na pohlavie. Direktíva vymedzila pravidlá pre oblasť poisťovníctva, podľa ktorých je nevyhnutné jednotné zaobchádzanie s osobami bez ohľadu na pohlavie, na základe čoho je nutné faktor pohlavia vylúčiť z oceňovania produktov, hoci je pri oceňovaní rozhodujúci. Pretože určité kategórie poistných rizík sa môžu pre pohlavia líšiť, bola spočiatku udelená poisťovníctvu výnimka, na základe ktorej mohli poisťovne naďalej využívať faktor pohlavia pri svojich výpočtoch, za predpokladu podloženia odlišností medzi pohlaviami na základne aktuársko-štatistických údajov. Okrem toho komisia požadovala, aby štáty Európskej únie, ktoré využijú túto výnimku, zverejňovali informácie týkajúce sa použitia faktora pohlavia pri výpočtoch. Výnimku sa rozhodli prijať všetky členské štáty, a to najmä pre dôchodkové a rizikové životné poistenie. Vzhľadom na štatistické zistenie, podľa ktorého sú muži rizikovejší vodiči, sa niektoré štáty priklonili k použitiu výnimky aj pri povinnom zmluvnom alebo havarijnom poistení. V marci 2011 nastal v tejto oblasti zlomový moment, kedy Európsky súdny dvor rozhodol o odstránení výnimky z nariadenia, čím nakázal absolútny zákaz použitia faktora pohlavia pri všetkých aktuárskych výpočtoch, ktoré sa týkajú výpočtu poisteného a poistných plnení. Rozsudok Súdneho dvora EÚ č. C-236/09 európskeho súdu o zrušení výnimky v oblasti poisťovníctva pre členské štáty Európskej Únie sa stal právoplatným ku dňu 21. decembra 2012.

Verdikt súdu sa stretol s negatívnymi reakciami zo strany poisťovní a aj samotných ich klientov, ktorí nesúhlasili, aby sa prestalo pohlavie využívať pri oceňovaní rizík a poistných produktov. Samotná Slovenská asociácia poisťovní (SLASPO), ktorá združuje slovenské poisťovne, uverejnila na svojej webovej stránke svoj postoj k téme antidiskriminácie. SLASPO chápe a podporuje zákaz diskriminácie ako takej, avšak nepovažuje použitie faktora pohlavia na účely poistenia za diskrimináciu. Ich argumentom bol fakt, že úlohou poisťovne je práca s rizikami, ktoré je potrebné čo najpresnejšie oceniť podľa všetkých možných faktorov, ešte skôr než poisťovňa toto riziko preberie. Kľúčovými faktormi oceňovania v životnom poistení je podľa nej vek, zdravotný stav, ale osobitne aj pohlavie. Predpokladaným vplyvom zrušenia výnimky na poisťovne sa venovala aj spoločnosť

OXERA (The Oxford Economic Research Associates) – Oxfordská asociácia ekonomického výskumu. Publikovala štúdiu o očakávanom účinku vylúčenia faktora pohlavia na poistencov. Štúdia spoločnosti OXERA obsahovala argumenty ako napríklad to, že riziko úmrtnosti mužov a žien je charakterizované znateľnými rozdielmi, ktoré vplyvajú na produkty životnej poisťovne alebo fakt, že rozdielne je aj riziko zapríčinenia dopravnej nehody. Výsledkom štúdie boli percentuálne zmeny vo výške poistného v prípade, že sa schváli absolútna rovnoprávnosť pohlavia v oblasti poisťovníctva. Pre ženy vo veku okolo 40 rokov by došlo k nárastu poistného pri produktoch životného poistenia v priemere o viac ako 30 %. V prípade poistenia na dožitie by sa u mužov vo veku okolo 65 rokov znížilo poistné plnenie o približne 5 %. U mladých šoférov by sa zvýšilo poistné v priemere o viac ako 11 % pri poistení motorového vozidla. Okrem toho sa OXERA obávala, že dôjde k tzv. fenoménu nepriaznivého výberu, ktorý by zapríčiniť, že poistenci s nižším rizikom by platili poistné na úkor rizikovejších poistených, čo by spôsobilo pokles záujmu u menej rizikových poistených, ktorí by následne zrušili existujúce poistné zmluvy, pretože poistenie by sa pre nich stalo pridrahé. To by malo za následok zvýšenie podielu viac rizikovejších poistencov, čo by znova viedlo k nárastu poistného a postupnému odchodu ďalších poistených (obrázok č. 2).



Zdroj: OXERA, The impact of a ban on the use of gender in insurance

Najväčšou obavou spotrebiteľov po nastolení rovnoprávnosti medzi mužmi a ženami bola zmena ceny poistných produktov. Zo strany poisťovní bolo obavou najmä fungovanie nového systému ako celku. Dnes už vieme povedať, ktoré obavy sa naplnili, a ktoré naopak nie. Poisťovniam táto zmena spôsobila dodatočné náklady súvisiace s oceňovaním a tvorbou nových poistných produktov, či preceňovaním už existujúcich. Výzvou pre poisťovne bolo

pri oceňovaní rizika zamerať pozornosť na iné faktory ako je pohlavie. Aby predišli obavám spotrebiteľov o raste cien, poisťovne zlacneli ceny pripoistení pre vybrané produkty, ktorých sa týkali zmeny. Okrem modernizácie starých produktov získali poisťovne príležitosť zamerať sa na produkty, ktoré sú typické pre jednotlivé pohlavia. Za významnú hrozbu všetkých obmedzení v legislatíve EÚ sa považuje oslabenie ekonomickej efektívnosti poisťovníctva.

Z pohľadu účtovníctva sú poisťovne na Slovensku povinné zostavovať účtovnú závierku v súlade s medzinárodnými štandardmi finančného výkazníctva IFRS (International Financial Reporting Standards), ktoré sú vydávané medzinárodnou radou pre účtovné štandardy IASB (International Accounting Standard Board). Tieto štandardy majú za úlohu zvýšiť transparentnosť účtovných výkazov poisťovní. 18. mája 2017 Rada pre medzinárodné účtovné štandardy vydala nový štandard IFRS 17 – Poistné zmluvy. Ten nahrádza terajší štandard IFRS 4, ktorý v súčasnosti pripúšťa širokú škálu možných postupov.

Štandard IFRS 17 upravuje princípy ohodnocovania, vykazovania a zverejňovania poistných zmlúv a zásadným spôsobom mení účtovanie pri všetkých účtovných jednotkách, ktoré vystavujú poistné zmluvy či investičné zmluvy s prvkami podielu na prebytku. Tento štandard bude účinný pre ročné účtovné obdobia začínajúce sa 1. januára 2023 alebo neskôr, pričom ho môžeme považovať za prvý takmer globálny štandard pre poistné zmluvy v Európskej únii. Hlavným cieľom je analýza ziskovosti poistných zmlúv a definovanie ich smerovania. Ďalším dôležitým cieľom, z dôvodu ktorého bol štandard vytvorený, je to, aby poisťovne rozložili zisk z poistných zmlúv do viacerých období, s ktorými vecne a časovo súvisí. Totižto podľa IFRS 4, ktorý je aktuálne v platnosti, sa zisk nevykazuje v čase, kedy skutočne vznikol a IFRS 17 sa pokúša tento nedostatok odstrániť.

Nový štandard definuje tri metódy ohodnocovania poistných zmlúv na základe typu produktu. Základnou metódou ohodnocovania je metóda GMM (General model measurement), ktorá by mala byť použiteľná pre väčšinu dlhodobých poistných zmlúv či už životného alebo neživotného poistenia. Táto metóda je vhodná pre tradičné životné postenie bez podielu na zisku, prípadne s negarantovaným podielom na zisku, kam zaradujeme napríklad doživotné poistenie na úmrtie alebo zmiešané poistenie. Druhou metódou je metóda VFA (Variable fee approach), ktorú majú poisťovne použiť pre poistenia s priamym podielom na zisku a pre investičné životné poistenie. Táto metóda je modifikáciou základnej

GMM metódy. Poslednou dobrovoľnou a zároveň najjednoduchšou metódou, ktorú poisťovne nemusia využiť, je metóda PAA (Premium allocation approach). Vhodná je hlavne pre krátkodobé poistné zmluvy, ktoré sú spojené najmä s neživotným poistením.

Špecialita tohto štandardu spočíva v detailnom delení poistných zmlúv. Zatiaľ čo pri štandarde IFRS 4 sa zmluvy delili na poistné a investičné, pri IFRS 17 sa poistné zmluvy agregujú na základe troch úrovní, a to podľa typu rizika, podľa obdobia vzniku (zmluvy uzatvorené v priebehu jedného roka) a podľa ziskovosti (ziskové, stratové a iné).

IFRS 17 kombinuje súčasne ocenenie budúcich peňažných tokov s vykázaním zisku za obdobie, v ktorom sú na základe zmluvy služby poskytované. Výsledky poisťovacích služieb (vrátane prezentácie výnosov z poistenia) sú podľa tohto štandardu prezentované oddelene od výnosov alebo výdavkov z poistenia. Štandard IFRS 17 ovplyvní poisťovne vo viacerých oblastiach, od finančníctva, poistnej matematiky, účtovníctva a informačných systémov až po dane a dividendy. Pre poisťovňu to bude znamenať zvýšenie nákladov na implementáciu štandardu v podobe nových pracovníkov a ich vzdelávania, ale aj na celkový chod poisťovne. IFRS 17 so sebou prinesie veľké množstvo údajov, ktoré bude treba spracovávať, uchovávať a analyzovať, čo bude spojené s vylepšením nákladnej IT infraštruktúry.

2 Cieľ práce

Hlavných cieľom dizertačnej práce je teoreticky aj prakticky preskúmať problematiku modelovania úmrtnosti pomocou parametrických modelov z aktuárskeho hľadiska. Následná aplikácia nadobudnutých poznatkov spolu s analýzou vplyvov, ktoré majú dopad na modelovanie úmrtnosti, naplňajú hlavný zámer záverečnej práce.

Úmrtnosť obyvateľstva sa neustále mení a vyvíja. Aj v dôsledku koronavírusu sme si ako ľudia uvedomili, aké následky môže mať zvýšená miera úmrtnosti na finančný stav domácností, ale aj iných subjektov. Tieto aspekty sa týkajú aj životných poisťovní, u ktorých smrť patrí k najviac poisťovaným rizikám. Životná poisťovňa sa snaží čo najpresnejšie poznať úmrtnostné správanie populácie a je pre ňu dôležité vedieť odhadnúť aj jeho vývoj do budúcnosti, aby bola schopná predchádzať finančným stratám, ktoré sú s rizikom úmrtnosti spojené. Jednou z možností sú práve parametrické modely úmrtnosti, ktoré sú predmetom tejto dizertačnej práce.

V teoretickej časti práce je najskôr naším cieľom predstaviť úmrtnosť ako súčasť aktuárskej bázy, jej ukazovatele a aspekty, ktoré na ňu vplývajú, ako aj analyzovať jej vývoj od vzniku Slovenskej republiky v roku 1993 až po súčasnosť. Pre splnenie hlavného cieľa je potrebné zosumarizovať deterministickú a stochastickú metodológiu modelovania úmrtnosti populácie, aby sme boli schopní naštudované metódy a postupy využiť v analýze úmrtnosti populácie na Slovensku podľa veku a pohlavia a poukázať na ďalší vývoj úmrtnosti v čase. Súčasťou tejto metodológie sú samotné deterministické a stochastické modely. Na odhad parametrov deterministických modelov sa používajú metódy nelineárnej regresie a na odhad stochastických modelov použijeme metódu najmenších štvorcov za pomoci singulárneho rozkladu matice.

V praktickej časti práce je cieľom využiť teoretické poznatky na modelovanie úmrtnostného správania populácie Slovenskej republiky pomocou vybraných deterministických a stochastických modelov v rokoch od 1993 až po súčasnosť. Predmetom modelovania sú vekovo špecifické miery úmrtnosti mužov a žien ako aj spoločné vekovo špecifické miery úmrtnosti v rozpätí od 0 do 100 rokov vypočítané pomocou údajov o úmrtnosti získaných zo Štatistického úradu Slovenskej republiky. Jednotlivé modely je potrebné medzi sebou porovnať, aby sme zistili, ktorý z modelov vie najlepšie zachytiť úmrtnostné správanie populácie Slovenska. Posledným, no veľmi podstatným zámerom tejto

práce, je poukázať na vplyv výberu modelu, efekt zavedenia gender smernice a tiež na dopad koronavírusu na poistné sadzby základných poistných produktov, ktorými sú dočasné poistenie na úmrtie a dočasný predľhotný dôchodok.

Na dosiahnutie všetkých cieľov využijeme všestranný štatistický softvér SAS Enterprise Guide, pričom v záverečnej práci sa zameriame aj na predstavenie tohto softvéru s ohľadom na jeho použitie pri modelovaní úmrtnosti. Ako pomocné nástroje sme použili aj softvérové rozhranie R a Microsoft Excel.

3 Metodika práce a metódy skúmania

V tejto kapitole nájdeme prehľad deterministických a stochastických modelov úmrtnosti, ktoré sa používajú na modelovanie úmrtnosti nielen ľudskej populácie. Naším cieľom je odhadnúť parametre týchto modelov pomocou štatistického softvéru SAS Enterprise Guide. Uvedieme metódy odhadu parametrov deterministických modelov (metódy nelineárnej regresie) ako aj stochastických modelov, ktoré sú v tomto analytickom softvéri dostupné. Kapitola obsahuje aj základné miery kvality modelov a presnosti ich prognóz (priemerný súčet štvorcov rezíduí MSE, štandardná odchýlka rezíduí RMSE, Akaikovo informačné kritérium AIC, Bayesovské informačné kritérium BIC), ktoré sme v praktickej časti použili na vzájomné porovnanie modelov úmrtnosti a tiež na porovnanie ich prognostických schopností.

3.1 Modely úmrtnosti

Modelov úmrtnosti je mnoho, no vo všeobecnosti ich vieme roztriediť na dve skupiny, a to na deterministické a stochastické modely úmrtnosti. Najčastejšie sú modely vyjadrené ako funkcie úmrtnosti, ktorými sú miera úmrtnosti q_x , intenzita úmrtnosti μ_x a centrálna miera úmrtnosti m_x . Miera úmrtnosti q_x alebo presnejšie ročná miera úmrtnosti vyjadruje pravdepodobnosť, že sa x ročná osoba nedožije veku $(x + 1)$, teda ďalšieho roku. Intenzita úmrtnosti μ_x je kľúčovým pojmom pre stochasticky chápanú analýzu prežitia a je analogickým vyjadrením miery úmrtnosti pomocou spojitej náhodnej veličiny. Môžeme povedať, že intenzita úmrtnosti predstavuje pravdepodobnosť ukončenia života v časovom intervale $\langle t, t + dt \rangle$, pričom $dt \rightarrow 0$, za podmienky, že doba života osoby bola až do času t . Inými slovami pravdepodobnosť úmrtnosti v závislosti od doby trvania života (na dĺžke života). Centrálna miera úmrtnosti m_x je v podstate z demografického hľadiska všeobecná miera úmrtnosti, ktorú sme zadefinovali v časti 1.2.

V dizertačnej práci vychádzame z údajov o úmrtnosti obyvateľstva Slovenskej republiky v rokoch od jej vzniku (1993) až po rok 2020 získaných zo stránky Štatistického úradu Slovenskej republiky. Štatistický úrad SR je ústredným orgánom štátnej správy Slovenskej republiky pre oblasť štátnej štatistiky a v jeho databáze DATAcube zverejňuje údaje pre ukazovatele hospodárskeho a sociálno-ekonomického vývoja dostupné

v mesačných, štvrtročných alebo ročných časových radoch. Dáta sú štatisticky spracované v územných štruktúrach za Slovenskú republiku, oblasti, kraje a okresy. Z tejto databázy sme získali údaje o počte úmrtí v stanovenom veku na Slovensku vo vybraných rokoch v závislosti od pohlavia a stredný stav obyvateľstva Slovenska takisto v závislosti od veku a pohlavia vo vybraných rokoch. Na základe týchto údajov sme vypočítali vekovo špecifickú centrálnu mieru úmrtnosti pre mužov a ženy, ale aj spoločnú bez ohľadu na pohlavie, pomocou vzťahu

$$m_{x,t} = \frac{D_{x,t}}{L_{x,t}} \quad (3.1)$$

kde $D_{x,t}$ je počet zomretých vo veku x v čase t a $L_{x,t}$ predstavuje stredný stav obyvateľstva vo veku x v čase t . Ako sme spomenuli, nie všetky modely využívajú pri odhadoch priamo centrálnu mieru úmrtnosti, a teda potrebujeme odvodiť ešte mieru úmrtnosti q_x a intenzitu úmrtnosti μ_x . Existujú dva spôsoby aproximácie centrálnej miery úmrtnosti $m_{x,t}$ pomocou pravdepodobnosti úmrtia osoby vo veku x v čase t , ktorú označujeme $q_{x,t}$. Prvý prístup

$$q_{x,t} = 1 - e^{-m_{x,t}} \quad (3.2)$$

je založený na predpoklade o konštantnej úmrtnosti medzi dvoma celočíselnými vekmi. Druhý prístup, $q_{x,t} = \frac{m_{x,t}}{1+0,5m_{x,t}}$, funguje za predpokladu rovnomerného rozdelenia úmrtí medzi dvoma vekmi. V dizertačnej práci je použitá prvá aproximácia pre transformáciu dát medzi vekovo špecifickou centrálnou mierou úmrtnosti $m_{x,t}$ a jednoročnou mierou úmrtnosti $q_{x,t}$. Pri využití toho istého predpokladu o konštantnej miere úmrtnosti, pre intenzitu úmrtnosti platí, že $\mu_{x,t} = m_{x,t}$.

3.1.1 *Deterministické modely – zákony úmrtnosti*

Zákony úmrtnosti majú tvar matematických spojitých funkcií so závislosťou na veku. Zvyčajne sa počet ich parametrov pohybuje medzi dvomi až deviatimi. Niektoré zákony úmrtnosti sa vzťahujú pre celé vekové rozpätie, iné sú použiteľné iba pre konkrétny vekový interval. Využitie nachádzajú napríklad pri vyrovnávaní údajov, zmenšovani štatistických chýb aj pri tvorbe úmrtnostných tabuliek. Uplatniteľné sú aj pri dopĺňaní neúplných údajov.

Tabuľka č. 2 - Zákony úmrtnosti

Autor	Rok	Model
<i>De Moivre</i>	1725	$\mu_x = \frac{1}{\omega - x}$
<i>Gompertz</i>	1825	$\mu_x = Bc^x$
<i>Makeham</i>	1867	$\mu_x = A + Bc^x$
<i>Makeham</i>	–	$\mu_x = A + Hx + Bc^x$
<i>Oppermann</i>	≤1870	$\mu_x = Ae^{Bx} + (C + Dx)e^{-Ex}$
<i>Oppermann</i>	1870	$\mu_x = Ax^{-\frac{1}{2}} + B + Cx^{\frac{1}{3}}$
<i>Thiele</i>	1871	$\mu_x = Ae^{-Bx} + Ce^{-D(x-E)^2} + FG^x$
<i>Wittstein & Bumstead</i>	1883	$q_x = \frac{1}{B}Ae^{-(Bx)^N} + A^{-(M-x)^N}$
<i>Perks</i>	1932	$\mu_x = \frac{A + Bc^x}{Kc^{-x} + 1} + Dc^x$
<i>Weibull</i>	1951	$\mu_x = kx^n$
<i>Van der Maen</i>	1943	$\mu_x = A + Bx + Cx^2 + \frac{I}{N - x}$
<i>Van der Maen</i>	1943	$\mu_x = A + Bx + \frac{I}{N - x}$
<i>Dvojitý geometrický model</i>	–	$\mu_x = A + Bc^x + Mn^x$
<i>Beard</i>	1971	$q_x = A + \frac{Bc^x}{Ec^{-2x} + 1 + Dc^x}$
<i>Beard-Makeham</i>	1971	$\mu_x = \frac{KAe^{Bx}}{1 + Ae^{Bx}} + C$
<i>Gamma-Gompertz</i>	1979	$\mu_x = \frac{Ae^{Bx}}{1 + \frac{AG}{B}(e^{Bx} - 1)}$
<i>Heligman-Pollard</i>	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^x}$
<i>Heligman-Pollard</i>	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^x}$
<i>Heligman-Pollard</i>	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + KGH^x}$
<i>Heligman-Pollard</i>	1980	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^{x^K}}$
<i>Siler</i>	1983	$\mu_x = Ae^{-Bx} + C + De^{Ex}$
<i>Martinelle</i>	1987	$\mu_x = \frac{Ae^{Bx} + C}{1 + De^{Bx}} + Ke^{Bx}$
<i>Kannisto</i>	1994	$\mu_x = \frac{B \cdot e^{\alpha x}}{1 + B \cdot e^{\alpha x}}$
<i>Kannisto-Makeham</i>	–	$\mu_x = \frac{Ae^{Bx}}{1 + Ae^{Bx}} + C$
<i>Coale-Kisker</i>	–	$m_x = e^{Ax^2 + Bx + C}$

Zdroj: vlastné spracovanie

Za priekopníkov zákonov úmrtnosti možno považovať Abrahama de Moivre, ktorý predpokladal rovnomerné rozloženie úmrtí v celom vekovom intervale a Benjamina Gompertza, ktorého zákon úmrtnosti je dodnes jedným z najpoužívanějších, aj z dôvodu jeho jednoduchosti. Postupom času sa deterministické modely ďalej vyvíjali a upravovali. Matematici častokrát vychádzali z už existujúcich zákonov. Keďže mali viac informácií o úmrtnostnom správaní populácii, boli schopní pridaním alebo úpravou parametrov odstrániť nedostatky pôvodných modelov. Okrem toho vznikali aj nové zákony úmrtnosti, ktoré využívali iné predpoklady.

Tabuľka č. 2 obsahuje zákony úmrtnosti spolu s menom autora a rokom objavenia (pokiaľ je známy). Ide o parametrické funkcie opisujúce úmrtnosť populácie v pevnom časovom bode. Skutočnú úmrtnosť však môžeme považovať za stochastickú, pretože sa neustále vyvíja. Keďže zákony úmrtnosti sú statické, odporúča sa ich parametre určovať periodicky, aby vyhovovali vývoju úmrtnostného správania populácie. Parametre týchto zákonov nám potom môžu poslúžiť na porovnanie zmien úmrtnosti populácie v čase, prípadne navzájom medzi populáciami rôznych krajín.

3.1.1.1 Gompertzov - Makehamov zákon úmrtnosti

Benjamin Gompertz bol anglický matematik, ktorý sa venoval štúdiu ľudskej populácie. Dospel k záveru, že pravdepodobnosť úmrtia sa zdvojnásobuje každých 8 rokov a v roku 1825 ako jeden z prvých matematicky vyjadril závislosť úmrtnosti od veku, čím zostrojil zákon úmrtnosti, ktorý patrí stále k najpoužívanejším modelom úmrtnosti. Jeho zákon úmrtnosti vychádza z predpokladu, že odolnosť človeka voči smrti klesá s rastúcim vekom alebo inak, že úmrtnosť rastie exponenciálne v závislosti od veku. Pre intenzitu úmrtnosti platí tento vzťah

$$\mu_x = B \cdot c^x$$

v ktorom je intenzita úmrtnosti vyjadrená pomocou dvoch reálnych parametrov B a c , pre ktoré platí $B > 0$ a $c > 1$. Záver, ku ktorému došiel Gompertz, je platný pre veľký počet krajín, pre rôzne časové obdobia a dokonca tiež pre rôzne organizmy. Hoci sa skutočná priemerná dĺžka života medzi krajinami a organizmami odlišuje, stále platí pravidlo, že pravdepodobnosť úmrtia sa zdvojnásobuje každých x rokov.

Už Gompertz si uvedomoval, že zákon úmrtnosti by mohol okrem exponenciálne rastúcej úmrtnosti v závislosti od veku obsahovať aj člen, ktorý by nebol závislý od veku. Keďže však informácie o úmrtnostnom správaní populácie neboli úplné, rozhodol sa ponechať len exponenciálnu zložku. Až o 35 rokov neskôr, v roku 1860 tento zákon úmrtnosti doplnil o prvok nezávislý od veku ďalší poistný matematik *William Makeham*. Podľa neho je tento prvok pomenovaný ako Makehamov parameter a obvykle sa označuje písmenom A pričom platí, že nadobúda hodnoty väčšie ako 0. Tento parameter slúži na posun grafu Gompertzovej funkcie μ_x o konštantu A . A tak vznikol vzťah, ktorý je dnes označovaný ako Gompertzov-Makehamov zákon úmrtnosti

$$\mu_x = A + B \cdot c^x \quad (3.3)$$

Výhodou tohto deterministického modelu úmrtnosti je najmä jeho jednoduchosť, či už spojená s interpretáciou, s pomerne malým počtom parametrov alebo aj s faktom, že k ich odhadu nie je potrebný špeciálny softvér. Naopak slabou stránkou modelu je jeho prispôbovanie sa vyšším vekom, v ktorých mimoriadne nadhodnocuje úmrtnosť. Tento fakt si uvedomovali aj samotný tvorcovia modelu, Gompertz a Makeham, ktorý tvrdili, že ich vzťah pre intenzitu úmrtnosti (3.3) je vhodný pre vekový interval v rozmedzí od 20 až do 60 rokov. V súčasnosti sa model používa pre široké spektrum organizmov, vrátane ľudí, pričom vekový interval je posunutý k vyšším vekom, čo je spôsobené vývojom úmrtnostného správania.

3.1.1.2 Kannistov - Makehamov zákon úmrtnosti

Stále populárnejšími pri modelovaní úmrtnosti sa stávajú logistické modely, ku ktorým patrí aj Kannistov-Makehamov zákon úmrtnosti. Pôvodný tvar funkcie intenzity úmrtnosti navrhol *Väinö Kannisto* v roku 1994. Jeho cieľom bolo zohľadnenie pozorovaného spomalenia úmrtnosti v pokročilom veku. Jeho vzťah pre intenzitu úmrtnosti vyzeral takto

$$\mu_x = \frac{Ae^{Bx}}{1 + Ae^{Bx}} \quad (3.4)$$

a obsahoval dva parametre, označované A, B . Pre tento model platí, že jeho parametre nadobúdajú nezáporné hodnoty, pričom intenzita úmrtnosti sa asymptoticky približuje k 1.

Makeham neskôr pridal k vzťahu (3.4) nový parameter C , ktorý rovnako ako pri Gompertzovom-Makehamovom zákone má za následok posun krivky intenzity úmrtnosti o konštantu C v smere osi y . Platí

$$\mu_x = \frac{Ae^{Bx}}{1 + Ae^{Bx}} + C \quad (3.5)$$

V prípade, že by sme chceli Kannistov-Makehamov model previesť do logistickej podoby, dostali by sme vzťah

$$\text{logit}(q_x) = \ln(A) + Bx$$

Medzi výhody Kannistovho-Makehamovho modelu patrí jeho flexibilita, pretože pozoruhodne opisuje aktuálny trend úmrtnostného správania, pričom vykazuje veľmi dobrú spoľahlivosť pre pokročilé veku až do 120 roku života.

3.1.1.3 Heligmanov - Pollardov zákon úmrtnosti

Larry Heligman a John H. Pollard sú strojcami niekoľkých zložitejších parametrických modelov úmrtnosti, ktoré slúžia na modelovanie úmrtnostného správania populácie na celom vekovom intervale. Tieto modely patria k najpoužívanejším zákonom úmrtnosti v demografii najmä preto, lebo sú biologicky interpretovateľné. Ako prvý predstavili v roku 1980 model v tvare

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x$$

kde A, B, C, D, E, F, G, H je 8 parametrov, ktoré sa snažíme odhadnúť.

Heligmanov-Pollardov zákon úmrtnosti sa skladá z troch členov. Prvý člen $A^{(x+B)^C}$ je zameraný na úmrtnosť detí, pričom parameter A odráža dojčenskú úmrtnosť, parameter B sa zameriava na mieru úmrtnosti detí vo veku 1 roka a parameter C odzrkadľuje úmrtnosť až po dospelosť. Parametre A, B, C sa nachádzajú v intervale hodnôt $\langle 0,1 \rangle$. Druhá časť modelu $De^{-E(\ln x - \ln F)^2}$ sa snaží opísať zvýšenú úmrtnosť u mladých, tzv. “havarijný hrb”. Havarijnému hrbu sme sa venovali v časti 1.2 dizertačnej práce. Výška havarijného hrbu je určená parametrom D , ktorý nadobúda hodnoty v rozpätí od 0 po 1. Šírku havarijného hrbu

udáva parameter E , pre ktorý platí $E > 0$. Parameter F určuje, kde na vekovej osi sa havarijný hrb nachádza. Posledný člen je pri tomto variante Heligmanovho-Pollardovho modelu tvorený Gompertzovým zákonom úmrtnosti a reprezentuje úmrtnostné správanie dospelých ľudí. Hodnoty parametrov tohto člena sú z intervalu $(0, \infty)$. Parameter G pritom odráža úroveň úmrtnosti dospelých a H vyjadruje mieru nárastu úmrtnosti, teda sklon Gompertzovej krivky.

Heligman a Pollard publikovali ešte niekoľko tvarov modelu, ktoré sa vo všeobecnosti líšili iba v poslednom člene a ich najnovší a zároveň najpresnejší model vyzerá takto

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^{xK}} \quad (3.6)$$

Heligmanove-Pollardove zákony úmrtnosti môžeme považovať za alternatívu ku Gompertzovmu-Makehamovmu modelu a ich vhodnosť pre celý vekový interval bola otestovaná na empirických pozorovaniach úmrtnostného správania populácie Austrálie a Spojených štátov amerických.

3.1.1.4 CoDe zákon úmrtnosti

V roku 2016 predstavili holandskí výskumníci *Joop de Beer* a *Fanny Janssen* nový model úmrtnosti. Dali mu pomenovanie CoDe, čo predstavuje skratku anglických pojmov Compression (zhusťovanie) a Delay (oneskorenie), pretože význam tohto modelu spočíva v rozlišovaní efektov zhusťovania (menšia variabilita úmrtí podľa veku) a oneskorenia (posun rozdelenia veku v čase úmrtia na starší vek) úmrtnostného správania ľudskej populácie. Rovnako ako Heligmanov-Pollardov zákon úmrtnosti je tento model aplikovateľný pre celé vekové rozpätie.

CoDe zákon úmrtnosti sa skladá z piatich aditívnych členov reprezentujúcich úmrtnosť v za sebou idúcich fázach života (detská úmrtnosť, úmrtnosť mladistvých a úmrtnosť v období rannej dospelosti, neskorej dospelosti, ale aj v starobe) a zahŕňa modálny vek pri úmrtí ako parameter, ktorý zohľadňuje oneskorenie úmrtnosti. Model opisuje úmrtnosť detí a mladistvých pomocou dvoch jednoduchých matematických funkcií a v rôznych fázach dospelosti používa na opis úmrtnosti zmiešaný logistický model s rôznym sklonom

pre mladých dospelých, dospelých a ľudí vo vyššom veku. CoDe zákon úmrtnosti je definovaný takto

$$q_x = \frac{A}{x+B} + \frac{ae^{(x-m)}}{1+e^{(x-m)}} + I(x \leq M-h) \left[\frac{b_1 e^{b_1(x-M)}}{1 + \frac{b_1}{g} e^{b_1(x-M)}} \right] + \quad (3.7)$$

$$+ I(M-h < x \leq M) \left[\frac{b_2 e^{b_2(x-M)}}{1 + \frac{b_2}{g} e^{b_2(x-M)}} + c_1 \right] + I(x > M) \left[\frac{b_3 e^{b_3(x-M)}}{1 + \frac{b_3}{g} e^{b_3(x-M)}} + c_2 \right]$$

Zo vzťahu (3.7) môžeme vidieť, že CoDe model tvorí až 10 parametrov, pričom šesť z nich sa v čase mení. Prvý člen je zameraný na úmrtnosť v nízkom veku a predpokladá inverzný vzťah medzi úmrtnosťou a vekom. Obsahuje dva parametre, z ktorých A odráža dojčenskú úmrtnosť, B reprezentuje detskú úmrtnosť a zároveň je ovplyvnený mierou poklesu úmrtnosti detí po prvom roku života. Efekt zmien v hodnote parametra B je v čase relatívne malý, preto zmeny úmrtnosti v detstve možno prisúdiť najmä zmenám parametra A . Interpretácia tohto parametra je potom jednoduchá, čím nižšia je hodnota parametra A , tým nižšia je úroveň detskej úmrtnosti.

Havarijný hrb sa prejavuje prudkým nárastom miery úmrtnosti počas rokov dospievania, ktorá sa následne ustáli ešte pred začiatkom dospelosti. Tento jav je v modeli zachytený pomocou druhého člena CoDe zákona úmrtnosti $\frac{ae^{(x-m)}}{1+e^{(x-m)}}$. Parameter m predstavuje vek, v ktorom je nárast miery úmrtnosti najprudší a a slúži na opis úmrtnosti spôsobenej ďalšími faktormi. V našich údajoch je najprudší nárast viditeľný od veku 16 rokov, čo korešponduje aj s krajinami, na ktorých bol CoDe model testovaný (Japonsko, Francúzsko, Dánsko).

Na rozlíšenie rôznych fáz úmrtnosti v dospelosti sme použili funkciu indikátora $I(*)$, kde M je hodnota modálneho veku pri úmrtí a h si zvolíme 30, aby jednotlivé životné fázy boli dobre interpretovateľné. Priemerná hodnota modálneho veku pre naše údaje predstavuje okolo 80 rokov života, z čoho vyplýva, že parameter b_1 reprezentuje sklon úmrtnosti do veku 50 rokov (skorá dospelosť), b_2 opisuje úroveň úmrtnosti medzi 50 až 80 rokom života (neskorá dospelosť) a parameter b_3 vyjadruje nárast úmrtnosti v závislosti od veku pre osoby nad 80 rokov (staroba). Súčasťou modelu sú aj dve konštanty c_1 a c_2 , ktoré sa vypočítajú

z ostatných parametrov a slúžia na to, aby medzi úrovňami úmrtnosti v rôznych fázach života bol plynulý prechod. Ich hodnotu dostaneme dosadením do nasledujúcich vzťahov

$$c_1 = \frac{b_1 e^{b_1(-h)}}{1 + \frac{b_1}{g} e^{b_1(-h)}} - \frac{b_2 e^{b_2(-h)}}{1 + \frac{b_2}{g} e^{b_2(-h)}} \quad (3.8)$$

$$c_2 = \frac{b_2}{1 + \frac{b_2}{g}} + c_1 - \frac{b_3}{1 + \frac{b_3}{g}} \quad (3.9)$$

Posledným parametrom je parameter g , ktorý predstavuje hornú hranicu miery úmrtnosti, teda hodnotu, ku ktorej sa približuje pravdepodobnosť úmrtia osoby vo veku nad 110 rokov. Pre súbor krajín so spoľahlivými údajmi odhadol Gampe (2010), že pravdepodobnosť úmrtia nad 110 rokov sa rovná 0,5. HMD (Human mortality database) predpokladá, že vekovo špecifická centrálna miera úmrtnosti má hornú hranicu rovnú jednej (Wilmoth a kol., 2017), čo by znamenalo, že horná hranica pravdepodobnosti úmrtia sa rovná 0,632 (podľa vzťahu (3.2)). Keďže pravdepodobnosti úmrtia sa približujú k hornej hranici vo veľmi vysokom veku a v dizertačnej práci používame údaje o úmrtnosti do veku 100 rokov, očakáva sa, že náš predpoklad pre parameter g bude mať minimálny vplyv na výsledok, preto použijeme rovnaký predpoklad ako samotný predstaviteľ modelu Beer a Janssen (2016), podľa ktorých sa táto hodnota rovná približne 0,7. Po doplnení všetkých časovo nezávislých parametrov, bude mať výsledný CoDe zákon úmrtnosti nasledovný tvar

$$q_x = \frac{A}{x+B} + \frac{ae^{(x-16)}}{1+e^{(x-16)}} + I(x \leq 50) \left[\frac{b_1 e^{b_1(x-80)}}{1 + \frac{b_1}{0,7} e^{b_1(x-80)}} \right] +$$

$$+ I(50 < x \leq 80) \left[\frac{b_2 e^{b_2(x-80)}}{1 + \frac{b_2}{0,7} e^{b_2(x-80)}} + c_1 \right] + I(x > 80) \left[\frac{b_3 e^{b_3(x-80)}}{1 + \frac{b_3}{0,7} e^{b_3(x-80)}} + c_2 \right].$$

3.1.2 Stochastické modely

Stochastické modely sa kvôli zjednodušeniu zvyknú označovať M1, M2, atď. a sú zosumarizované v tabuľke č. 3. Niektoré z týchto modelov si vyžadujú určité obmedzenia parametrov. Počet obmedzení parametrov požadovaný pre každý model je uvedený v zátvorkách, taktiež v tabuľke č. 3.

Tabuľka č. 3 - Základné stochastické modely ľudskej úmrtnosti

M1: Lee-Carter model (1992)	
$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$	(2 obmedzenia)
M2: Renshaw-Haberman model (2006)	
$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} + \beta_x^{(3)} \gamma_{t-x}^{(3)}$	(4 obmedzenia)
M3: Currie Age-Period-Cohort model APC (2006)	
$\ln(m_{x,t}) = \beta_x^{(1)} + n_\alpha^{-1} \kappa_t^{(2)} + n_\alpha^{-1} \gamma_{t-x}^{(3)}$	(3 obmedzenia)
M4: Currie, Durban, and Eilers model P-splines (2004)	
$\ln(m_{x,t}) = \sum_{i,j} \theta_{ij} B_{ij}(t, x)$	(bez obmedzení)
M5: Cairns-Blake-Dowd model CBD-1 (2006)	
$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$	(bez obmedzení)
M6: Cairns et al. model (2009)	
$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}$	(2 obmedzenia)
M7: Cairns et al. model CBD-2 (2007)	
$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)}$	(3 obmedzenia)
M8: Cairns et al. model CBD-3 (2007)	
$\text{logit}(q_{x,t}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}(x_c - x)$	(1 obmedzenie)

Zdroj: vlastné spracovanie

V stochastických modeloch budeme používať nasledujúce označenia parametrov:

- β_x odráža efekt úmrtnosti spojený s vekom. Vekové efekty sa týkajú zmien v životných skúsenostiach v dôsledku chronologického veku.
- κ_t reprezentuje efekt periódy. Je výsledkom vonkajších faktorov, ktoré v určitom kalendárnom čase rovnako ovplyvňujú všetky vekové skupiny. Môže vzniknúť z radu environmentálnych, sociálnych a ekonomických faktorov, napr. vojna, hladomor, hospodárska kríza. Tento parameter je závislý od času.
- γ_c odzrkadľuje efekt týkajúci sa kohorty, pričom $c = t - x$ označuje rok narodenia. Kohorta predstavuje skupinu osôb, u ktorej došlo v rovnakom čase k nejakej významnej demografickej udalosti. Medzi efekty kohorty patria spoločenské

a historické zmeny ovplyvňujúce konkrétnu skupinu so zdieľanou udalosťou (napr. sobáš alebo vojna).

Modely M1 až M3 môžeme označiť ako modely patriace do rodiny zovšeobecnených modelov Lee-Carter a modely M5 – M8 patria k členom rodiny zovšeobecnených modelov CBD-Perks. Všetky modely M1 – M3 a M5 – M8 zdieľajú rovnaký základný predpoklad, že vekové, periódové a kohortné efekty sú svojou povahou kvalitatívne odlišné, a preto je potrebné ich modelovať rôznymi spôsobmi. Naproti tomu modely P-splines, M4, predpokladajú, že v podkladových povrchoch úmrtnosti je hladkosť tak v periódových, ako aj vo vekových a kohortných efektoch. Modely M5 – M8 sa líšia od modelov M1 – M3 tým, že predpokladajú funkčný vzťah (a teda hladkosť) medzi mierami úmrtnosti v susedných vekových skupinách v rovnakom roku.

Od všetkých stochastických modelov sa požaduje, aby parametre závislé od času boli vzhľadom k predĺženiu modelovaného obdobia úplne robustné. To znamená, že keď budú k dispozícii údaje o úmrtnosti za ďalší rok, tak sa podľa nich aktualizuje stochastický model, no časovo závislé parametre v predchádzajúcich rokoch to neovplyvní. Táto vlastnosť, ktorú nazývame invariantnosť voči novým údajom, je zásadne dôležitá, pretože sledovanie vývoja týchto parametrov by bolo nemožné, ak by sa menili ich historické hodnoty.

3.1.2.1 Lee-Carter trieda modelov

Ronald Demos Lee a Lawrence Carter predstavili prvý stochastický model úmrtnosti v roku 1992. Tento model spočíval v kombinácii demografického modelu úmrtnosti s metódami prognózovania časových radov. Lee a Carter (1992) predpokladali, že logaritmy centrálnych mier úmrtnosti vyhovujú vzťahu

$$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} + \varepsilon_{x,t} \quad (3.10)$$

Parameter $\beta_x^{(1)}$ opisuje všeobecný tvar logaritmu centrálnej miery úmrtnosti naprieč vekom. Tento parameter nie je závislý od času a predstavuje priemernú úmrtnosť ľudí v konkrétnom veku x . Jediný prvok závislý od času je parameter $\kappa_t^{(2)}$, ktorý charakterizuje fluktuáciu úrovne úmrtnosti v závislosti od času. Ak poklesne, znamená to, že úmrtnosť sa zlepšila a naopak, ak vzrastie, svedčí to o zhoršujúcej sa úrovni úmrtnosti. Parameter $\beta_x^{(2)}$

určuje, ako táto zmena úrovne úmrtnosti ovplyvní centrálnu mieru úmrtnosti $m_{x,t}$ v konkrétnom veku x . Ak je vysoká v porovnaní s určitým vekom, znamená to, že úmrtnosť sa v tomto veku zlepšuje rýchlejšie ako v iných vekových skupinách. Naopak, ak by bol tento parameter negatívny, znamenalo by to, že úmrtnosť sa v týchto vekoch zhoršuje. Člen $\varepsilon_{x,t}$ odráža náhodnú zložku spôsobenú historickými vplyvmi, ktoré model nezachytil.

Pri odhadovaní parametrov musíme brať do úvahy obmedzenia, ktorých výber nemá vplyv na kvalitu modelu, preto je ich stanovenie individuálne. V literatúre sa však najčastejšie stretávame s obmedzeniami, ktoré navrhli samotní autori modelu, a teda, že $\sum_x \beta_x^{(2)} = 1$ a $\sum_t \kappa_t^{(2)} = 0$.

Atraktivita tohto modelu súvisí najmä s jeho jednoduchosťou, avšak má aj nejaké slabé stránky. Hlavná nevýhoda modelu spočíva v tom, že má iba jeden faktor, čo vedie k dokonalej korelácii vývoja úmrtnosti vo všetkých vekových skupinách (triviálna korelačná štruktúra). Lee-Carterov model neobsahuje kohortný efekt a v prípade krajín, kde bol tento efekt pozorovaný v minulosti, tento model nie je vhodný na modelovanie, pretože nezodpovedá historickým údajom. Neistota budúcich mier úmrtnosti je proporcionálna k priemernej miere zlepšenia β_x , ktorá pre vysoký vek môže viesť k tomu, že táto neistota bude príliš nízka, pretože miera historického zlepšenia bola vo vysokom veku často nižšia. Aj napriek nevýhodám tohto modelu, jeho jednoduchosť viedla k tomu, že sa považuje za meradlo, na základe ktorého je možné hodnotiť ďalšie stochastické modely úmrtnosti. Existuje značné množstvo literatúry dopĺňujúcej alebo upravujúcej Lee-Carterov model, napr. Booth a kol. (2006), Hainaut (2012) alebo Wolf (2004). *Arthur Renshaw* a *Steven Haberman* v roku 2006 upravili Lee-Carterov model jednoduchým pridaním faktora na zachytenie kohortového efektu (M2). Tento model sa oveľa lepšie hodí pre krajiny ako Spojené kráľovstvo, kde bol identifikovaný kohortný efekt. Trpí však nedostatočnou robustnosťou, pravdepodobne kvôli prítomnosti viac ako jedného lokálneho maxima vo funkcii maximálnej vierohodnosti.

Iain Currie (2006) v publikácii poznamenal, že aj keď model obsahuje kohortný efekt, pre väčšinu simulovaných mier úmrtnosti je korelačná štruktúra stále triviálna, pričom simulované kohortné parametre sú relevantné iba pre vyšší vek na vzdialenom konci projekcie. Na základe tejto analýzy Currie použil na úmrtnosť zjednodušený model kohorty podľa veku (APC model – M3), ktorý odstránil problém robustnosti, ale na úkor presnosti

(kvality) modelu. Model APC má pôvod v lekárskej štatistike a predchádza modelu Lee-Cartera.

3.1.2.2 CBD-Perks trieda modelov

Andrew Cairns, David Blake a Kevin Dowd (2006) spozorovali, že pre údaje z Anglicka, Walesu a Spojených štátov sa javilo, že efekt kohorty odhadovaných mier úmrtnosti má trend podľa roku narodenia, a preto predstavili stochastický model, ktorý obsahuje dva časovo závislé parametre. Model využíva relatívnu jednoduchosť (t. j. takmer lineárnosť logaritmu $q_{x,t}$) krivky miery úmrtnosti vo vyššom veku. Hlavná nevýhoda modelu CBD spočíva v tom, že bol navrhnutý pre vyšší vek, a teda ignoruje potreby modelovania úmrtnosti v nižšom veku (napríklad havarijný hrb alebo dojčenskú úmrtnosť). V publikácii Cairns a kol. (2009) to autori argumentovali tým, že významné náklady spojené s úmrtnosťou sú v staršom veku, a preto sa ich modelovanie zameralo práve na vyšší vek. Model priamo využíva pravdepodobnosť úmrtia $q_{x,t}$ namiesto centrálnej miery úmrtnosti $m_{x,t}$ a môžeme ho vyjadriť v tvare logitovej, resp. logaritmickej funkcie takto

$$\text{logit}(q_{x,t}) = \ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \varepsilon_{x,t} \quad (3.11)$$

kde

- x je vek ($x = x_1, x_2, \dots, x_p$),
- t je rok ($t = t_1, t_2, \dots, t_q$),
- $q_{x,t}$ je pravdepodobnosť úmrtia, ktorá predstavuje pravdepodobnosť, že jedinec vo veku x v čase t zomrie v priebehu jedného roku (v časovom intervale medzi t a $t + 1$),
- $\ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right)$ reprezentuje logistickú transformáciu $q_{x,t}$, pričom $\frac{q_{x,t}}{1-q_{x,t}}$ vyjadruje šancu na úmrtie,
- $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$ sú dva stochastické parametre a predstavujú dva časové indexy modelu,
- $\bar{x}$ je priemerný modelovaný vek,
- $\varepsilon_{x,t}$ odráža náhodnú zložku spôsobenú historickými vplyvmi, ktoré model nezachytil.

Vzhľadom na jednoduchú štruktúru je možné indexy úmrtnosti, teda parametre CBD modelu, jednoducho interpretovať. Časový index $\kappa_t^{(1)}$ je úrovňová konštanta modelu, ktorá reprezentuje úroveň úmrtnosti v čase t a ovplyvňuje každý vek rovnakým spôsobom. Zníženie hodnoty $\kappa_t^{(1)}$ teda spôsobí paralelný posun logisticky transformovanej krivky úmrtnosti nadol. Inak povedané, ak sa časom znižuje, znamená to, že úmrtnosť v priebehu času klesala u všetkých vekových skupín. Parameter $\kappa_t^{(2)}$ predstavuje sklon logisticky transformovanej krivky. Každý vek je týmto parametrom ovplyvnený inak. Zvýšenie indexu úmrtnosti $\kappa_t^{(2)}$ spôsobí nárast strmosti logisticky transformovanej krivky úmrtnosti, čo znamená, že úmrtnosť (v logistickej mierke) v mladšom veku (pod priemerným vekom $\bar{x}$) sa zlepšuje rýchlejšie ako v staršom veku (nad priemerným vekom $\bar{x}$). Model nevyžaduje ďalšie obmedzenia. Na odhadnutie parametrov CBD modelu sa môže použiť metóda najmenších štvorcov. V tomto prípade sa dvojica parametrov $(\kappa_t^{(1)}, \kappa_t^{(2)})$ odhaduje zvlášť pre každé obdobie t , pričom presnosť odhadov rastie so zväčšovaním rozsahu vstupných údajov (s dlhším obdobím analýzy).

CBD model neobsahuje žiadny kohortný efekt, čo však napravili Cairns a kol. (2009), ktorí pridali kohortný efekt navyše k existujúcim dvom periódovým efektom (M6). Neskôr sa vytvorila aj ďalšia úprava umožňujúca, aby sa kohortný efekt časom zmenšoval (M8). Okrem toho existuje ešte jeden model z triedy CBD-Perks modelov, ktorý je zovšeobecnením modelu M6 a obsahuje navyše ešte kvadratický výraz kohortného efektu. Zahrnutie kvadratického výrazu je inšpirované možným zakrivením identifikovaným v grafoch *logit* ($q_{x,t}$).

3.1.2.3 Ďalšie stochastické modely

Prístup, ktorý sa stal populárnym vo Veľkej Británii, je použitie penalizovaných splines (P-splines – M4), ktoré navrhli v roku 2004 Iain Currie, Maria Durbán a Paul Eilers. Všeobecný model má tvar

$$\ln(m_{x,t}) = \sum_{i,j} \theta_{ij} B_{ij}(t, x)$$

kde $B_{ij}(t, x)$ sú vopred určené základné funkcie s pravidelne rozmiestnenými uzlami a θ_{ij} sú parametre, ktoré je potrebné odhadnúť.

Okrem základných stochastických modelov úmrtnosti sa v neskoršom období objavili také, ktoré sa snažili spojiť pozitívne aspekty už existujúcich modelov. Výsledkom toho bol napríklad Richardov Platov štvorfaktorový model, ktorý vychádzal z modelu Lee-Cartera, a ktorý pridal faktory na zachytenie druhého vekovo periodového efektu podľa modelu CBD a kohortového efektu podľa modelu Renshawa a Habermana. Inovácia v Platovom modeli potom spočívala v pridaní ďalšieho periodového efektu ovplyvňujúceho iba nižšie vekové kategórie a je navrhnutý tak, aby umožňoval prispôbenie modelu celému vekovému intervalu. Colin O'Hare a Youwei Li (2012) upravili Platov model tak, aby lepšie vyhovoval širšiemu vekovému rozpätiu. Vylepšili Platovu špecifikáciu pridaním kvadratického vekového efektu pre mladšie veky. Pre viac informácií ohľadom stochastických modelov úmrtnosti odporúčame napríklad publikáciu O'Hare a Li (2017).

3.2 Metódy odhadu parametrov deterministických modelov úmrtnosti

Zákony úmrtnosti patria k parametrickým modelom a na ich odhad môžeme použiť metódy nelineárnej regresnej analýzy. Cieľom nelineárnej regresnej analýzy je nájsť vzťah medzi závislou premennou (v našom prípade miera úmrtnosti q_x , intenzita úmrtnosti μ_x alebo centrálna miera úmrtnosti m_x) a jednou alebo viacerými nezávislými premennými (vek). V softvéri SAS Enterprise Guide sú dostupné tri najpoužívanejšie iteračné metódy: Gaussova-Newtonova metóda, Gradientná metóda a Levenbergova-Marquardtova metóda. Pri týchto metódach uvažujeme o nelineárnej regresnej funkcii s vektorom parametrov $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Podstata metód spočíva v nájdení takých odhadov parametrov regresnej funkcie, pre ktoré platí, že súčet štvorcov odchýlok regresnej funkcie od pozorovaných hodnôt je čo najmenší.

$$S(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2 \rightarrow \min$$

Východiskovým bodom metód sú počiatočné (vstupné) odhady parametrov $\theta^0 = (\theta_1^0, \theta_2^0, \theta_3^0, \dots, \theta_k^0)$, ktoré sa v jednotlivých iteráciách vylepšujú. V-te vylepšené odhady získané vo v-tej iterácii, pričom $v = 0, 1, 2, \dots$, sa označujú

$\boldsymbol{\theta}^v = (\theta_1^v, \theta_2^v, \theta_3^v, \dots, \theta_k^v)$. V každej iterácii sa počíta hodnota funkcie $S(\boldsymbol{\theta}^v)$, pričom by malo platiť $S(\boldsymbol{\theta}^{v+1}) \leq S(\boldsymbol{\theta}^v)$, čo znamená, že regresná funkcia s parametrami odhadnutými v iterácii $(v + 1)$ lepšie opisuje cieľovú premennú (vysvetľovanú premennú) ako regresná funkcia s parametrami odhadnutými vo v -tej iterácii.

3.2.1 Gaussova - Newtonova metóda

Gaussova-Newtonova metóda patrí medzi iteračné metódy, ktoré využívajú deriváciu funkcie. Nazýva sa niekedy aj metóda linearizácie, pretože vo v -tej iterácii využíva lineárnu aproximáciu funkcie $f(\mathbf{x}_i, \boldsymbol{\theta})$ prostredníctvom Taylorovho rozvoja 1. rádu v bode $\boldsymbol{\theta}^v$.

$$f(\mathbf{x}_i, \boldsymbol{\theta}) \cong f(\mathbf{x}_i, \boldsymbol{\theta}^v) + \sum_{j=1}^k \left\{ \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^v} \cdot (\theta_j - \theta_j^v) \right\} + \varepsilon_i \quad (3.12)$$

každej iterácii ($v = 0, 1, 2, \dots$) sa urobia substitúcie:

$$\begin{aligned} f_i^v &= f(\mathbf{x}_i, \boldsymbol{\theta}^v) \\ \beta_j^v &= \theta_j - \theta_j^v \\ J_{ij}^v &= \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^v} \end{aligned}$$

čo nám umožňuje nelineárny regresný model prepísať do lineárneho tvaru

$$y_i - f_i^v = \sum_{j=1}^k \beta_j^v J_{ij}^v + \varepsilon_i \quad (3.13)$$

Rovnicu (3.13) vieme tiež zapísať ako

$$\mathbf{y}^v = \mathbf{J}^v \boldsymbol{\beta}^v + \boldsymbol{\varepsilon}^v \quad (3.14)$$

Vektor parametrov $\boldsymbol{\beta}^v$ tohto regresného modelu sa odhadne metódou najmenších štvorcov,

$$\hat{\boldsymbol{\beta}}^v = (\mathbf{J}^{vT} \mathbf{J}^v)^{-1} \cdot \mathbf{J}^{vT} \cdot \mathbf{y}^v = (\mathbf{J}^{vT} \mathbf{J}^v)^{-1} \cdot \mathbf{J}^{vT} \cdot (\mathbf{y} - \mathbf{f}^v) \quad (3.15)$$

a z jeho odhadu $\hat{\boldsymbol{\beta}}^v$ sa vypočíta vektor vylepšených odhadov pre iteráciu $(v + 1)$ takto:

$$\hat{\boldsymbol{\theta}}^{v+1} = \hat{\boldsymbol{\theta}}^v + \hat{\boldsymbol{\beta}}^v \quad (3.16)$$

Iteračný postup pokračuje až dovtedy, kým nie je splnené konvergenčné kritérium, teda pokiaľ neplatí

$$\left| \frac{\hat{\theta}_j^{v+1} - \hat{\theta}_j^v}{\hat{\theta}_j^v} \right| < \delta \quad (3.17)$$

kde δ je vopred určené malé kladné číslo, najčastejšie 10^{-6} .

Gaussova-Newtonova metóda dosahuje dobré výsledky, ak bol dobrý počiatočný odhad parametrov, ak je regresná funkcia takmer lineárna alebo v prípade, že majú rezíduá striedavé znamienka. Nevýhodou tejto metódy je, že niekedy konverguje veľmi pomaly, čo vedie k veľkému počtu iterácií. V určitých prípadoch môže dokonca oscilovať. Vtedy sa rast a pokles súčtu štvorcov $S(\theta)$ opakuje, až pokiaľ sa hodnota nestabilizuje.

3.2.2 Gradientná metóda

Základná stratégia gradientnej metódy inak nazývanej aj metóda najstrmšieho zostupu spočíva v iteračnom spôsobe hľadania globálneho minima funkcie $S(\theta)$ priamou minimalizáciou. Začína sa v začiatočných odhadoch $\theta^0 = (\theta_1^0, \theta_2^0, \theta_3^0, \dots, \theta_k^0)$, pre ktoré sa vypočíta súčet štvorcov $S(\theta)$. Hodnota parametra sa zvýši o malú hodnotu. Ak súčet štvorcov poklesne, pokračuje sa vo zvyšovaní hodnoty parametra. Ak však hodnota $S(\theta)$ vzrastie, hodnota parametra sa vráti na pôvodnú úroveň a následne sa zníži. Tento postup sa opakuje veľakrát, pričom každý krok by mal viesť k zníženiu hodnoty súčtu štvorcov $S(\theta)$. Ak hodnota $S(\theta)$ namiesto toho vzrastie, tak krok bol veľký a preskočilo sa minimum funkcie $S(\theta)$. V takomto prípade sa veľkosť kroku zmenší a pokračuje sa uvedeným algoritmom. Proces sa končí, ak sa nájde bod, ktorý je podľa konvergenčného kritéria dostatočne blízky globálnemu minimu funkcie $S(\theta)$.

Gradientná metóda je typ iteračnej metódy, ktorá využíva gradient funkcie. Funkcia $S(\theta)$ klesá v smere záporného gradientu. V prípade, že máme hodnotu súčtu štvorcov funkcie $f(x_i, \theta)$, potom smer funkcie sa pohybuje v smere

$$\frac{1}{2} \frac{\partial S(\theta)}{\partial \theta^i} = -J^T [y - f(\theta^i)] \quad (3.18)$$

kde matica J je Jakobián funkcie $f(x_i, \theta)$.

Všeobecný zápis na určenie vektora vylepšených odhadov pre iteráciu ($v + 1$) vyzerá takto

$$\boldsymbol{\theta}^{v+1} = \boldsymbol{\theta}^v + \lambda \mathbf{J}^T \cdot [\mathbf{y} - f(\boldsymbol{\theta}^v)] \quad (3.19)$$

kde λ predstavuje skalár, ktorý zvolíme tak, aby platilo

$$S(\boldsymbol{\theta}^v + \lambda \mathbf{J}^T [\mathbf{y} - f(\boldsymbol{\theta}^v)]) < S(\boldsymbol{\theta}^v) \quad (3.20)$$

Gradientná metóda pracuje efektívne v prvých iteráciách, kde je schopná aj v prípade zlých počiatočných odhadov nájsť vhodný smer k minimu funkcie $S(\boldsymbol{\theta})$. Avšak v posledných iteráciách, v oblasti okolo minima funkcie $S(\boldsymbol{\theta})$, má tendenciu konvergovať pomaly alebo oscilovať. Aj z tohto dôvodu sa gradientná metóda používa častokrát na určenie oblasti, v ktorej je najvyššia pravdepodobnosť dosiahnutia optimálnej hodnoty pri postupe z východiskového počiatočného odhadu optimálnych podmienok.

3.2.3 Levenbergova - Marquardtova metóda

Levenbergova-Marquardtova metóda je rovnako ako Gaussova-Newtonova metóda modifikáciou jednoduchšej linearizácie. Túto metódu odhadu parametrov v roku 1944 publikoval *Kenneth Levenberg*. Neskôr v roku 1963 ju znovuobjavil americký štatistik *Donald Marquardt*. Občas zvykne byť označovaná pojmom Marquardtov kompromis, pretože sa pokúša sklbiť výhody vyššie spomínaných metód (Gaussovej-Newtonovej metódy a Gradientnej metódy).

Rovnako ako pri predošlých metódach vychádzame z počiatočného vektora parametrov $\boldsymbol{\theta}^0$. Všeobecne môžeme nasledujúcu iteráciu vektora vylepšených odhadov zapísať takto

$$\boldsymbol{\theta}^1 = \boldsymbol{\theta}^0 + \delta \quad (3.21)$$

pričom δ je parameter, na určenie ktorého je funkcia $f(\mathbf{x}_i, \boldsymbol{\theta} + \delta)$ aproximovaná linearizačným vzťahom

$$f(\mathbf{x}_i, \boldsymbol{\theta} + \delta) \approx f(\mathbf{x}_i, \boldsymbol{\theta}) + \mathbf{J}\delta \quad (3.22)$$

v ktorom $J = \frac{\partial f(x_i, \theta)}{\partial \theta_j}$ určuje gradient funkcie $f(x_i, \theta)$ vzhľadom na vektor parametrov θ .

Vo funkcii minimálnej sumy štvorcov $S(\theta)$ bude gradient vzhľadom na vektor θ nulový. Pomocou aproximácie (3.22) vyjadríme

$$S(\theta + \delta) \approx \|y - f(\theta) - J\delta\|^2 \quad (3.23)$$

Následne, keď funkciu zderivujeme podľa δ a položíme ju rovnú nule, úpravou dostaneme vzťah

$$(J^T J)\delta = J^T [y - f(\theta)] \quad (3.24)$$

výsledkom čoho je sústava lineárnych rovníc, ktorej riešením je parameter δ . (Marquardt, 1963). Levenberg ďalej rozvinul metódu tým, že nahradil rovnicu (3.24), tzv. tlmeným variantom

$$(J^T J + \lambda I)\delta = J^T [y - f(\theta)] \quad (3.25)$$

kde I je jednotková matica, ktorá vyjadruje prírastok δ k odhadu vektora parametrov θ a λ je tlmený faktor, pre ktorý platí $\lambda \geq 0$. V prípade, že dochádza k výraznému úbytku funkcie súčtu štvorcov $S(\theta)$, využívajú sa nižšie hodnoty tlmeného faktora a algoritmus sa približuje ku Gaussovej-Newtonovej metóde. V opačnom prípade, keď pri iterácii neklesá súčet štvorcov dostatočne rýchlo, tlmený faktor λ sa zvyšuje a algoritmus prepne na algoritmus Gradientnej metódy.

Nevýhodou Levenbergovho algoritmu je, že pokiaľ je parameter δ priveľký, nemá význam použitie inverzie matice $(J^T J + \lambda I)$. Preto Marquardt rovnicu (3.25) upravil tak, že jednotkovú maticu I nahradil diagonálnou maticou súčinu $J^T J$. Finálny Levenbergov-Marquardtov algoritmus vieme následne vyjadriť vzťahom

$$(J^T J + \lambda \cdot \text{diag}(J^T J))\delta = J^T [y - f(\theta)] \quad (3.26)$$

pričom $\text{diag}(J^T J)$ je spomínaná diagonálna matica súčinu matíc $J^T J$ (Levenberga, 1944).

Hodnoty tlmeného faktora λ môžu byť rôzne v závislosti od použitého softvéru. V softvéri SAS Enterprise Guide je východiskovou hodnotou $\lambda = 10^{-8}$, pričom pri každej ďalšej iterácii je tlmený faktor násobený desiatkou, až kým neplatí

$$S(\hat{\theta}^{k+1}) < S(\hat{\theta}^k) \quad (3.27)$$

Podstatou výpočtu je zachovať čo najnižšiu možnú hodnotu tlmeného faktora a zároveň zabezpečiť, aby funkcia súčtu štvorcov klesala každou iteráciou. Výsledný vektor odhadnutých parametrov vypočítaný pomocou Levenbergovej-Marquardtovej metódy sa spravidla nachádza medzi vektormi vypočítanými pomocou Gaussovej-Newtonovej a Gradientnej metódy.

3.3 Metódy odhadu parametrov stochastických modelov úmrtnosti

Na odhad parametrov stochastických modelov úmrtnosti môžeme použiť metódu najmenších štvorcov. Vzhľadom na to, že Cairns-Blake-Dowdov model využíva takmer lineárnosť logaritmu miery úmrtnosti $q_{x,t}$ vo vyššom veku, budeme k nemu pristupovať ako k lineárnemu regresnému modelu. Parametre Lee-Carterovho modelu nie je možné odhadnúť pomocou bežných regresných metód, a preto, aby sme minimalizovali súčet štvorcov, použijeme blízku aproximáciu pomocou metódy singulárneho rozkladu matice za predpokladu, že reziduá sú homoskedastické. Časovo závislé parametre týchto modelov sú stochastické procesy, na prognózovanie ktorých použijeme metódu autoregresného integrovaného kľzavého priemeru ARIMA.

3.3.1 Metóda najmenších štvorcov

Metóda najmenších štvorcov je matematicko-štatistická metóda pre aproximáciu dvojice nameraných údajov $[x, y]$. Aproximáciou je myslené čo najlepšie priblíženie sa ku skutočnej hodnote. V dnešnej dobe sa metóda najmenších štvorcov často používa na vyhľadanie alebo odhad numerických hodnôt parametrov na prispôsobenie funkcie množine údajov a na charakterizáciu štatistických vlastností odhadov. Metóda existuje s niekoľkými variáciami:

- obyčajná metóda najmenších štvorcov (OLS – Ordinary least square),
- vážená metóda najmenších štvorcov (WLS – Weithed least square),
- alternatívna metóda najmenších štvorcov (ALS – Alternative least square),
- parciálna metóda najmenších štvorcov (PLS – Partial least square).

Najčastejším, a tiež najstarším použitím metódy najmenších štvorcov je regresná analýza,

ktorá spočíva v nájdení krivky, ktorá najlepšie vyhovuje množine dátových bodov. Pre dvojrozmernú náhodnú premennú $[x, y]$ je úlohou nájsť funkčnú závislosť $y = f(x, \beta_0, \dots, \beta_k)$ a odhadnúť neznáme parametre $\beta_0, \dots, \beta_k$, pričom odhady neznámych parametrov volíme tak, aby minimalizovali rezíduá, ktoré predstavujú rozdiely medzi teoretickými a skutočnými (nameranými) hodnotami premennej y . Keďže rozdiely môžu nadobúdať ako kladné, tak aj záporné hodnoty, umocňujú sa na druhú a počítajú sa ich súčty. Minimalizujeme vlastne funkciu T , čo je súčet kvadratických odchýlok (štvorcov) teoretických a skutočných hodnôt premennej y .

$$T(\beta_0, \dots, \beta_k) = \sum_{i=1}^n [y_i - f(x_i, \beta_0, \dots, \beta_k)]^2 \quad (3.28)$$

Funkcia f sa nazýva regresná funkcia. Ak má regresná funkcia konečné prvé parciálne derivácie podľa všetkých parametrov $\beta_0, \dots, \beta_k$, potom odhady parametrov získame riešením $(k + 1)$ rovníc o $(k + 1)$ neznámych, teda

$$\frac{\partial T}{\partial \beta_i} = 0 \text{ pre } i = 1, \dots, k \quad (3.29)$$

Najjednoduchším prípadom použitia metódy najmenších štvorcov je preloženie (aproximácia) nameraných údajov priamkou (lineárna regresia). Pri lineárnej regresii s jednou premennou je regresná funkcia vyjadrená vzt'ahom

$$\hat{y}_i = \beta_0 + \beta_1 x_i \quad (3.30)$$

Rovnica zahŕňa dva parametre, ktoré vyjadrujú priesečník (β_0) s y -osou a sklon regresnej priamky (β_1), čo je tzv. smernica priamky. Parameter β_0 sa zvykne nazývať aj ako lokujúca alebo úrovňová konštanta a vyjadruje očakávanú úroveň závislej premennej pri nulovej hodnote nezávislej premennej. β_1 charakterizuje zmenu závislej premennej, ktorá zodpovedá zmene nezávislej premennej o jednu jej jednotku. Metóda najmenších štvorcov definuje odhad týchto parametrov ako hodnoty, ktoré minimalizujú výraz

$$T = \sum_{i=1}^n [y_i - \hat{y}_i]^2 = \sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i]^2 \quad (3.31)$$

Matematicky ide o určenie lokálneho extrému funkcie, t. j. nájdenie bodu, v ktorom prvá derivácia funkcie je rovná nule. Preto výraz (3.31) derivujeme podľa jednotlivých parametrov a položíme ho rovný nule. Takto dostaneme sústavu rovníc:

$$\frac{\partial T}{\partial \beta_0} = \frac{\partial (\sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i]^2)}{\partial \beta_0} = 0 \quad (3.32)$$

$$\frac{\partial T}{\partial \beta_1} = \frac{\partial (\sum_{i=1}^n [y_i - \beta_0 - \beta_1 x_i]^2)}{\partial \beta_1} = 0 \quad (3.33)$$

Riešením príslušnej sústavy dvoch rovníc o dvoch neznámych sú odhady parametrov $\hat{\beta}_0$ a $\hat{\beta}_1$ v tvare

$$\hat{\beta}_0 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n y_i x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (3.34)$$

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (3.35)$$

Odhady metódy najmenších štvorcov majú silné štatistické vlastnosti. Konkrétne, keď

- (1) získané údaje tvoria náhodnú vzorku z presne definovanej populácie,
- (2) populačný model je lineárny,
- (3) rezíduum má nulovú očakávanú hodnotu,
- (4) vysvetľujúce premenné sú lineárne nezávislé,
- (5) rezíduá majú normálne rozdelenie, pričom sú tiež nekorelované s nezávislými premennými a majú konštantný rozptyl (takzvaný predpoklad homoskedasticity).

Napriek svojej popularite a všestrannosti má metóda najmenších štvorcov svoje problémy. Pravdepodobne najdôležitejšou nevýhodou je jej vysoká citlivosť na extrémne pozorovania. Je to dôsledok použitia “štvorcov“, pretože druhé umocnenie zväčšuje veľkosť rozdielov, a preto dáva extrémnym pozorovaniam oveľa väčší význam.

Aplikácia metódy najmenších štvorcov na Lee-Carter model

Singulárny rozklad matice (SVD)

Postup, ktorý použili aj samotní autori Lee-Carterovho modelu, je spojením metódy najmenších štvorcov a singulárneho rozkladu matice. Vychádzame zo štatistického modelu v tvare Lee-Carterovho modelu (3.10). Potom funkcia, ktorú minimalizujeme, vyzerá takto

$$Q = \sum_{x=x_1}^{x_p} \sum_{t=t_1}^{t_q} \left(\ln(\hat{m}_{x,t}) - \beta_x^{(1)} - \beta_x^{(2)} \kappa_t^{(2)} \right)^2 \quad (3.36)$$

Parameter $\beta_x^{(1)}$ dostaneme, keď deriváciu funkcie Q podľa $\beta_x^{(1)}$ položíme rovnú 0.

$$\sum_{x=x_1}^{x_p} \sum_{t=t_1}^{t_q} \left(\ln(\hat{m}_{x,t}) - \beta_x^{(1)} - \beta_x^{(2)} \kappa_t^{(2)} \right) = 0 \quad (3.37)$$

Po úpravách pri zohľadnení obmedzení tohto modelu $\sum_x \beta_x^{(2)} = 1$ a $\sum_t \kappa_t^{(2)} = 0$ dostaneme, že parameter $\hat{\beta}_x^{(1)}$ je rovný

$$\hat{\beta}_x^{(1)} = \frac{1}{t_q - t_1 + 1} \sum_{t=t_1}^{t_q} \ln(\hat{m}_{x,t}) \quad (3.38)$$

Na odhad zvyšných parametrov Lee-Carterovho modelu nie je možné použiť jednoduchý regresný model, pretože tento model nemá bežnú štruktúru, a tak nie je možné parametre $\beta_x^{(2)}$ a $\kappa_t^{(2)}$ jednoznačne vyjadriť. Z tohto dôvodu sa na odhad najčastejšie používa singulárny rozklad matice (Singular value decomposition – SVD) (Haberman a Russolillo, 2005). Singulárny rozklad matice je metóda pozostávajúca z transformácie korelovaných premenných na množinu nekorelovaných premenných. Je založená na vete z lineárnej algebry, ktorá hovorí, že maticu $\mathbf{A}$ možno rozložiť na tri matice - ortogonálnu maticu $\mathbf{U}$, diagonálnu maticu $\mathbf{D}$ a transponovanú ortogonálnu maticu $\mathbf{V}$. Keďže už máme odhadnutý parameter $\hat{\beta}_x^{(1)}$, singulárny rozklad matice použijeme na maticu $\mathbf{Z}$, pre ktorú platí

$$\mathbf{Z} = \ln(\hat{m}_{x,t}) - \hat{\beta}_x^{(1)} = \beta_x^{(2)} \kappa_t^{(2)} \quad (3.39)$$

Našou úlohou je nájsť odhady parametrov $\beta_x^{(2)}$ a $\kappa_t^{(2)}$, ktoré minimalizujú funkciu

$$R = \sum_{x=x_1}^{x_p} \sum_{t=t_1}^{t_q} \left(\mathbf{Z} - \beta_x^{(2)} \kappa_t^{(2)} \right)^2 \quad (3.40)$$

Na základe singulárneho rozkladu matice stanovíme maticu $\mathbf{Z}^T \mathbf{Z}$, ktorá má rozmer $[t_q - t_1 + 1] \times [t_q - t_1 + 1]$. Rovnako vymedzíme maticu $\mathbf{Z} \mathbf{Z}^T$ s rozmerom

$[x_p - x_1 + 1] \times [t_q - t_1 + 1]$. Maticu $\mathbf{Z}$ môžeme aproximovať pomocou jej najväčších hodnôt, a preto označme $\mathbf{u}_1$ vlastný vektor prislúchajúci najväčšej vlastnej hodnote vektora matice $\mathbf{Z}^T \mathbf{Z}$ a podobne označíme $\mathbf{v}_1$ vlastný vektor, ktorý zodpovedá najväčšiemu vlastnému číslu matice $\mathbf{Z} \mathbf{Z}^T$. Zodpovedajúcu aproximáciu matice $\mathbf{Z}$ podľa metódy najmenších štvorcov za použitia prvej najväčšej hodnoty predstavuje

$$\mathbf{Z} \approx \mathbf{Z}^* = \sqrt{\lambda_1} \mathbf{v}_1 \mathbf{u}_1^T \quad (3.41)$$

Východiskom aproximácie sú odhady parametrov, ktoré vyzerajú takto

$$\hat{\beta}_x^{(2)} = \frac{\mathbf{v}_1}{\sum_{j=1}^{x_p - x_1 + 1} \mathbf{v}_{1j}} \quad (3.42)$$

$$\hat{\kappa}_t^{(2)} = \sqrt{\lambda_1} \left(\sum_{j=1}^{x_p - x_1 + 1} \mathbf{v}_{1j} \right) \mathbf{u}_1 \quad (3.43)$$

pričom λ_1 predstavuje prvú zložku singulárneho rozkladu matice $\mathbf{Z}$. V podstate je možné zahrnúť do modelu aj viac zložiek singulárneho rozkladu matice, avšak zatiaľ sa nepodarilo dokázať, že by sa tým zlepšila kvalita modelu. Alternatívou k metóde singulárneho rozkladu matice je napríklad Newton-Raphsonová metóda. Pre viac informácií ohľadom Newton-Raphsonovej metódy odporúčame pozrieť publikáciu Pitacco (2009).

Haberman-Russolillo metóda

S použitím blízkej aproximácie singulárneho rozkladu matice (SVD) za predpokladu, že reziduá sú homoskedastické, a že sú splnené obmedzenia $\sum_x \beta_x^{(2)} = 1$ a $\sum_t \kappa_t^{(2)} = 0$, ale stále pomocou aplikovania metódy najmenších štvorcov prišli Haberman a Russolillo (2005) k iným záverom. Podľa nich je postup výpočtu parametrov Lee-Carterovho modelu nasledovný:

1. Výpočet $\hat{\beta}_x^{(1)}$ – vzhľadom na fakt, že tento parameter je fixný v čase, ale rozdielny naprieč vekovými skupinami, je vhodné ho vypočítať rovnako ako pri predchádzajúcej metóde ako priemer všetkých $\beta_x^{(1)}$ v priebehu času.

$$\hat{\beta}_x^{(1)} = \frac{1}{T} \sum_t \ln(m_{x,t}) \quad (3.44)$$

2. Len čo nájdeme $\hat{\beta}_x^{(1)}$ pre každý vek, je možné vypočítať $\hat{\kappa}_t^{(2)}$. Vo všeobecnosti odhadujeme hodnoty parametra $\kappa_t^{(2)}$ pre každý rok t podľa vzťahu:

$$\hat{\kappa}_t^{(2)} = \sum_x (\ln(m_{x,t}) - \beta_x^{(1)}) \quad (3.45)$$

3. $\hat{\beta}_x^{(2)}$ v rovnici Lee-Carterovho modelu (3.10) je v podstate sklon rozdielu medzi logaritmom miery úmrtnosti a $\hat{\beta}_x^{(1)}$ v čase t v konkrétnom veku x a hodnotami parametra $\hat{\kappa}_t^{(2)}$.

$$\hat{\beta}_x^{(2)} = \frac{\sum_t \kappa_t^{(2)} (\ln(m_{x,t}) - \beta_x^{(1)})}{\sum_t \kappa_t^{(2)2}} \quad (3.46)$$

4. V poslednom kroku sa odporúča prehodnotiť $\hat{\kappa}_t^{(2)}$ takým spôsobom, aby skutočná pozorovaná úroveň úmrtnosti bola v súlade s parametrami $\hat{\beta}_x^{(1)}$ a $\hat{\beta}_x^{(2)}$ získanými pre každý vek x . Iteratívnym vyhľadávaním nájdeme nový odhad parametra $\hat{\kappa}_t^{(2)}$, pričom pôvodný odhad upravíme tak, aby skutočný celkový pozorovaný počet úmrtí $\sum_x D_{x,t}$ sa rovnal celkovým očakávaným úmrtiam $\sum_x L_{x,t} e^{\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}}$ pre každý rok t .

Iteračná metóda v bode 4. prebieha nasledovne:

- 1) Porovnáme celkové očakávané úmrtia $\sum_x L_{x,t} e^{\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}}$ a skutočné pozorované úmrtia $\sum_x D_{x,t}$ v každom roku t .
- 2) Toto porovnanie odhalí jeden z troch možných stavov:
 - i. V prípade, že platí $\sum_x L_{x,t} e^{\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}} > \sum_x D_{x,t}$, potrebujeme znížiť očakávané úmrtia a upraviť odhadovaný parameter $\hat{\kappa}_t^{(2)}$ tak, že nový odhad $\hat{\kappa}_t'^{(2)}$ bude:
$$\hat{\kappa}_t'^{(2)} = \hat{\kappa}_t^{(2)} (1 - d), \text{ ak } \hat{\kappa}_t^{(2)} > 0$$

$$\hat{\kappa}_t'^{(2)} = \hat{\kappa}_t^{(2)} (1 + d), \text{ ak } \hat{\kappa}_t^{(2)} < 0$$

pričom d predstavuje nejaké malé číslo.
 - ii. Pokiaľ platí $\sum_x L_{x,t} e^{\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}} = \sum_x D_{x,t}$, tak zastavíme iteračný proces.

- iii. Napokon keď platí $\sum_x L_{x,t} e^{\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}} < \sum_x D_{x,t}$, potrebujeme zvýšiť očakávané úmrtia a upraviť odhadovaný parameter $\hat{\kappa}_t^{(2)}$ tak, že nový odhad $\hat{\kappa}_t'^{(2)}$ bude:

$$\hat{\kappa}_t'^{(2)} = \hat{\kappa}_t^{(2)}(1 + d), \text{ ak } \hat{\kappa}_t^{(2)} > 0$$

$$\hat{\kappa}_t'^{(2)} = \hat{\kappa}_t^{(2)}(1 - d), \text{ ak } \hat{\kappa}_t^{(2)} < 0$$

3) Vrátime sa ku kroku 1).

Výsledok iteračného postupu sa líši od pôvodných odhadov. Dôvodom je, že nízke miery úmrtnosti mladých ľudí majú oveľa nižší vplyv na celkový počet úmrtí, ale pri odhadovaní prirodzených logaritmov miery úmrtnosti majú rovnakú váhu ako vysoké miery úmrtnosti starších ľudí. Je tiež potrebné poznamenať, že rozdiely vo veľkostiach vekových skupín tiež vedú k rôznym váham pri prehodnocovaní parametra $\hat{\kappa}_t^{(2)}$.

Aplikácia metódy najmenších štvorcov na CBD model

Metóda najmenšieho štvorca sa používa na odhad parametrov minimalizáciou súčtu štvorcov rezíduí. Aplikácia tejto metódy pre model CBD je nasledovná. Minimalizačná funkcia má tvar:

$$J = \sum_{x=x_1}^{x_p} \sum_{t=t_1}^{t_q} \left(\log \left(\frac{q_{x,t}}{1 - q_{x,t}} \right) - \kappa_t^{(1)} - \kappa_t^{(2)}(x - \bar{x}) \right)^2 \quad (3.47)$$

Aby sme minimalizovali funkciu J , zderivujeme funkciu (3.47) podľa $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$ a derivácie položíme rovné nule, aby sme našli stacionárny bod funkcie. Dostaneme

$$\frac{\partial J}{\partial \kappa_t^{(1)}} = \sum_{x=x_1}^{x_p} 2 \left(\log \left(\frac{q_{x,t}}{1 - q_{x,t}} \right) - \kappa_t^{(1)} - \kappa_t^{(2)}(x - \bar{x}) \right) = 0 \quad (3.48)$$

$$\frac{\partial J}{\partial \kappa_t^{(2)}} = \sum_{x=x_1}^{x_p} 2(x - \bar{x}) \left(\log \left(\frac{q_{x,t}}{1 - q_{x,t}} \right) - \kappa_t^{(1)} - \kappa_t^{(2)}(x - \bar{x}) \right) = 0 \quad (3.49)$$

Riešením systému lineárnych rovníc získame najmenšie štvorcové odhady parametrov modelu CBD.

$$\hat{\kappa}_t^{(1)} = \frac{\sum_{x=x_1}^{x_p} \log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) - \kappa_t^{(2)} \sum_{x=x_1}^{x_p} (x - \bar{x})}{p} \quad (3.50)$$

$$\hat{\kappa}_t^{(2)} = \frac{p \sum_{x=x_1}^{x_p} (x - \bar{x}) \log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) - \sum_{x=x_1}^{x_p} \log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) \sum_{x=x_1}^{x_p} (x - \bar{x})}{p \sum_{x=x_1}^{x_p} (x - \bar{x})^2 - \left(\sum_{x=x_1}^{x_p} (x - \bar{x})\right)^2} \quad (3.51)$$

3.3.2 Prognózovanie – ARIMA modely

ARIMA modely patria do skupiny Box-Jenkinsových metód. Tieto metódy sú základom pre modelovanie rôznych druhov časových radov, od nestacionárnych cez sezónne až po stacionárne a nestacionárne viacrozmerne časové rady. Metodológia Box-Jenkins využíva reziduálnu zložku ako východiskovú zložku konštrukcie časových radov. Reziduálnu zložku môžu tvoriť aj korelované náhodné veličiny, z čoho vyplýva, že prostredníctvom tejto metodológie je možné pracovať s časovými radmi, ktoré majú navzájom závislé pozorovania. Metódy Box-Jenkinsa vytvárajú veľmi presné prognostické modely pre rôzne druhy údajov, pričom platí, že využívajú najnovšie pozorovania pri analyzovaní minulých chýb prognóz pre správne odhadnutie budúceho vývoja.

Výhodou Box-Jenkinsových modelov je ich flexibilita, pretože sa veľmi rýchlo prispôbujú zmenám v povahe procesu modelovania. Ďalšou veľkou výhodou je schopnosť modelovať trend a aj sezónnu zložku. Algoritmy sú však pomerne zložité a zahŕňajú niekoľko iteračných postupov s komplikovanými výpočtami. Okrem toho použitie metodológie vyžaduje vysokú mieru subjektivity, pri ktorej sú potrebné dlhodobejšie skúsenosti, pretože je nutné vedieť si vybrať zo širokej škály autoregresných modelov a modelov integrovaných kľzavých priemerov, ktorých je ARIMA súčasťou.

Prvkami Box-Jenkinsových modelov sú:

- **Autoregresné členy (AR)** – táto zložka vyjadruje vzťah modelu k predchádzajúcim hodnotám (histórii) časového radu a má za následok použitie posunutej hodnoty reziduí.
- **Členy kľzavých priemerov (MA)** – člen, ktorý slúži na zlepšovanie prognózy, pričom využíva hodnoty reziduí z predchádzajúcich predpovedí.
- **Integračný člen (I)** – prvok vyjadrujúci rád, ktorým sa časový rad diferencuje so zámerom eliminovania nestacionarity.

- **Sezónna zložka (S)** – v prípade, že časové rady sú v kratších intervaloch ako je jeden rok (štvrtročných, mesačných, týždenných alebo denných), odporúča sa použiť sezónne autoregresné členy (SAR) alebo sezónne členy kĺzavých priemerov (SMA).

Vo všeobecnosti môžeme postup tvorby Box-Jenkins modelov zapísať do štyroch krokov:

1. Identifikácia modelu – do identifikácie modelu patrí voľba vyhovujúceho modelu a určenie počtu parametrov a druhu parametrov. Voľba modelu vyplýva z grafickej analýzy autokorelačnej funkcie (AFC) a funkcie parciálnej autokorelácie (PAFC). Pridaním parametra do modelu rastie presnosť odhadu, avšak znižuje sa jeho zrozumiteľnosť. Podstatou je nájsť optimálnu rovnováhu medzi presnosťou a zrozumiteľnosťou (jednoduchosťou) modelu.

2. Odhad modelu – v ďalšom kroku odhadneme vyhovujúce modely a otestujeme významnosť ich parametrov. Metóda najmenších štvorcov je jedna z metód na odhad parametrov. Jej opis nájdeme v časti 3.3.1. Na posúdenie kvality modelu sa používajú napríklad Akaikovo informačné kritérium (AIC) a Swarz-Bayesovo kritérium (SBC), ktoré vyhodnocujú optimálnu rovnováhu medzi presnosťou a jednoduchosťou modelu. Vo všeobecnosti platí, že keď porovnávame vhodné modely, menšie hodnoty AIC a SBC signalizujú viac vyhovujúci model. Viac informácií o kritériách nájdeme v časti 3.4.

3. Verifikácia modelu – následne je dôležité overiť kvalitu modelu prostredníctvom rezíduí. Rezíduá by mali byť nezávislé od času, a teda mali by byť takmer náhodné. V prípade, že máme k dispozícii dlhé časové rady, môžeme posúdiť stabilitu modelu pomocou spätného testovania (backtestingu). V takom prípade odhadneme model na kratšom časovom intervale a posúdime kvalitu odhadu na zvyšnom časovom intervale.

4. Prognózovanie – význam ARIMA modelov spočíva v ich prognostických vlastnostiach. Kvalitu presnosti prognózy možno posúdiť na základe porovnania odhadovaných hodnôt so skutočnou hodnotou sledovanej premennej. Na posúdenie presnosti prognózy sa používajú najčastejšie MSE a RMSE, ktoré sú opísané v časti 3.4.

Proces ARIMA (p, d, q)

Proces ARIMA (p, d, q) sa používa na modelovanie nestacionárnych časových radov. Nestacionarita je typická pre ekonomické časové rady a má za následok prítomnosť trendu. Môže byť spôsobená zmenami strednej hodnoty procesu alebo rozptylu procesu v čase,

prípadne oboje. Aby sme model stacionarizovali, môžeme použiť diferenciu hodnôt d -teho stupňa. V praxi sa najčastejšie stretávame s diferenciou prvého a druhého stupňa.

Poradie modelu ARIMA (autoregresný integrovaný kĺzavý priemer) sa zvyčajne označuje notáciou $ARIMA(p, d, q)$, kde

- p je poradie autoregresnej časti,
- d je poradie diferenciácie,
- q je poradie procesu kĺzavého priemeru.

V prípade, že v modeli nie je žiadna diferenciácia hodnôt časového radu ($d = 0$), modely sa zvyčajne označujú ako modely $ARMA(p, q)$.

Matematicky vieme ARIMA model zapísať v tvare

$$W_t = \mu + \frac{\theta(B)}{\phi(B)} a_t \quad (3.52)$$

kde t predstavuje časový index, W_t je rad odozvy alebo rozdiel hodnôt radu odozvy, μ označuje strednú hodnotu, B je operátor posunu (tzn. $BX_t = X_{t-1}$), $\phi(B)$ znamená autoregresný operátor reprezentovaný ako polynóm operátora posunu B

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$\theta(B)$ signalizuje operátor kĺzavého priemeru, ktorý je rovnako vyjadrený ako polynóm operátora posunu B

$$\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$$

a napokon a_t je nezávislá porucha, ktorá sa označuje ako náhodná chyba. Pre jednoduchý (nesezónny) model s možnosťou diferencovania môžeme W_t vyjadriť takto

$$W_t = (1 - B)^d y_t \quad (3.53)$$

Potom ďalšia možnosť zápisu tohto modelu vyzerá nasledovne

$$\phi(B)(W_t - \mu) = \theta(B)a_t \quad (3.54)$$

alebo

$$\phi(B)W_t = \text{konštanta} + \theta(B)a_t \quad (3.55)$$

pričom $\text{konštanta} = \phi(B)\mu = \mu - \phi_1\mu - \dots - \phi_p\mu$.

Určenie stupňa diferenciencie:

Stupeň diferenciencie môžeme určiť viacerými spôsobmi. Najjednoduchším je posúdenie grafu časového radu. Ak má lineárny trend, tak na jeho stacionarizáciu nám postačí prvá diferenciencia. V prípade, že ide o kvadratický trend, vhodné je použiť druhú diferencienciu.

Druhým variantom je zhodnotenie tvaru výberovej autokorelačnej funkcie (VACF) a výberovej parciálnej autokorelačnej funkcie (VPACF). Ak je začiatočná hodnota VACF blízka jednej a ďalej hodnoty klesajú pomaly (takmer lineárne), hodnoty VPACF naopak klesajú exponenciálne k nule, je vhodné diferencovať časový rad, pretože je nestacionárny. Časové rady diferencujeme, pokiaľ sa nezmení tvar autokorelačných funkcií. Optimálny stupeň diferencovania je ten, pri ktorom je štandardná odchýlka diferencovaných časových radov najnižšia. Nepatrné prediferencovanie časového radu môžeme kompenzovať zaradením $MA(q)$ procesu a poddiferencovanie naopak vieme vyvážiť zaradením $AR(p)$ procesu do modelu.

Určenie stupňa autoregresného polynómu p a polynómu klzavého priemeru q :

Po vymedzení stupňa diferencovania je potrebné stanoviť stupeň autoregresného polynómu p , a tiež stupeň polynómu klzavého priemeru q . Najpoužívanější postup spočíva v hľadaní podobností medzi tvarmi funkcií VACF a VPACF s ich teoretickými tvarmi. Vo výberových ACF a PACF môže nastať situácia, že obe funkcie klesajú rýchlo k nule, a to znamená, že žiadna z nich nemá charakter exponenciálneho klesania. Vtedy treba spozorovať, ktorá z funkcií klesá k nule rýchlejšie, a teda má menší počet štatisticky významných koeficientov. Ak VACF klesá rýchlejšie, potom zvolíme model $MA(q)$. Naopak, ak VPACF klesá rýchlejšie, potom si vyberieme model $AR(p)$. Ak obe funkcie klesajú k nule rovnako rýchlo, potom odhadneme oba modely a na základe verifikácie oboch odhadnutých modelov rozhodneme, ktorý model je vhodnejší.

V $ARIMA(p, d, q)$ modeli môže dochádzať k tomu, že procesy $AR(p)$ a $MA(q)$ si budú vzájomne rušiť svoje účinky, hoci oba budú štatisticky významné, no v modeli budú fungovať proti sebe. Vtedy hovoríme, že model je preparametrizovaný. Prejaví sa to tak, že odhady parametrov budú nejednoznačné a model bude mať problém skonvergovať. Predísť tomuto problému môžeme tým, že odhadneme model s nižším stupňom AR alebo MA procesom a porovnáme ich. Z tohto dôvodu sa pri výbere parametrov modelov používa

metóda postupného zaraďovania jednotlivých procesov do modelu (forward selection method).

3.4 Meranie kvality modelov a presnosti ich prognóz

Priemerný súčet štvorcov rezíduí (MSE) je dôležitým štatistickým kritériom používaným na meranie kvality regresných modelov a posúdenie presnosti ich prognóz. Rezíduá predstavujú v regresnej analýze rozdiely medzi odhadovanou a skutočnou hodnotou závislej premennej. MSE potom môžeme označiť ako rizikovú funkciu, ktorá vyjadruje očakávanú hodnotu rozdielu štvorcovej chyby. Príčinami tohto rozdielu sú náhodnosť, alebo fakt, že odhady nezohľadňujú informácie, ktoré by mohli viesť k presnejším odhadom.

Označme $\hat{\mathbf{y}}_i$ vektor očakávaných hodnôt závislej premennej a $\mathbf{y}_i$ vektor skutočných hodnôt závislej premennej. Potom priemerný súčet štvorcov rezíduí môžeme zapísať v tvare

$$MSE = \frac{1}{n - k} \sum_{i=1}^n (\hat{\mathbf{y}}_i - \mathbf{y}_i)^2$$

kde k je počet parametrov. Všeobecný zápis pre priemerný súčet štvorcov rezíduí vyzerá takto

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] \quad (3.56)$$

pričom θ je konkrétny neznámy parameter a $\hat{\theta}$ zodpovedá odhadu tohto parametra. V prípade, že $MSE = 0$, potom odhadovaná hodnota sa rovná skutočnej hodnote parametra θ , čo je v praxi v podstate nemožný jav. MSE je nepriamo úmerný k presnosti modelu. Čím je nižšia hodnota MSE, tým je model presnejší. Môžeme ho považovať za vhodné kritérium aj na porovnávanie dvoch alebo viacerých štatistických modelov. MSE nachádza uplatnenie aj v metódach krokovej regresie, kde slúži k stanoveniu počtu premenných, ktoré majú byť zaradené do regresného modelu. Takisto je aj súčasťou štatistickej analýzy rozptylu.

Štandardná odchýlka rezíduí (RMSE) je ďalším zo spôsobov kvantifikácie rezíduí. Štandardnú odchýlku rezíduí definujeme ako druhú odmocninu rozdielov medzi odhadovanými hodnotami závislej premennej a jej pozorovanými skutočnými hodnotami. Podobne ako MSE sa RMSE používa na odhad veľkosti chyby modelov. Je to populárna miera presnosti na porovnanie kvality regresných modelov. Ak $\hat{\theta}$ je odhadom

vybraného parametra θ , potom RMSE v skutočnosti predstavuje druhú odmocninu priemerného súčtu štvorcov rezíduí MSE. Platí, že

$$\begin{aligned} RMSE(\hat{\theta}) &= \sqrt{MSE(\hat{\theta})} \\ RMSE(\hat{\theta}) &= \sqrt{E[(\hat{\theta} - \theta)^2]} \end{aligned} \quad (3.57)$$

Existujú ešte ďalšie miery na posúdenie kvality modelov a presnosti ich prognóz. Spomenieme napríklad ešte priemernú absolútnu chybu rezíduí (MAE), priemernú absolútnu percentuálnu chybu rezíduí (MAPE), relatívnu chybu nazývanú tiež Theilov koeficient a iné.

Alternatívny prístup k výberu modelu zahŕňa použitie pravdepodobnostných štatistických mier, ku ktorým patria napríklad Akaikovo a Bayesovské informačné kritérium. V našej práci budeme tento prístup používať pri výbere vhodného ARIMA modelu na prognózovanie vývoja mier úmrtnosti stochastických modelov. Informačné kritériá poskytujú analytickú techniku na hodnotenie modelov a výber medzi nimi. Modely sú hodnotené tak na základe ich výkonu na množine cvičných údajov, ako aj na základe zložitosti modelu. Výkonnosť modelu sa dá vyhodnotiť pomocou log-likelihood (vierohodnosti), ktorá vychádza z metódy maximálnej vierohodnosti. Zložitosť modelu možno vyhodnotiť podľa počtu stupňov voľnosti alebo parametrov v modeli.

Akaikovo informačné kritérium, skrátené AIC, je pomenované podľa objaviteľa *Hirotsugu Akaike*. Vychádza zo základov teórie informácií a dedukcií založených na frekvencii. Štatistika AIC je definovaná takto

$$AIC = \log \hat{\sigma}_k^2 + \frac{n + 2k}{n} \quad (3.58)$$

kde k je počet parametrov v modeli, n je rozsah súboru (počet sledovaní) a $\hat{\sigma}_k^2$ je reziduálny rozptyl. Ak chceme na výber modelu použiť AIC, jednoducho si vyberieme model, ktorý dáva najmenšiu AIC oproti množine uvažovaných modelov. V porovnaní s BIC (3.59) toto kritérium neberie do úvahy zložitosť modelu.

Bayesovské informačné kritérium alebo skrátené BIC, je kritérium odvodené z Bayesovskej teórie pravdepodobnosti, na základe čoho dostalo aj svoje pomenovanie. Štatistiku môžeme vyjadriť takto

$$BIC = \log \hat{\sigma}_k^2 + \frac{k \log(n)}{n} \quad (3.59)$$

Rovnako ako pri AIC aj pri tomto kritériu vyberáme model s najnižšou hodnotou štatistiky. Vypočítaná hodnota sa líši od AIC, avšak dá sa dokázať, že tieto kritériá sú navzájom proporcionálne. Na rozdiel od AIC, BIC berie do úvahy aj zložitosť modelu, pričom zložitejšie modely majú vypočítanú štatistiku horšiu (vyššiu), a tak je menej pravdepodobné, že takéto modely budú vybrané.

3.5 Niektoré základné druhy poistných produktov

Súčasťou cieľov dizertačnej práce je poukázať na rozdiely pri použití rôznych modelov úmrtnosti, a tiež vplyv koronavírusu a gender smernice na jednorazové netto poistné. Aby sme boli schopní naplniť tento cieľ, potrebujeme si ešte zadať základné poistné produkty. Najčastejšími druhmi poistenia osôb sú poistenie pre prípad smrti, poistenie pre prípad dožitia, zmiešané poistenie a dôchodkové poistenie. Pričom na trhu existujú rôzne varianty týchto druhov poistení. Pri analýzach v praktickej časti sa zameriame len na poistenie na úmrtie a dôchodkové poistenie, pretože tieto dva druhy produktov sú najviac ovplyvnené rizikom nesprávneho stanovenia predpokladu úmrtnosti poistného kmeňa.

Poistenie na úmrtie sa častokrát nazýva aj rizikové poistenie a je preň typická jednorazová výplata poistnej sumy. Poistenie na úmrtie môžeme uzatvoriť dočasné alebo doživotné. Už z názvu vyplýva, že rozdiel spočíva v dĺžke poistnej doby. My na svoju analýzu použijeme dočasné poistenie na úmrtie, ktoré sa zvykne označovať aj ako úverové, pretože býva často spojené s hypotekárnymi úvermi. Poisťovňa preberá zodpovednosť za splatenie úveru, v prípade, že by poistená osoba zomrela počas doby splácania tohto úveru (n). **Dočasné poistenie na úmrtie** má označenie $A_{x:\overline{n}|}^1$ a jeho súčasnú hodnotu vypočítame podľa vzťahu:

$$A_{x:\overline{n}|}^1 = \sum_{k=0}^{n-1} {}_k p_x \cdot q_{x+k} \cdot v^{k+1} \quad (3.60)$$

pričom p_x je ročná miera dožitia, q_x je miera úmrtnosti a v je odúročiteľ.

Súčasná hodnota poistenia je aktuálna cena očakávaných čiastok, ktoré bude musieť poisťovňa v rámci poistenia vyplatiť. Na odhad súčasnej hodnoty sa používajú úmrtnostné

tabuľky a poistno-technická úroková miera určená v čase uzavretia poistnej zmluvy. Súčasná hodnota je zároveň jednorazovým netto poistným, teda ide o poistné zaplatené naraz na celý čas platnosti poistenia, ktoré je splatné v okamihu uzavretia poistnej zmluvy a nezohľadňuje náklady poisťovne. Poistná sadzba je potom súčasná hodnota vyjadrená na jednotku poistnej hodnoty. Vzťah (3.60) teda vyjadruje jednorazové netto poistné, ktoré musí zaplatiť x - ročná osoba, aby v prípade, že zomrie v priebehu najbližších n rokov, bola jej dedičom vyplatená na konci roka úmrtia poistná suma vo výške 1 € (Šoltésová, 2019).

Dôchodkové poistenia sú naopak charakteristické priebežnými výplatami s určitou pravidelnosťou (napr. mesačne, ročne, ...) počas určitého obdobia, pokiaľ je poistená osoba nažive. To znamená, že nevieme, aké dlhé toto obdobie bude. Typickými príkladmi sú napríklad starobné dôchodky, sirotské dôchodky alebo v niektorých prípadoch aj invalidné dôchodky. Dôchodkové poistenie nemusí byť len doživotné, je možné si založiť aj dočasné dôchodkové poistenia, u ktorého je doba výplaty vopred dohodnutá v poistnej zmluve. Okrem tohto členenia môžeme dôchodkové poistenie rozlišovať ešte na predlehotné a polehotné. Pri predlehodnom dôchodku sa dávky vyplácajú na začiatku obdobia a pri polehotnom je to na konci obdobia. Na analýzu si ako zástupcu dôchodkových poistení vyberieme dočasný predlehotný dôchodok. **Dočasný predlehotný dôchodok** označujeme aktuárskym symbolom $\ddot{a}_{x:\overline{n}|}$ a pre jeho súčasnú hodnotu platí vzťah

$$\ddot{a}_{x:\overline{n}|} = \sum_{k=0}^{n-1} {}_k p_x \cdot v^k, \quad (3.61)$$

kde p_x je ročná miera dožitia a v je odúročiteľ. Pomocou vzťahu (3.61) vypočítame jednorazové netto poistné, ktoré musí zaplatiť x - ročná osoba v prípade, že chce, aby jej po dobu n rokov bola vyplácaná dávka vo výške 1 € na začiatku každého roka pokiaľ bude nažive (Šoltésová, 2019).

3.6 Softvér SAS Enterprise Guide

SAS (Statistical Analysis Software) je štatistický softvér vyvinutý SAS Inštitútom za účelom dátového manažmentu, pokročilých analýz, prediktívnych analýz, viacrozmerných analýz, business intelligence, atď. Vývoj tohto softvéru trval 10 rokov, od roku 1966 do roku 1976, kedy bol založený Inštitút SAS. SAS poskytuje aj grafické

používateľské rozhranie typu point-and-click pre netechnických používateľov a iné možnosti prostredníctvom jazyka SAS. Má viac ako 200 modulov s veľmi širokým rozsahom použitia a jednoduchým používaním, medzi ktoré patria napríklad tieto:

- Base SAS - Základné procedúry a správa údajov
- SAS/STAT - Štatistická analýza
- SAS/GRAF - Grafika a prezentácie
- SAS/ETS - Ekonometria a analýza časových radov
- SAS/QC - Kontrola kvality
- Enterprise Miner - Dolovanie dát
- Enterprise Guide - Grafické používateľské rozhranie
- SAS EBI - Business Intelligence.

Viac informácií o softvéri SAS a jeho moduloch nájdete na webovej stránke www.sas.com, ktorá okrem iného obsahuje informácie o produktoch, konferenciách organizovaných používateľskými skupinami, ale aj programy poskytované ich tvorcami.

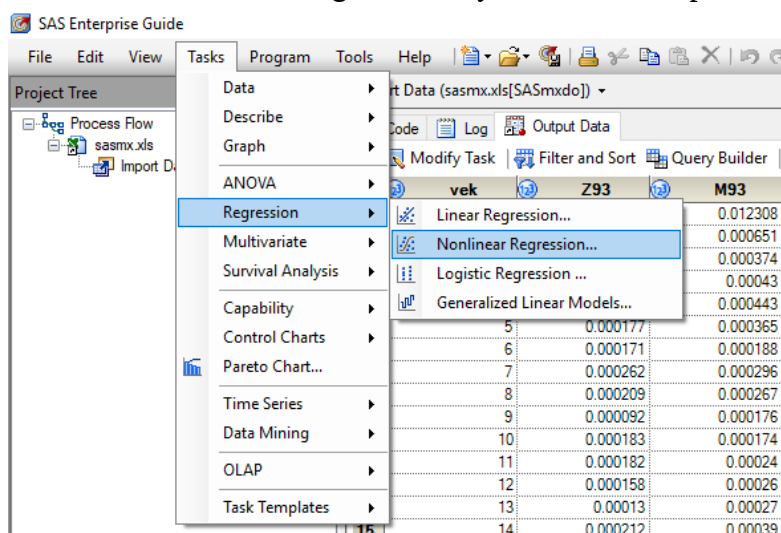
Jedným zo samostatných nástrojov je už spomínaný SAS Enterprise Guide. Ide o grafické používateľské rozhranie, ktoré nachádza uplatnenie v oblasti marketingu, telekomunikácii, poisťovníctva, ale aj v iných oblastiach, v ktorých sa pracuje s väčším množstvom údajov. Výhodou tohto prostredia je rozhranie point-and-click, ktoré zabezpečuje, že používateľ pri práci s dátami nepotrebuje nutne poznať programovací jazyk SAS, ale postačuje mu znalosť základných nástrojov. SAS EG umožňuje okrem iného:

- spracovanie a analýzu dát (štatistická, ekonometrická, matematická) – poskytuje rôzne úlohy, ktoré môžeme vykonávať na dátach (okrem iného analýza rozptylu, regresia, analýza časových radov, popisná štatistika),
- transformácia, triedenie, kombinovanie súborov údajov,
- grafická konštrukcia SQL dotazov (spolu s generovaním SQL) pomocou sprievodcu,
- vytváranie grafov a grafických reportov,
- vytváranie procesov a úloh vo forme projektu,
- zdieľanie výsledkov a analýz, export výsledkov do aplikácií Adobe Acrobat, Microsoft Excel, Microsoft Access, Microsoft Word, Microsoft PowerPoint, automatické odosielanie e-mailov so správami a výsledkami,
- prístup k štruktúram SAS OLAP cez SAS OLAP Server alebo iné štruktúry dodávateľov a vizualizácia OLAP dát vo forme diagramov.

SAS EG je bezpečné pracovné prostredie umožňujúce prístup viacerých používateľov k rovnakým údajom. Jednoduché používateľské rozhranie a kompatibilita s inými databázami zabezpečuje intuitívnu prácu s údajmi a komplexnú analýzu prostredníctvom iba jedného nástroja.

Nelineárna regresná analýza v SAS Enterprise Guide

Obrázok č. 3 - Nelineárna regresná analýza v SAS Enterprise Guide



Zdroj: vlastné spracovanie, softvér SAS Enterprise Guide

Súčasťou používateľského rozhrania SAS Enterprise Guide je aj nelineárna regresia. Na obrázku č. 3 môžeme vidieť jej zaradenie v softvérovom prostredí. Pri odhade parametrov regresného modelu je možné si vybrať jednu z troch metód opísaných v kapitole 3.2 Metódy odhadu parametrov deterministických modelov úmrtnosti (Gaussova-Newtonova metóda, Gradientná metóda, Levenbergova-Marquardtova metóda).

Procedúra nelineárnej regresie v SAS Enterprise Guide poskytuje širokú ponuku možností nastavenia výstupu. Za jej hlavnú nevýhodu považujeme malý výber predvolených tvarov nelineárnych modelov, ktoré nezodpovedajú vzťahom deterministických modelov úmrtnosti (zákonom úmrtnosti) opísaným v časti 3.1, ktoré sú predmetom tejto záverečnej práce. To však nie je problém, pretože v prostredí SAS EG sa dajú upravovať ale aj písať vlastné programy pomocou programovacieho jazyka SAS. Ukážku programu nelineárnej regresnej analýzy nájdeme v prílohe 1.

Na naprogramovanie vlastného kódu nelineárnej regresnej analýzy sa predpokladá znalosť štyroch procedúr PROC SQL, PROC NLIN, prípadne PROC SQRT a PROC

GLOT. Procedúra PROC GLOT slúži na tvorbu grafických výstupov. V prípade nelineárnej regresie ide konkrétne o grafy predikovaných hodnôt, rezíduí, a iné. PROC SORT nachádza využitie pri usporiadaní pozorovaných údajov podľa hodnôt jednej alebo viacerých premenných. Tieto procedúry nie sú bezpodmienečne nevyhnutné pre vytvorenie programu nelineárnej regresie, ale napomáhajú k hlbšej analýze. PROC SQL patrí k východiskovým procedúram, ktoré obsahuje každý program a používa sa na vytváranie, udržiavanie, získavanie a aktualizáciu údajov. Najpodstatnejšou procedúrou vytvorenou priamo pre nelineárnu regresnú analýzu je PROC NLIN. Obsahuje množstvo príkazov, ktorých prehľad nájdeme v (SAS Institute Inc, 2015). Za najdôležitejšie príkazy považujeme tieto

- **METHOD** (= *MARQUARDT*, *NEWTON*, *GRADIENT*) – slúži na špecifikáciu iteračnej metódy, ktorá má byť použitá na odhad parametrov nelineárneho regresného modelu.
- **MAXITER** (= *n*) – pomocou tohto príkazu definujeme maximálny počet iterácií optimalizačného algoritmu. Predvolená hodnota je $n = 100$.
- **CONVERGE** (= *c*) – používa sa na stanovenie konvergenčného kritéria pre procedúru. Prednastavená hodnota konvergenčného kritéria je 10^{-5} .
- **MODEL** – pomocou príkazu vymedzíme požadovaný tvar regresného modelu.
- **PARMS** – dáva priestor pre bližšiu špecifikáciu parametrov a nastavenie ich počiatočných hodnôt. Počiatočné hodnoty sú prednastavené na 0.001.

Obrázok č. 4 - Fázy iteračného postupu
Nonlinear Regression

Results				
The NLIN Procedure				
Dependent Variable 2020				
Method: Marquardt				
Iterative Phase				
Iter	a	b	c	Sum of Squares
0	0.000190	0.000018	1.1135	1.486E-6
1	0.000187	0.000014	1.1192	1.358E-6
2	0.000203	0.000012	1.1232	1.29E-6
3	0.000214	0.000011	1.1258	1.27E-6
4	0.000224	9.856E-6	1.1283	1.265E-6
5	0.000227	9.64E-6	1.1290	1.262E-6
6	0.000227	9.6E-6	1.1291	1.262E-6
7	0.000227	9.594E-6	1.1291	1.262E-6
8	0.000227	9.593E-6	1.1291	1.262E-6

NOTE: Convergence criterion met.

Zdroj: výstup zo softvéru SAS EG

V tejto časti predstavíme ukážku výstupu procedúry nelineárnej regresie pre Gompertzov-Makehamov model úmrtnosti v roku 2020 bez zohľadnenia pohlavia. Súčasťou výstupu výslednej procedúry nelineárnej regresnej analýzy sú jednotlivé fázy iteračnej metódy (obrázok č. 4), ktoré obsahujú aj súčet štvorcov rezíduí. Ďalej vo výstupe nájdeme sumár procesu odhadovania premenných (obrázok č. 5), kde je definovaná použitá metóda, počet iterácií na dosiahnutie konvergenčného kritéria, počet pozorovaní a iné charakteristiky.

Obrázok č. 5 - Sumár procesu odhadu parametrov

Estimation Summary	
Method	Marquardt
Iterations	8
Subiterations	8
Average Subiterations	1
R	7.754E-6
PPC(b)	0.000014
RPC(b)	0.000093
Object	1.673E-9
Objective	1.262E-6
Observations Read	50
Observations Used	50
Observations Missing	0

Zdroj: výstup zo softvéru SAS EG

Výstup zahŕňa aj analýzu najmenších štvorcov (obrázok č. 6), v ktorej sa nachádzajú charakteristiky pre model a reziduálnu zložku ako počet stupňov voľnosti, súčet štvorcov, priemerný súčet štvorcov. Pomocou nich sa navyše dopracujeme k F štatistike a jej p -hodnote, na základe ktorých sme schopní otestovať štatistickú významnosť nelineárneho regresného modelu.

Obrázok č. 6 - Analýza najmenších štvorcov

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	2	0.000054	0.000027	1002.43	<.0001
Error	47	1.262E-6	2.685E-8		
Corrected Total	49	0.000055			

Zdroj: výstup zo softvéru SAS EG

Dôležitým výstupom sú aj odhady parametrov spolu s ich štandardnými odchýlkami a s intervalmi spoľahlivosti odhadu parametrov (obrázok č. 7). Okrem toho výstup obsahuje

aj korelačnú maticu parametrov (obrázok č. 8) a nakoniec graf pozorovaných hodnôt miery úmrtnosti v závislosti od veku (obrázok č. 9).

Obrázok č. 7 - Odhad parametrov

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Skewness
a	0.000227	0.000039	0.000150	0.000305	-0.0750
b	9.593E-6	2.825E-6	3.909E-6	0.000015	0.7694
c	1.1291	0.00686	1.1153	1.1429	0.1294

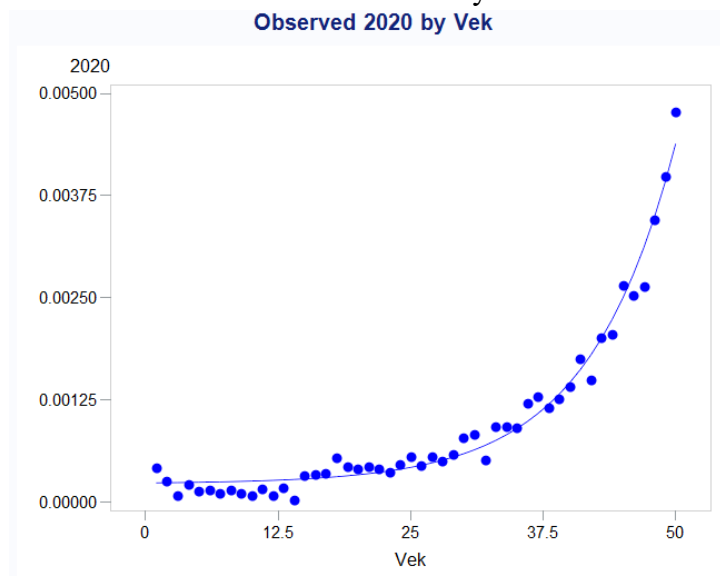
Zdroj: výstup zo softvéru SAS EG

Obrázok č. 8 - Korelačná matica

Approximate Correlation Matrix			
	a	b	c
a	1.0000000	-0.1391414	-0.1261535
b	-0.1391414	1.0000000	0.1266709
c	-0.1261535	0.1266709	1.0000000

Zdroj: výstup zo softvéru SAS EG

Obrázok č. 9 - Graf závislosti miery úmrtnosti od veku



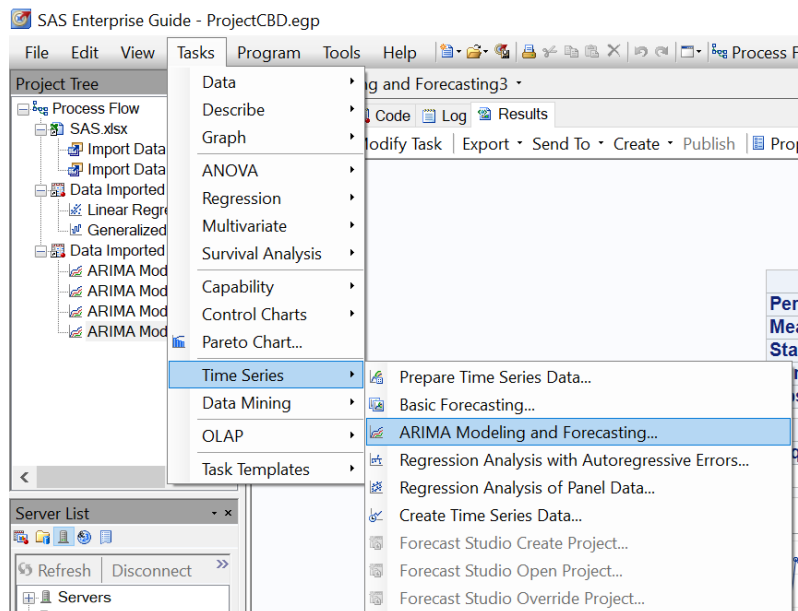
Zdroj: výstup zo softvéru SAS EG

ARIMA

Procedúra ARIMA poskytuje komplexnú sadu nástrojov na jednorozmernú identifikáciu modelu časových radov, odhad parametrov a prognózu a zároveň ponúka veľkú flexibilitu v druhoch modelov ARIMA, ktoré je možné analyzovať. V časti 3.3.2 sme uviedli postup tvorby ARIMA modelu podľa Box-Jenkinsovej metódy, pričom táto metóda sa

využíva aj v programe SAS Enterprise Guide. Na obrázku č. 10 vidíme, že ARIMA modely sú v softvérovom prostredí začlenené medzi Časové rady.

Obrázok č. 10 - ARIMA modelovanie v SAS Enterprise Guide



Zdroj: vlastné spracovanie, softvér SAS Enterprise Guide

Pre naprogramovanie kódu ARIMA modelov sa predpokladá znalosť práce s procedúrami PROC SORT, PROC GPLOT, PROC SQL a PROC ARIMA. Prvé tri procedúry sú známe už z nelineárnej regresie v prvej časti tejto kapitoly. PROC ARIMA je procedúra tvorby ARIMA modelov, ktorá sa skladá z troch fáz – **IDENTIFY**, **ESTIMATE**, **FORECAST**. Ukážku programu ARIMA modelovania nájdeme v prílohe 1.

Vo fáze identifikácie sa aplikuje príkaz **IDENTIFY** na zisťovanie možných vyhovujúcich modelov ARIMA. Príkaz IDENTIFY číta časové rady, ktoré sa majú použiť v ďalších príkazoch, s možnosťou diferencovania a počíta autokorelácie, inverzné autokorelácie, čiastočné autokorelácie a krížové korelácie. V tomto kroku je možné vykonať testy stacionarity, aby sa zistilo, či je potrebné diferencovanie časového radu. Analýza výstupu príkazu IDENTIFY zvyčajne vedie k jednému alebo viacerým modelom ARIMA, ktoré by mohli byť vhodné.

Vo fáze odhadu a diagnostickej kontroly používame príkaz **ESTIMATE** na určenie modelu ARIMA, ktorý by zodpovedal premenným špecifikovaným na základe formulácie IDENTIFY a na odhad parametrov tohto modelu. Príkaz ESTIMATE tiež vytvára

diagnostické štatistiky, ktoré nám pomáhajú posúdiť primeranosť modelu a na základe nich porovnávame presnosť modelov navzájom.

Testy významnosti pre odhady parametrov signalizujú, či niektoré premenné v modeli nie sú zbytočné. Testy bieleho šumu rezíduí zasa naznačujú, či rezíduá neobsahujú ďalšie informácie, ktoré by mohli viesť k zložitejšiemu modelu. Špecifikácia **OUTLIER** poskytuje ďalší užitočný nástroj na kontrolu toho, či súčasný odhadovaný model zohľadňuje všetky odchýlky v časovom rade. Ak diagnostické testy poukazujú na problém s modelom, je potrebné vyskúšať iný model a následne zopakovať túto fázu.

Vo fáze prognózovania sa používa príkaz **FORECAST** na predpovedanie budúcich hodnôt časových radov a na generovanie intervalov spoľahlivosti týchto predpovedí z modelu ARIMA vytvoreného v predchádzajúcom kroku príkazom ESTIMATE.

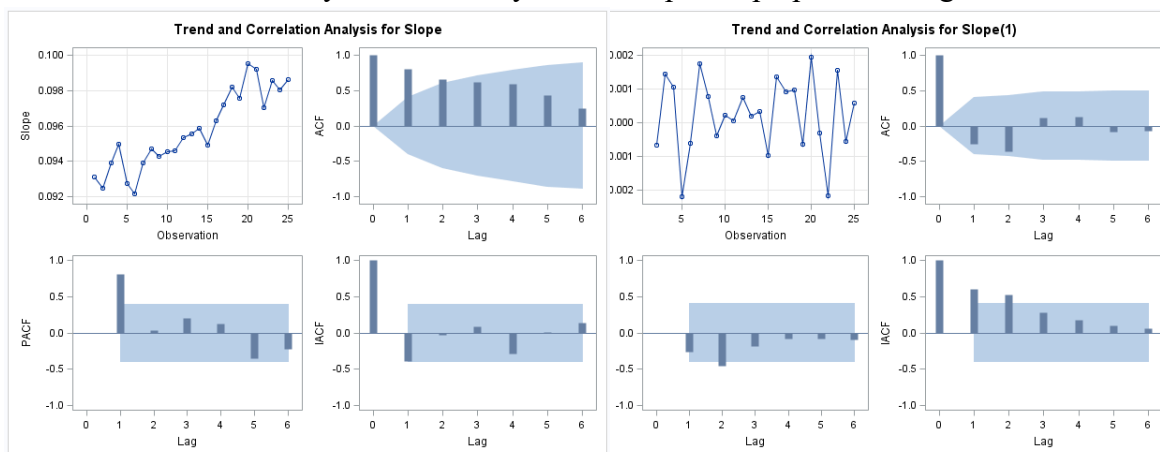
Obrázok č. 11 - Opisné štatistiky

Name of Variable = Slope2	
Period(s) of Differencing	1
Mean of Working Series	0.000229
Standard Deviation	0.001096
Number of Observations	24
Observation(s) eliminated by differencing	1

Zdroj: výstup zo softvéru SAS EG

Potom ako sme sa venovali základným príkazom pri tvorbe ARIMA modelov, predvedieme aj ukážku výstupu procedúry pre parameter $\kappa_t^{(2)}$ CBD modelu úmrtnosti bez zohľadnenia pohlavia. Príkaz IDENTIFY najskôr zobrazí hodnoty opisných štatistík (obrázok č. 11), pozostávajúcich zo stupňa diferenciácie, počtu pozorovaní, strednej hodnoty a štandardnej odchýlky časového radu. Ďalej tento príkaz vytvorí panel grafov na posúdenie autokorelácie a analýzu trendov, pričom príkaz **NLAG** riadi počet oneskorení, pre ktoré sú zobrazené autokorelácie. V predvolenom nastavení sú funkcie autokorelácie vykreslené na oneskorenie 24. Panel obsahuje graf časového radu, graf výberovej autokorelačnej funkcie (VACF), graf inverznej výberovej autokorelačnej funkcie (VIACF) a graf parciálnej výberovej autokorelačnej funkcie (VPACF).

Obrázok č. 12 - Grafy autokorelačných funkcií pred a po pridaní integračného člena



Zdroj: výstup zo softvéru SAS EG

V prvom rade potrebujeme z nestacionárneho časového radu spraviť stacionárny, k čomu nám pomôže integračný člen (obrázok č. 12). Na identifikovanie stupňa diferenciácie nám môže pomôcť tento sumár pravidiel (Robert Nau, 2020):

Pravidlo 1 - Ak má časový rad pozitívnu autokoreláciu s vysokým počtom oneskorení (povedzme 10 alebo viac), pravdepodobne potrebuje vyšší stupeň diferenciácie.

Pravidlo 2 - Ak je autokorelácia s oneskorením -1 nulová alebo negatívna alebo sú všetky autokorelácie malé, potom časový rad nepotrebuje vyššie poradie diferenciácie. Ak je táto autokorelácia negatívna s hodnotou -0,5 alebo viac, rad môže byť prediferencovaný.

Pravidlo 3 - Optimálny stupeň diferenciácie je často ten, pri ktorom je štandardná odchýlka diferencovaných časových radov najnižšia. Mierne prediferencovanie alebo poddiferencovanie je možné napraviť pomocou členov AR alebo MA.

Pravidlo 4 - Nediferencovaný model predpokladá, že pôvodný časový rad je stacionárny (okrem iného). Model, ktorý je raz diferencovaný, predpokladá, že pôvodný rad má konštantný trend (napr. model náhodnej prechádzky). Model dvakrát diferencovaný predpokladá, že pôvodný rad má trend meniaci sa v čase (napr. náhodný trend).

Po vymedzení stupňa diferencovania je potrebné stanoviť stupeň autoregresného polynómu p , a tiež stupeň polynómu klzavého priemeru q . Najpoužívanejší postup spočíva v hľadaní podobností medzi tvarmi autokorelačných funkcií ACF a PACF s ich teoretickými tvarmi. Na identifikovanie stupňov AR a MA členov existuje tiež niekoľko pravidiel, ktoré nám pomôžu pri ich stanovení (Robert Nau, 2020):

Pravidlo 5 - Ak funkcia parciálnej výberovej autokorelácie (PACF) diferencovaného časového radu zobrazuje ostrý skok a/alebo je autokorelácia s oneskorením -1 pozitívna, teda ak sa rad javí mierne poddiferencovaný, odporúča sa zvážiť prídanie jedného alebo viacerých členov AR do modelu. Oneskorenie, po prekročení ktorého PACF ostro skočí, indikuje počet výrazov AR.

Pravidlo 6 - Ak výberová autokorelačná funkcia (ACF) diferencovaného radu znázorňuje ostrý skok a/alebo autokorelácia s oneskorením -1 je negatívna, teda ak sa rad javí mierne prediferencovaný, navrhuje sa pridať do modelu člen MA. Oneskorenie, po prekročení ktorého ACF ostro skočí, vyjadruje počet výrazov MA.

Pravidlo 7 - Členy AR a MA môžu navzájom rušiť svoje efekty, takže ak sa zdá, že model s oboma členmi vyhovuje údajom, je potrebné vyskúšať aj model s jedným menším počtom členov AR a jedným menším počtom členov MA, najmä ak odhady parametrov v pôvodnom modeli si vyžadujú na skonvergovanie viac ako 10 iterácií.

Pravidlo 8 - Ak je v časti modelu AR jednotkový koreň, to znamená, ak je súčet koeficientov AR takmer presne 1, mali by sme znížiť počet členov AR o jeden a zvýšiť stupeň diferenciácie o jeden.

Pravidlo 9 - Ak je v časti modelu MA jednotkový koreň, teda ak je súčet koeficientov MA takmer presne 1, mali by sme znížiť počet členov MA o jeden a znížiť tiež stupeň diferenciácie o jeden.

Pravidlo 10 - Ak sa dlhodobé prognózy javia ako nepravidelné alebo nestabilné, v koeficientoch AR alebo MA môžu mať jednotkový koreň.

Obrázok č. 13 - Test bieleho šumu

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	62.05	6	<.0001	0.803	0.658	0.613	0.596	0.428	0.243

Zdroj: výstup zo softvéru SAS EG

Posledná časť predvoleného výstupu príkazu IDENTIFY je kontrola bieleho šumu. Toto je aproximatívny test štatistickej hypotézy, že žiadna z autokorelácií časového radu až do daného oneskorenia sa významne nelíši od 0. V prípade, že je táto hypotéza pravdivá, ARIMA model nie je potrebný pre tento časový rad. V tomto prípade hypotézu bieleho šumu zamietame, keďže p -hodnota testu je menšia ako 0,0001 (obrázok č. 13).

Obrázok č. 14 - Odhad parametrov ARIMA modelu

Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MU	0.0002492	0.00003816	6.53	<.0001	0
MA1,1	0.85277	0.11388	7.49	<.0001	1

Zdroj: výstup zo softvéru SAS EG

Tabuľka odhadov parametrov na obrázku č. 14 má názov „Conditional Least Squares Estimation“, ktorý označuje použitú metódu odhadu. S príkazom **METHOD** môžeme nastaviť aj iné metódy odhadu. V SAS Enterprise Guide máme na výber metódu maximálnej vierohodnosti, podmienenú metódu najmenších štvorcov a nepodmienenú metódu najmenších štvorcov. Tabuľka odhadov parametrov obsahuje zoznam parametrov v modeli. Pre každý parameter tam nájdeme jeho odhadovanú hodnotu, štandardnú odchýlku a *t* štatistiku pre odhad, na základe ktorej vieme posúdiť štatistickú významnosť parametra.

Obrázok č. 15 - Štatistiky goodness of fit

Constant Estimate	0.000249
Variance Estimate	9.814E-7
Std Error Estimate	0.000991
AIC	-262.002
SBC	-259.646
Number of Residuals	24

Zdroj: výstup zo softvéru SAS EG

Ďalšou časťou výstupu príkazu ESTIMATE je tabuľka štatistík goodness-of-fit na obrázku č. 15, ktoré pomáhajú pri porovnaní tohto modelu s inými modelmi.

Obrázok č. 16 - Korelačná matica

Correlations of Parameter Estimates		
Parameter	MU	MA1,1
MU	1.000	-0.206
MA1,1	-0.206	1.000

Zdroj: výstup zo softvéru SAS EG

Okrem týchto výstupov tiež obsahuje tabuľku korelácií odhadov parametrov, ako je znázornené na obrázku č. 16. Táto tabuľka slúži na posúdenie, do akej miery môže

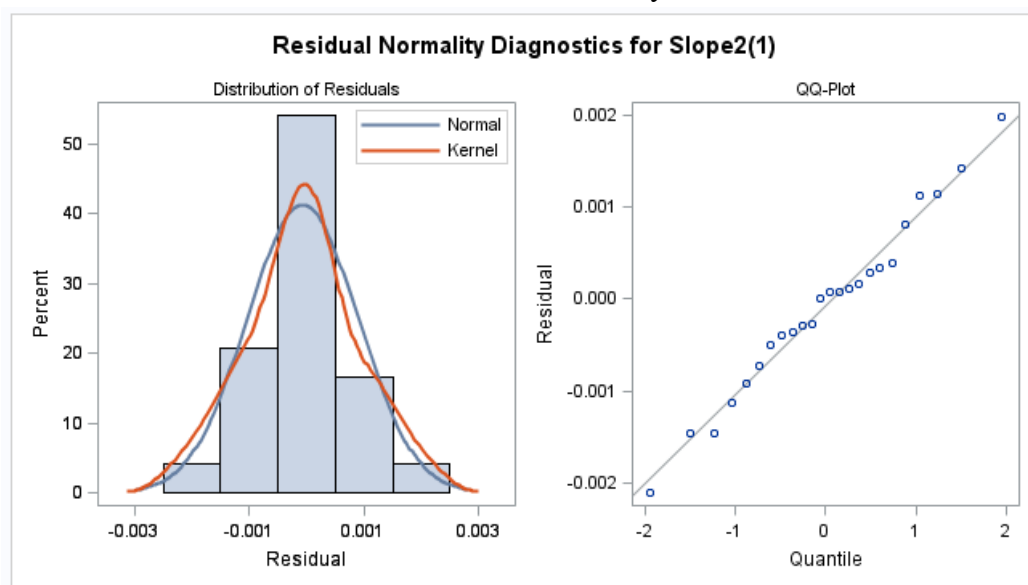
kolinearita ovplyvniť výsledky. Ak sú dva odhady parametrov veľmi vysoko korelované, je potrebné zvážiť vylúčenie jedného z nich z modelu.

Obrázok č. 17 - Testy autokorelácie rezíduí

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	2.94	5	0.7085	0.228	-0.139	0.064	0.065	-0.093	-0.114
12	7.06	11	0.7940	-0.049	-0.050	0.096	0.011	-0.185	-0.191
18	15.47	17	0.5618	-0.007	-0.150	-0.197	0.110	0.179	-0.071

Zdroj: výstup zo softvéru SAS EG

Obrázok č. 18 - Grafická analýza rezíduí



Zdroj: výstup zo softvéru SAS EG

Ďalšou súčasťou výstupu je kontrola autokorelácií rezíduí (obrázok č. 17), ktorá je dôležitá na posúdenie vhodnosti odhadnutého modelu. Tento výstup má rovnakú formu ako kontrola autokorelácie bieleho šumu. Chí-kvadrát štatistika testu slúži na posúdenie toho, či rezíduá sú korelované alebo či neobsahujú ďalšie informácie, ktoré by mohli viesť k zložitejším modelom. V tomto prípade zamietame autokoreláciu rezíduí v ARIMA modeli na vysokej hladine významnosti, čo znamená, že daný model je adekvátnym modelom pre tento časový rad. Výstup obsahuje aj grafickú analýzu rezíduí (obrázok č. 18), ktorá je najjednoduchším prostriedkom na posúdenie normality náhodnej zložky. Q-Q graf, skratka pre kvantil-kvantil graf, sa používa na posúdenie, či súbor údajov potenciálne pochádza z nejakého teoretického rozdelenia. V našom prípade je teoretickým rozdelením práve

normálne rozdelenie. Ak sú rezíduá normálne rozdelené, body v Q-Q grafe budú ležať na priamej diagonálnej čiare. Naopak, čím viac sa body v grafe výrazne odchyľujú od priamej diagonálnej čiary, tým je menej pravdepodobné, že rezíduá sa riadia normálnym rozdelením. Odchýlky kvantilov od osi 45° sú zanedbateľné, teda môžeme konštatovať, že reziduálna zložka má blízko k normálnemu rozdeleniu. Na základe posúdenia štatistickej významnosti parametrov, Akaikovho informačného kritéria (AIC), korelačnej matice, ale aj analýzy rezíduí vyberieme model, ktorý je najvhodnejší pre daný štatistický rad. Matematické vyjadrenie výsledného modelu zobrazené na obrázku č. 19 je takisto súčasťou výstupu ARIMA procedúry.

Obrázok č. 19 - Matematické vyjadrenie modelu

Model for variable Slope2	
Estimated Mean	0.000249
Period(s) of Differencing	1

Moving Average Factors	
Factor 1:	1 - 0.85277 B**(1)

Zdroj: výstup zo softvéru SAS EG

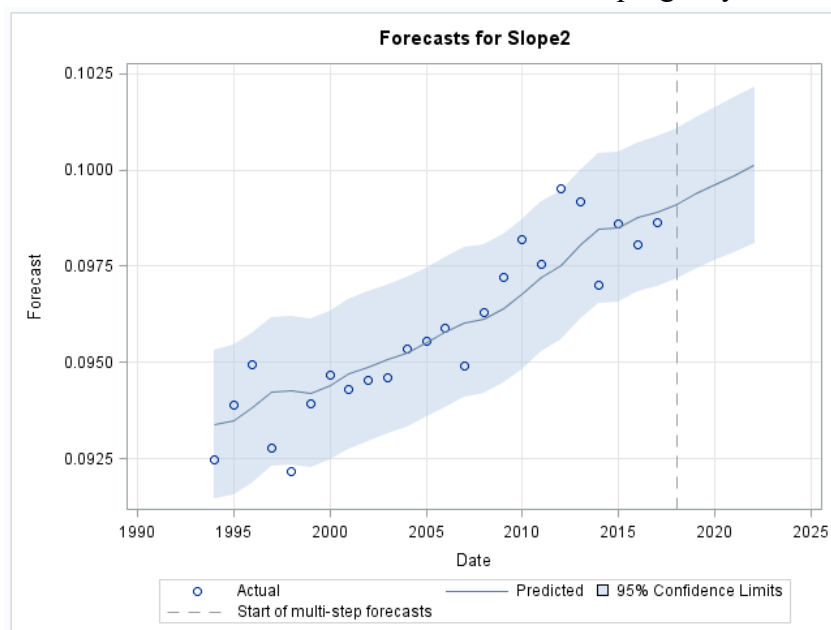
Na vytvorenie prognózy použijeme príkaz **FORECAST** po výbere vo fáze ESTIMATE modelu, ktorý považujeme za najlepší. Možnosť **LEAD** určuje, na koľko období má byť prognóza. Voľba **ID** predstavuje premennú, ktorá je zvyčajne v tvare dátumu, času. Predvolený príkaz FORECAST zobrazuje prognózované hodnoty, ako je to znázornené v tabuľke na obrázku č. 20. Tabuľka prognózy obsahuje pre každé prognózované obdobie počet pozorovaní, hodnotu prognózy, štandardnú odchýlku a 95 % interval spoľahlivosti. Výstup obsahuje aj grafické zobrazenie prognózovaných hodnôt (obrázok č. 21).

Obrázok č. 20 - Prognóza

Forecasts for variable Slope2				
Obs	Forecast	Std Error	95% Confidence Limits	
26	0.099123	0.0009907	0.097182	0.101065
27	0.099372	0.001001	0.097410	0.101335
28	0.099622	0.001012	0.097638	0.101605
29	0.099871	0.001022	0.097867	0.101875
30	0.100120	0.001033	0.098096	0.102144

Zdroj: výstup zo softvéru SAS EG

Obrázok č. 21 - Grafické znázornenie prognózy



Zdroj: výstup zo softvéru SAS EG

Poznámka:

Jednou z hlavných nevýhod ARIMA modelov, ako sme už spomenuli v časti 3.3.2, je, že metodológia si vyžaduje značnú mieru znalostí a dlhodobé skúsenosti. Naviesť na správny smer nám môže pomôcť funkcia `auto.arima()`, ktorá je súčasťou **forecast** knižnice v softvéri R. Funkcia `auto.arima()` používa modifikáciu Hyndman-Khandakarovho algoritmu (Hyndman a Khandakar, 2008), ktorý kombinuje test jednotkového koreňa, minimalizáciu AIC a metódu maximálnej vierohodnosti na získanie modelu ARIMA. 35 možných argumentov pre `auto.arima()` poskytuje veľa variácií tohto algoritmu. Viac informácií o argumentoch je možné nájsť v RDocumentation [29].

Algoritmus Hyndman-Khandakara pre automatické modelovanie ARIMA je nasledovný:

1. Diferencia d ($0 \leq d \leq 2$) sa určí pomocou opakovaných Kwiatkowski–Phillips–Schmidt–Shin (KPSS) testov. Tento test sa používa na zistenie, či je časový rad stacionárny okolo strednej hodnoty alebo lineárneho trendu, alebo je nestacionárny v dôsledku jednotkového koreňa. Stacionárny časový rad je taký, ktorého priemer a rozptyl sú v čase konštantné.

2. Hodnoty p a q potom získame pomocou minimalizácie AIC po diferencovaní údajov. Namiesto toho, aby sa zvažovali všetky možné kombinácie p a q , algoritmus používa postupné vyhľadávanie/zaradovanie.

a. Najskôr sa odhadnú štyri počiatočné modely:

- $ARIMA(0, d, 0)$
- $ARIMA(2, d, 2)$
- $ARIMA(1, d, 0)$
- $ARIMA(0, d, 1)$,

pričom každý model obsahuje konštantu pokiaľ sa d nerovná 2. Ak $d \leq 1$, odhaduje sa ešte aj dodatočný model $ARIMA(0, d, 0)$ bez konštanty.

b. Najlepší model (s najmenšou hodnotou AIC) odhadnutý v kroku a. sa nastaví ako „aktuálny model“.

c. Ďalej sa zvažia variácie súčasného modelu:

- pozmenenie p a/alebo q aktuálneho modelu o ± 1 ;
- zahrnutie/vylúčenie konštanty z aktuálneho modelu.

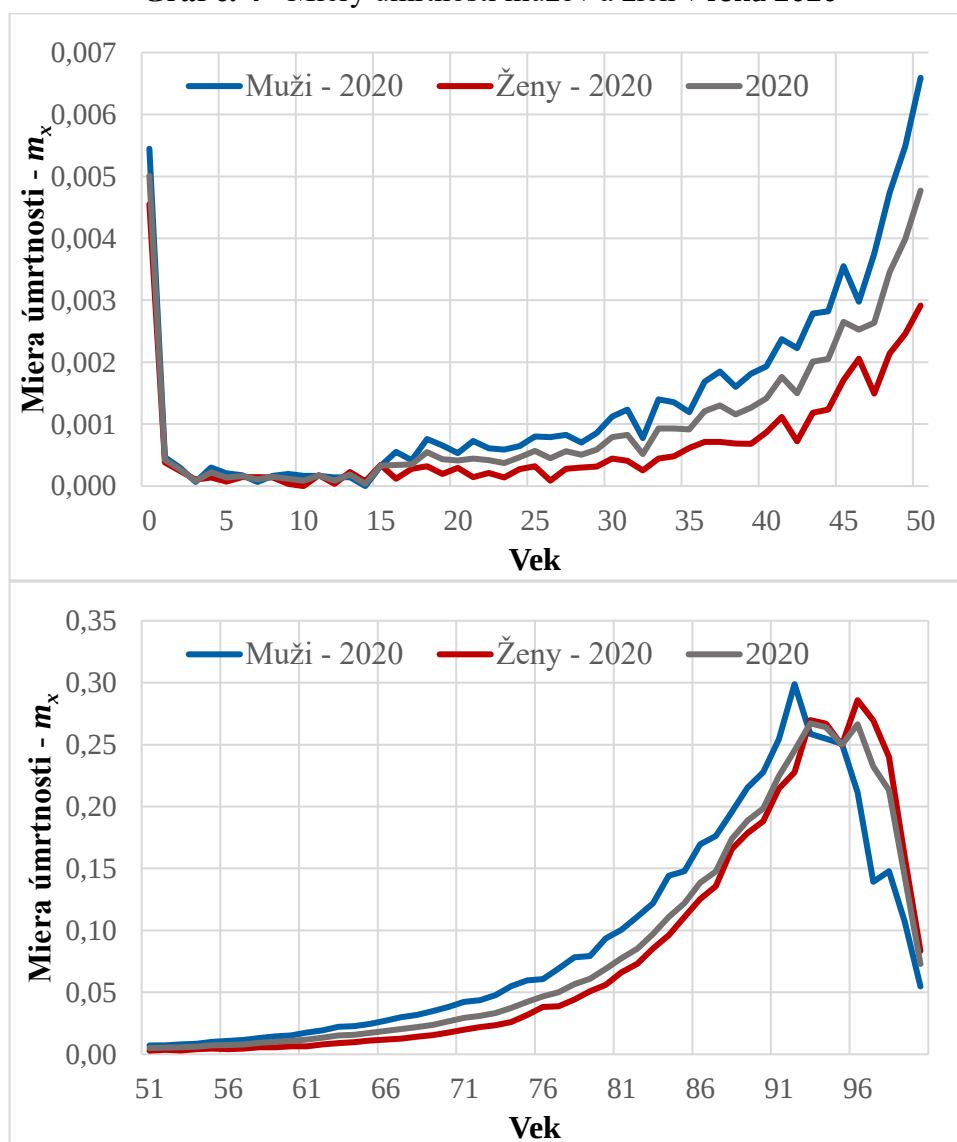
Najlepší model, o ktorom sa doteraz uvažovalo (buď aktuálny model alebo jedna z týchto variácií), sa stáva novým aktuálnym modelom.

Predvolený postup používa niektoré aproximované hodnoty na urýchlenie vyhľadávania. Týmto aproximáciám sa dá vyhnúť argumentom *approximation = FALSE*. Je možné, že model s minimálnym AIC nebude nájdený v dôsledku týchto aproximácií alebo kvôli použitiu postupného zaradovania. Ak sa použije argument *stepwise = FALSE*, bude sa hľadať najlepší model z oveľa väčšej množiny modelov. Ukážku algoritmu v softvéri R spolu s výstupmi je možné nájsť v prílohe 2.

4 Výsledky práce a diskusia

Princíp modelovania miery úmrtnosti, intenzity úmrtnosti, ale aj pravdepodobnosti úmrtia spočíva v tom, že spojité (okamžité) teoretické miery sú odhadnuté pomocou mier vypočítaných z reálnych dát pre ročný vekový interval, čo vedie k nesystematickým nepravidelnostiam. Graf č. 4 obsahuje vekovo špecifické miery úmrtnosti vypočítané z údajov o úmrtnosti mužov a žien na Slovensku v roku 2020 so zohľadnením pohlavia a aj jeho vynechaním.

Graf č. 4 - Miera úmrtnosti mužov a žien v roku 2020



Zdroj: vlastné spracovanie

Graf č. 4 poukazuje na niekoľko dôležitých faktov, ktoré treba brať do úvahy pri modelovaní úmrtnosti. Jedným z nich sú už spomínané nesystematické nepravidelnosti. Na grafe sú viditeľné najviac pre vek nad 90 rokov, kde klesajúci trend miery úmrtnosti je nereálny a nedáva zmysel. Môžeme to vysvetliť tak, že takéhoto veku sa nedožíva veľa ľudí, a teda dochádza k malému počtu úmrtí, pričom sú v nich aj veľké výkyvy. Tento jav si môžeme všimnúť tiež v mladších vekoch, kde to ale vzhľadom na menšie hodnoty výkyvov nie je až tak výrazné. Na odstránenie výkyvov, a teda priblíženie sa k spojitým hodnotám mier úmrtnosti sa používa graduácia (vyrovnávanie). Jednou z metód graduácie sú aj zákony úmrtnosti, ktorým sa v tejto dizertačnej práci venujeme.

Ďalšou dôležitou skutočnosťou zobrazenou na grafe č. 4 je rozdiel vo vekovo špecifických mierach úmrtnosti v závislosti od pohlavia. Ak neberieme do úvahy výkyvy od veku 90 rokov, môžeme povedať, že na celom intervale je miera úmrtnosti mužov vyššia ako miera úmrtnosti žien, čo nás privádza k otázke, či je spravodlivé zjednotenie pohlavia pri analýze úmrtnosti, ktorému sme sa venovali v kapitole 1.3. Odlišnosti vo vekovo špecifických mierach úmrtnosti v závislosti od pohlavia vedú k rozdielnym výškam poistného hlavne pri dôchodkovom poistení a poistení na dožitie. Zanedbanie pohlavia vedie k podhodnoteniu úmrtnosti mužskej populácie a nadhodnoteniu tej ženskej, výsledkom čoho dospejeme ku krivke, ktorá je akýmsi váženým priemerom úmrtí mužov a žien v danom veku pre daný rok. Vynechanie pohlavia pri modelovaní úmrtnosti je prijateľné do 18 roku života, pretože rozdiely v mierach úmrtnosti mužov a žien sú po tento vek zanedbateľné. V druhej polovici intervalu, medzi vekmi 51 až 100, si môžeme všimnúť, že krivka, ktorá nezohľadňuje pohlavie, sa viac približuje ku krivke reprezentujúcej úmrtnosť žien. To môže mať za následok značné podhodnotenie úmrtnosti mužskej populácie pre tento vekový interval.

V nasledujúcej kapitole dizertačnej práce sa budeme venovať jednotlivým deterministickým a stochastickým modelom úmrtnosti spomínaným v predchádzajúcej časti a aplikujeme ich na konkrétne údaje o úmrtnosti na Slovensku v rokoch 1993 až 2020 pri zohľadnení aj zanedbaní pohlavia ako faktora modelovania úmrtnosti. Modelovaným vekovým intervalom bude vekové rozpätie od 0 do 100 rokov. Aby sme zaručili prehľadnosť a vyššiu presnosť odhadu, vekový interval sme rozdelili na dve časti, oblasť do 50 rokov vrátane a oblasť nad 50 rokov. Ďalej sa zameriame na vzájomne porovnanie modelov úmrtnosti s vynechaním pohlavia ako faktora modelovania. Vyberieme model,

ktorý najvierohodnejšie opisuje údaje o úmrtnostnom správaní populácie Slovenskej republiky na celom vekovom intervale, pričom vhodnosť modelov posúdime na základe štatistických kritérií – priemerný súčet štvorcov rezíduí a štandardná odchýlka rezíduí opísaných v kapitole 3.4. V závere kapitoly analyzujeme, aký dopad má výber konkrétneho modelu na poistné sadzby, ale aj vplyv zavedenia rodovej rovnosti a efekt zhoršenej úmrtnosti v dôsledku koronavírusu na konci sledovaného obdobia na deterministický a stochastický model. Pri analýze vychádzame zo základných poistných produktov, ktorými sú dočasné poistenie na úmrtie a dočasný predľahotný dôchodok.

4.1 Deterministické modely úmrtnosti

4.1.1 Gompertzov – Makehamov model úmrtnosti

Modelovanie úmrtnostného správania populácie Slovenskej republiky sme začali jedným z prvých existujúcich deterministických modelov úmrtnosti – Gompertzovým-Makehamovým zákonom úmrtnosti. Môžeme ho vyjadriť pomocou vzťahu (3.3), pričom graficky je znázornený vo forme rastúcej exponenciálnej krivky. Vzhľadom k tvrdeniam samotných autorov modelu, očakávame, že najlepšie výsledky bude dosahovať na vekovom intervale od 20 do 60 rokov. Na odhad parametrov sme použili metódy nelineárnej regresnej analýzy opísané v časti 3.2. Keďže ponuku nelineárnych modelov softvéru SAS Enterprise Guide tvorí len niekoľko predvolených typov, ktoré nie sú v súlade s tvarmi zákonov úmrtnosti, pri vytváraní úmrtnostných modelov bolo nevyhnutné naprogramovanie vlastného kódu na odhad parametrov týchto modelov. Príklad výsledného programu nájdeme v prílohe 1.

Tabuľka č. 4 obsahuje súhrn výsledkov odhadovaných parametrov Gompertzovho-Makehamovho zákona úmrtnosti zvlášť, ale aj spoločne pre mužov a ženy za roky 1993 a 2020 spolu so štatistickými mierami na posúdenie vhodnosti modelov MSE a RMSE. Výsledky sú uvedené osobitne pre vekový interval $\langle 0; 50 \rangle$ a $\langle 50; 100 \rangle$ na dosiahnutie presnejších odhadov so zameraním na prehľadnosť výstupov.

Tabuľka č. 4 - Odhad parametrov GM modelu pre roky 1993 a 2020

Gompertz- Makeham model		1993			2020		
		Ženy	Muži	Spolu	Ženy	Muži	Spolu
Parametre 50-	A	0,00023	0,00037	0,00029	0,00015	0,0003	0,00023
	B	1,20E-05	4,2E-05	3,1E-05	3,47E-06	1,6E-05	9,6E-06
	c	1,1205	1,1182	1,1156	1,1419	1,1245	1,1291
Parametre 50+	A	-0,0182	-0,0374	-0,0316	-0,007	-0,0176	-0,1105
	B	5,45E-04	0,00501	0,00261	3,60E-05	0,00058	0,0224
	c	1,0671	1,0429	1,0494	1,0999	1,0686	1,028
MSE		0,0004	0,00083	0,00018	0,00277	0,00332	0,00079
RMSE		0,01998	0,02881	0,01331	0,05266	0,0576	0,02819

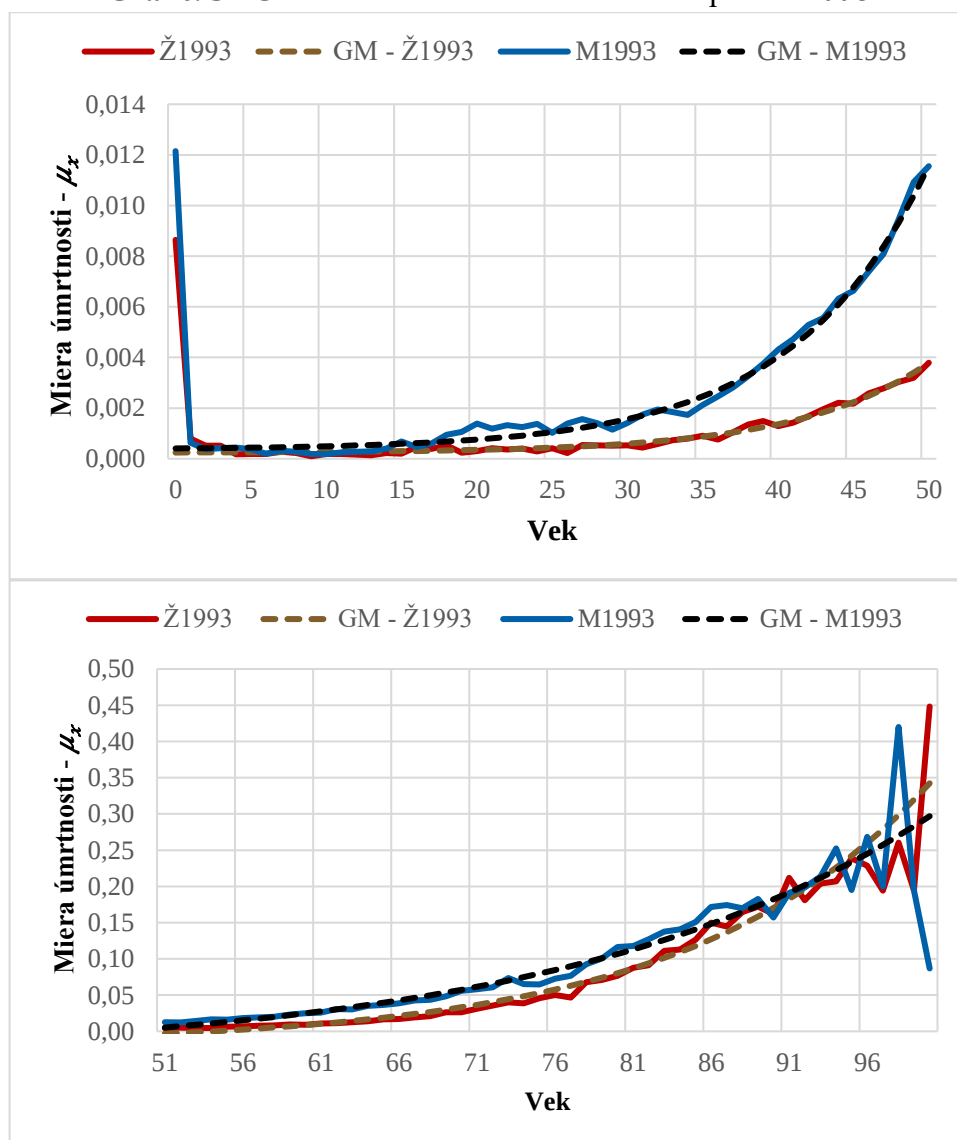
Zdroj: vlastné spracovanie

Výsledky modelovania pomocou Gompertzovho-Makehamovho zákona úmrtnosti v závislosti od pohlavia pre rok 1993 sú znázornené na grafe č. 5. Súčasťou grafu sú skutočné pozorované vekovo špecifické miery úmrtnosti mužov (M1993) a žien (Ž1993) vypočítané na základe údajov zo Štatistického úradu Slovenskej republiky a zároveň ich vyrovnané hodnoty za použitia Gompertzovho-Makehamovho zákona (GM – M1993, GM – Ž1993). V prvej kapitole sme sa venovali anomálii dojčenskej úmrtnosti, ktorá je výrazne vyššia ako v nasledujúcich rokoch života. Podľa grafu č. 5, ale už aj na základe tvaru Gompertzovho-Makehamovho zákona úmrtnosti (rastúca exponenciálna krivka) vieme povedať, že GM model úmrtnosti nespĺňa požiadavky na modelovanie dojčenskej úmrtnosti, pretože skutočné pozorované vekovo špecifické miery úmrtnosti sa výrazne odlišujú od vyrovnaných mier.

Ak sa zameriame na výsledky so zohľadnením pohlavia ako faktora modelovania, v grafe č. 5 pozorujeme, že v prípade mužskej populácie v roku 1993, model zľahka nadhodnocuje mieru úmrtnosti mužov po začiatok dospelosti, teda do veku 18 roku života. Následne je viditeľné mierne podhodnotenie úmrtnosti v oblasti havarijného hrbu. Medzi vekmi 18 až 90 rokov Gompertzov-Makehamov zákon úmrtnosti takmer absolútne opisuje úmrtnostné správanie mužskej časti populácie. Vo vysokých vekoch nad 90 rokov sa na grafe č. 5 objavujú výrazné nepravidelnosti, ktoré spôsobujú, že pri vyrovnaných mierach úmrtnosti mužov dochádza k podhodnoteniu skutočných mier úmrtnosti.

Čo sa týka opačného pohlavia, u žien platí, že volatilita údajov je značne miernejšia v porovnaní s mužmi a takisto v tomto roku nie je viditeľná na krivke zvýšená úmrtnosť v oblasti havarijného hrbu, čo spôsobilo zlepšenie kvality Gompertzovho-Makehamovho modelu. Môžeme povedať, že prakticky na celom vekovom intervale, tento zákon úmrtnosti spoľahlivo opisuje úmrtnostné správanie ženskej časti populácie. O tom, že Gompertzov-Makehamov model viac zodpovedá úmrtnosti ženskej populácie v roku 1993, svedčia aj hodnoty priemerného súčtu rezíduí a reziduálnej odchýlky.

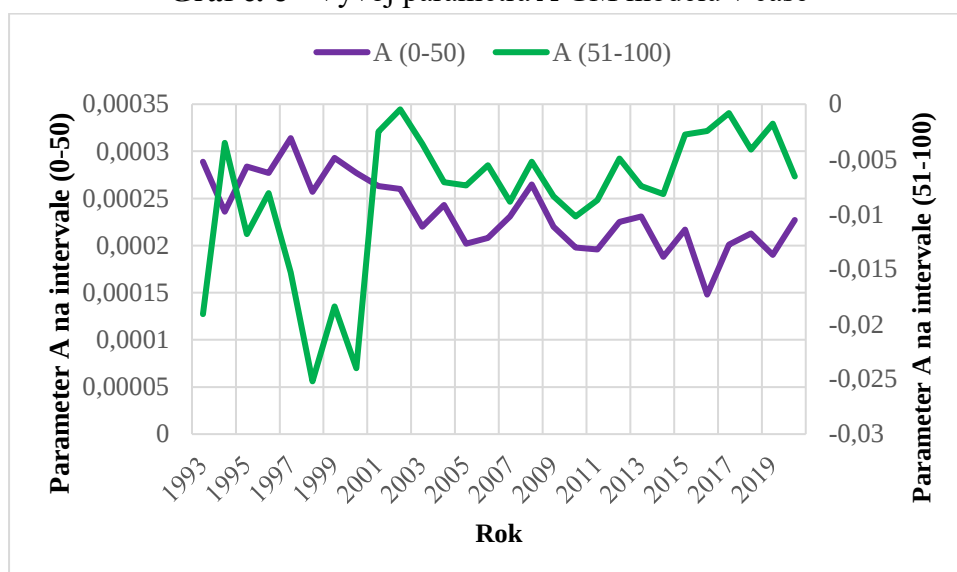
Graf č. 5 - GM model úmrtnosti mužov a žien pre rok 1993



Zdroj: vlastné spracovanie

Ak sa pozrieme do tabuľky č. 4 odhliadnuc od pohlavia, zistíme, že Gompertzov-
 Makehamov model vykazuje ešte lepšie výsledky s $MSE = 0,00018$ ako v prípade, ak
 uvažujeme oddelene úmrtnosť mužov a žien. Teraz sa ďalej zameriame na vývoj parametrov
 Gompertzovho-Makehamovho zákona úmrtnosti v čase pre pozorované obdobie od roku
 1993 do 2020 na vekových intervaloch $\langle 0; 50 \rangle$ a $\langle 50; 100 \rangle$, ktoré sú znázornené na grafoch
 č. 6, 7 a 8. Keďže hodnoty parametrov na týchto dvoch intervaloch sú pomerne rozdielne,
 na ľavej strane sa nachádza os hodnoty parametra pre vekový interval od 0 do 50 rokov
 vrátane a na pravej strane grafu nájdeme os pre vekový interval nad 50 rokov.

Graf č. 6 - Vývoj parametra A GM modelu v čase

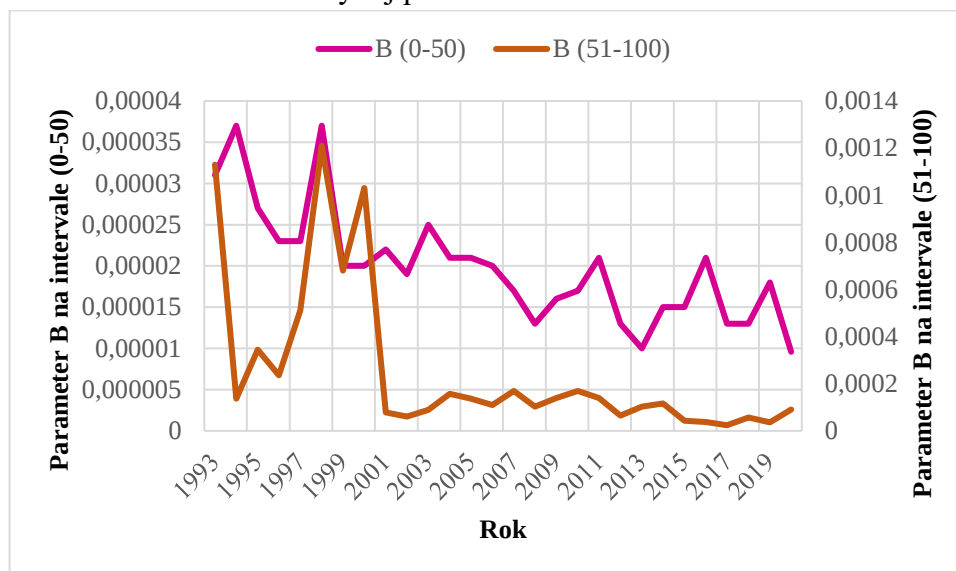


Zdroj: vlastné spracovanie

Makehamov parameter A predstavuje konštantu, ktorá slúži na posun grafu funkcie
 miery úmrtnosti. Na základe grafu č. 6 vieme povedať, že na vekovom intervale od 0 do 50
 rokov nadobúda tento parameter veľmi nízke hodnoty a vo všeobecnosti má klesajúcu
 tendenciu. Na intervale od 51 do 100 rokov sa parameter správa úplne odlišne. Dosahuje
 záporné hodnoty, ktoré sú okrem iného veľmi volatilné. Predpokladáme, že je to spôsobené
 volatilitou mier úmrtnosti vo vysokých vekoch, vďaka ktorej sa nepodarilo spoľahlivo
 odhadnúť výsledné parametre modelu. Parameter B Gompertzovho-Makehamovho zákona
 úmrtnosti má vplyv na zakrivenie (strmosť) krivky. Je logické preto, že na intervale
 $\langle 50; 100 \rangle$ sú hodnoty tohto parametra vyššie ako do veku 50 rokov (graf č. 7). Rovnako ako
 parameter A, má klesajúci trend. Parameter c takisto pôsobí na tvar krivky miery úmrtnosti,

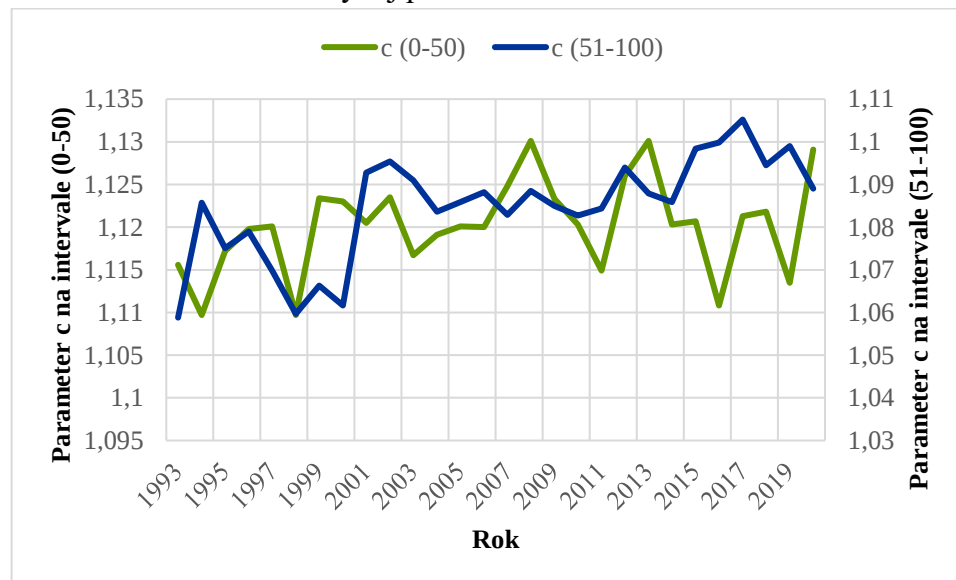
pričom vyjadruje funkčný vzťah medzi vekom a mierou úmrtnosti. Ako jediný z parametrov má porovnateľné hodnoty na celom vekovom intervale, pričom mierne rastie v čase (graf č. 8).

Graf č. 7 - Vývoj parametra B GM modelu v čase



Zdroj: vlastné spracovanie

Graf č. 8 - Vývoj parametra c GM modelu v čase

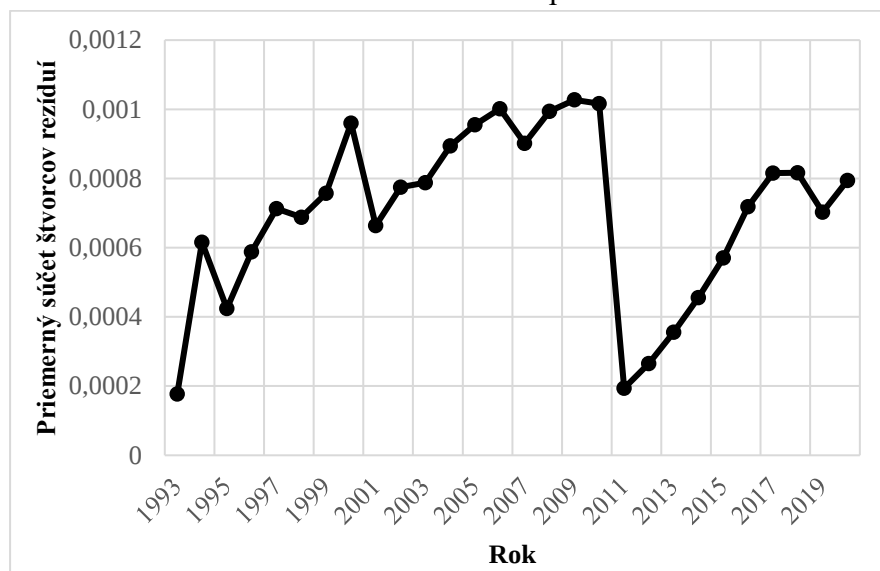


Zdroj: vlastné spracovanie

V kapitole sa budeme zaoberať aj vzájomným porovnaním deterministických a stochastických modelov úmrtnosti, načo nám poslúži priemerný súčet štvorcov rezíduí

a jeho odmocnina (štandardná odchýlka rezíduí). Graf č. 9 zachytáva MSE v jednotlivých rokoch modelovaného obdobia. Najspoľahlivejšie sa nám podarilo vyrovnať krivku skutočnej vekovo špecifickej miery úmrtnosti pomocou Gompertzovho-Makehamovho zákona úmrtnosti v rokoch 1993 a 2011. Naopak najvyššie hodnoty priemerného súčtu štvorcov rezíduí sú namerané v rokoch 2009 a 2010.

Graf č. 9 - MSE GM modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020



Zdroj: vlastné spracovanie

Aby sme zosumarizovali uplatnenie Gompertzovho-Makehamovho zákona úmrtnosti, môžeme povedať, že jeho pozorovaním pri modelovaní v rokoch 1993 – 2020 sme dospeli k záveru, že je vhodný hlavne pre vekové rozpätie od 25 do 90 rokov života. V nižších vekoch zľahka nadhodnocuje mieru úmrtnosti populácie, a naopak, vo vyšších vekoch ju podhodnocuje. V ďalších častiach preveríme fakt, či novšie modely úmrtnosti lepšie odzrkadľujú úmrtnostné správanie slovenskej populácie.

4.1.2 Kannistov – Makehamov model úmrtnosti

Reprezentantom logistických modelov je Kannistov-Makehamov zákon úmrtnosti. Vzťah na vyjadrenie tohto zákona (3.5) obsahuje tri parametre, ktoré bolo potrebné odhadnúť pomocou metód nelineárnej regresie, teda patrí stále k tým jednoduchším modelom. Vyznačuje sa podobnými charakteristickými črtami ako Gompertzov-

Makehamov zákon úmrtnosti. Keďže cieľom autorov pri tvorbe modelu bolo zohľadniť spomalenie úmrtnosti v pokročilom veku, očakávali by sme, že bude dosahovať lepšie výsledky práve vo vyšších vekoch. Pri vzájomnom porovnaní jednotlivých modelov v ďalších častiach práce sa však presvedčíme o podobnosti týchto modelov, ktoré sú takmer totožné, ak neberieme do úvahy pohlavie ako faktor, ktorý má vplyv na úmrtnosť.

V tabuľke č. 5 nájdeme výsledné odhadované parametre a miery presnosti Kannistovho-Makehamovho zákona úmrtnosti zvlášť, ale aj spoločne pre mužov a ženy za roky 1993 a 2020 na vekovom intervale do 50 rokov a na intervale (50;100).

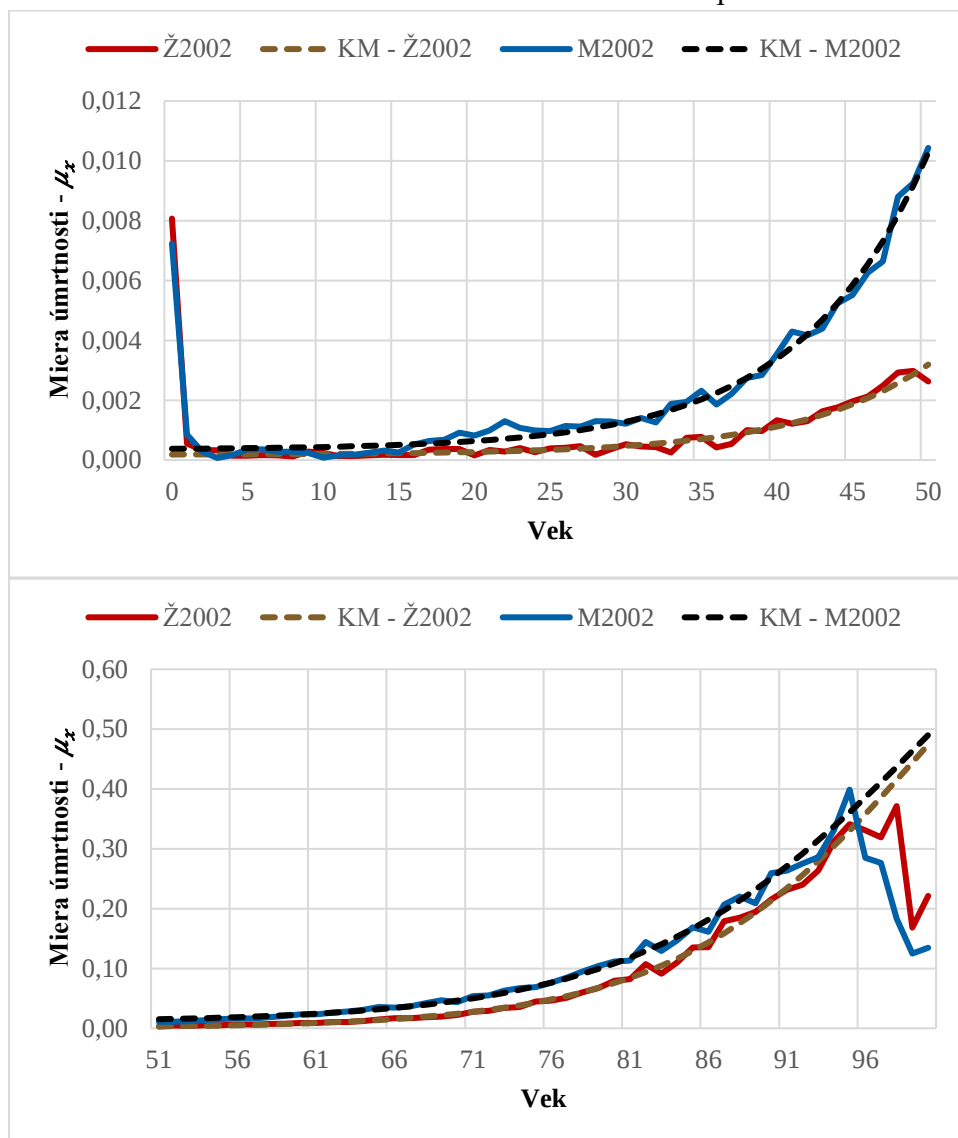
Tabuľka č. 5 - Odhad parametrov KM modelu pre roky 1993 a 2020

<i>Kannisto-Makeham model</i>		1993			2020		
		Ženy	Muži	Spolu	Ženy	Muži	Spolu
<i>Parametre 50-</i>	A	1,2E-05	4,1E-05	0,00003	3,45E-06	1,6E-05	9,5E-06
	B	1,14E-01	0,1125	0,1098	1,33E-01	0,1177	0,1217
	C	0,00023	0,00037	0,00029	0,00015	0,0003	0,00023
<i>Parametre 50+</i>	A	0,00016	0,00174	0,00071	5,4E-06	0,00011	0,0055
	B	8,06E-02	0,0556	0,0644	1,18E-01	0,0874	0,0441
	C	-0,011	-0,0218	-0,0171	-0,002	-0,0064	-0,0655
MSE		0,00038	0,0008	0,00017	0,00197	0,00291	0,00077
RMSE		0,01956	0,02823	0,01307	0,04443	0,05394	0,02771

Zdroj: vlastné spracovanie

Pri grafickej analýze sa bližšie pozrieme na rok 2002. Na grafe č. 10 sú znázornené výsledky modelovania pomocou Kannistovho-Makehamovho zákona úmrtnosti v závislosti od pohlavia. Modrou plnou čiarou sú zachytené pozorované vekovo špecifické miery úmrtnosti mužov (M2002) a miery úmrtnosti žien (Ž2002) sú reprezentované červenou plnou čiarou. Hodnoty týchto mier úmrtnosti vyrovnané pomocou Kannistovho-Makehamovho zákona úmrtnosti predstavujú prerušované čiary s označením KM – M2002, KM – Ž2002. Takisto ako pri predchádzajúcom Gompertzovom-Makehamovom modeli, Kannistov-Makehamov zákon úmrtnosti nie je vhodný na modelovanie dojčenskej úmrtnosti, čo patrí k jeho hlavným nevýhodám.

Graf č. 10 - KM model úmrtnosti mužov a žien pre rok 2002

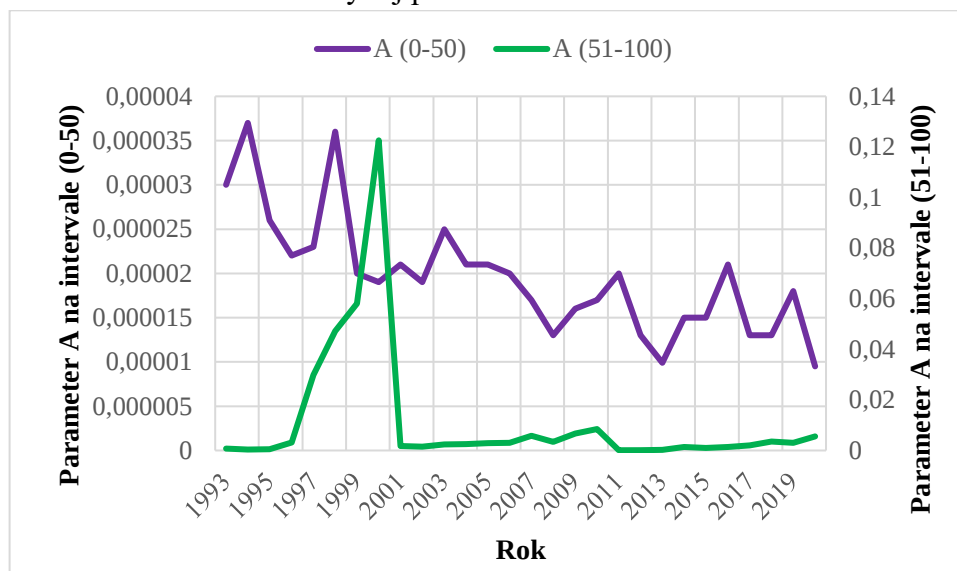


Zdroj: vlastné spracovanie

Na grafe č. 10 môžeme pozorovať, že vekovo špecifické miery úmrtnosti mužov a žien sa do 15 roku života značne približujú. Rovnaký jav môžeme vidieť aj u populácie nad 90 rokov, kedy sa dokonca krivky skutočných mier úmrtnosti pretáli vo veku 96 rokov a došlo k javu nazývanému ženská nadúmrtnosť. Tento jav v takto vysokých vekoch je spôsobený najmä tým, že ženy sa dožívajú vyššieho veku, čo znamená, že počet žien v tomto veku je vyšší, a teda aj zomiera viac žien ako mužov. Pravdepodobnosť úmrtia v starobe by mala byť u mužov väčšia, čo korešponduje s mierami úmrtnosti vyrovnanými pomocou Kannistovho-Makehamovho zákona úmrtnosti. Môžeme povedať, že prakticky na celom vekovom intervale, tento zákon úmrtnosti veľmi dobre opisuje úmrtnostné správanie

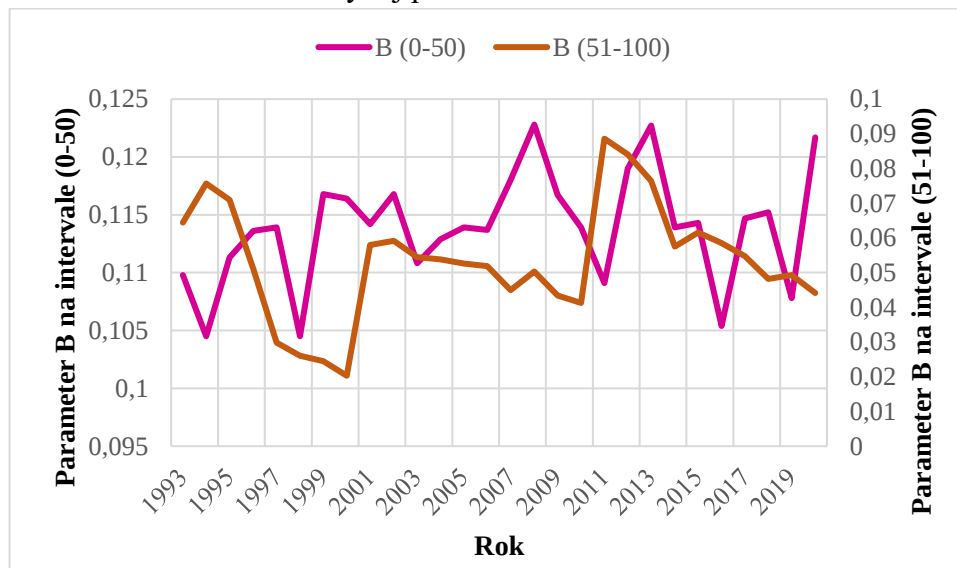
populácie mužov aj žien s miernym podhodnotením úmrtnosti v oblasti havarijného hrbu v prípade mužskej časti populácie. Okrem toho je úmrtnosť výrazne nadhodnotená pri oboch pohlaviach vo veku od 96-teho roku života.

Graf č. 11 - Vývoj parametra A KM modelu v čase



Zdroj: vlastné spracovanie

Graf č. 12 - Vývoj parametra B KM modelu v čase

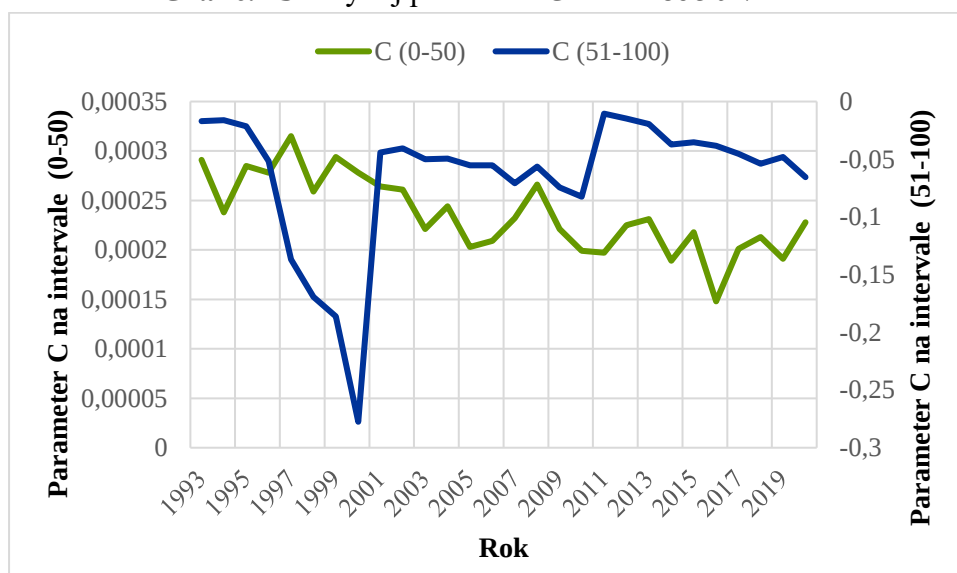


Zdroj: vlastné spracovanie

Vývoj parametrov Kannistovho-Makehamovho zákona úmrtnosti v čase pre pozorované obdobie od roku 1993 do 2020 na vekových intervaloch $\langle 0; 50 \rangle$ a $\langle 50; 100 \rangle$

je zobrazený na grafoch č. 11, 12 a 13. Hodnoty parametrov na týchto vekových intervaloch sú výrazne odlišné, preto sa na grafoch nachádzajú znova dve osy. Parametre A a B Kannistovho-Makehamovho modelu vplyvajú na tvar krivky. Hodnoty týchto parametrov sú veľmi volatilné, pričom parameter A má na intervale $\langle 0; 50 \rangle$ klesajúci trend. Medzi rokmi 1997 až 2000 sa Kannistovmu-Makehamovmu zákonu úmrtnosti nepodarilo spoľahlivo priblížiť tvaru krivky vekovo špecifickej miery úmrtnosti pre vek od 51 do 100 rokov, preto parametre v tomto období sa výrazne odlišujú od vývoja v ďalších rokoch. Úlohou Makehamovej konštanty C je posun krivky funkcie miery úmrtnosti v smere osi y . Tento parameter nadobúda hodnoty blízke nule a vo všeobecnosti má klesajúcu tendenciu.

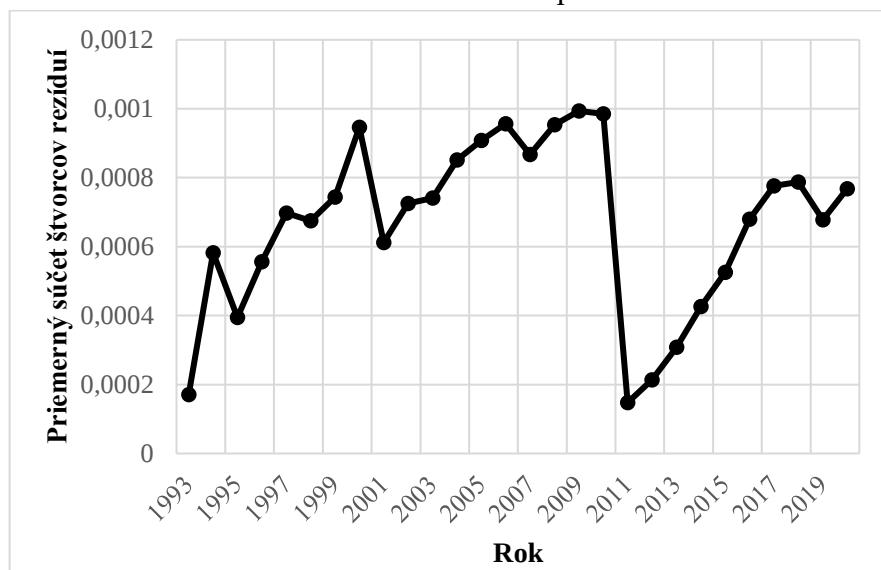
Graf č. 13 - Vývoj parametra C KM modelu v čase



Zdroj: vlastné spracovanie

O už niekoľkokrát spomenutej podobnosti Kannistovho-Makehamovho modelu úmrtnosti s Gompertzovým-Makehamovým zákonom svedčí aj graf č. 14, ktorý zobrazuje vývoj priemerného súčtu štvorcov rezíduí KM modelu. Tento graf je prakticky identický s grafom č. 9 zobrazujúcim MSE GM modelu v predchádzajúcej kapitole. Pri porovnaní priemerného súčtu rezíduí týchto dvoch modelov je rozdiel rozpoznateľný až na piatom desatinnom mieste. Najlepšie sa nám teda opäť podarilo vyrovnať krivku skutočnej vekovo špecifickej miery úmrtnosti pomocou Kannistovho-Makehamovho zákona úmrtnosti v rokoch 1993 a 2011 a najhoršie to na základe hodnôt MSE dopadlo v rokoch 2009 a 2010.

Graf č. 14 - MSE KM modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020



Zdroj: vlastné spracovanie

V konečnom dôsledku je Kannistov-Makehamov zákon úmrtnosti veľmi podobný Gompertzovmu-Makehamovmu zákonu úmrtnosti. Hoci patrí k logistickým modelom a medzi vynájdením týchto dvoch zákonov úmrtnosti je viac ako 100 rokov, ich výsledky na populácii Slovenska sú prakticky zhodné.

4.1.3 Heligmanov - Pollardov model úmrtnosti

Heligmanov-Pollardov zákon úmrtnosti sa skladá z deviatich biologicky interpretovateľných parametrov, a teda môžeme ho zaradiť už ku komplexnejším modelom. Matematicky je vyjadrený vzťahom (3.6). Komplexnosť tohto zákona sa odzrkadľuje v jeho tvare, ktorý nepredstavuje už len jednoduchú exponenciálne rastúcu krivku. Naopak je možné ho omnoho viac tvarovať, čo má za následok, že sa vie lepšie prispôbiť trendu úmrtnosti v oblasti havarijného hrbu, ako aj vo vysokých vekoch. Keďže v ostatných štátoch sa používa pre všetky vekové kategórie, predpokladáme, že aj pre Slovenskú republiku sa osvedčí ako vhodný kandidát na modelovanie úmrtnosti na celom vekovom intervale. Tabuľka č. 6 pozostáva z odhadovaných parametrov Heligmanovho-Pollardovho zákona úmrtnosti, osobitne pre mužov aj ženy, ale tiež so zanedbaním pohlavia ako faktora úmrtnosti za roky 1993 a 2020 na intervale vekov od narodenia po 50 rokov a od 51 až do 100 rokov. Súčasťou tabuľky č. 6 sú aj štatistické miery presnosti modelov MSE a RMSE

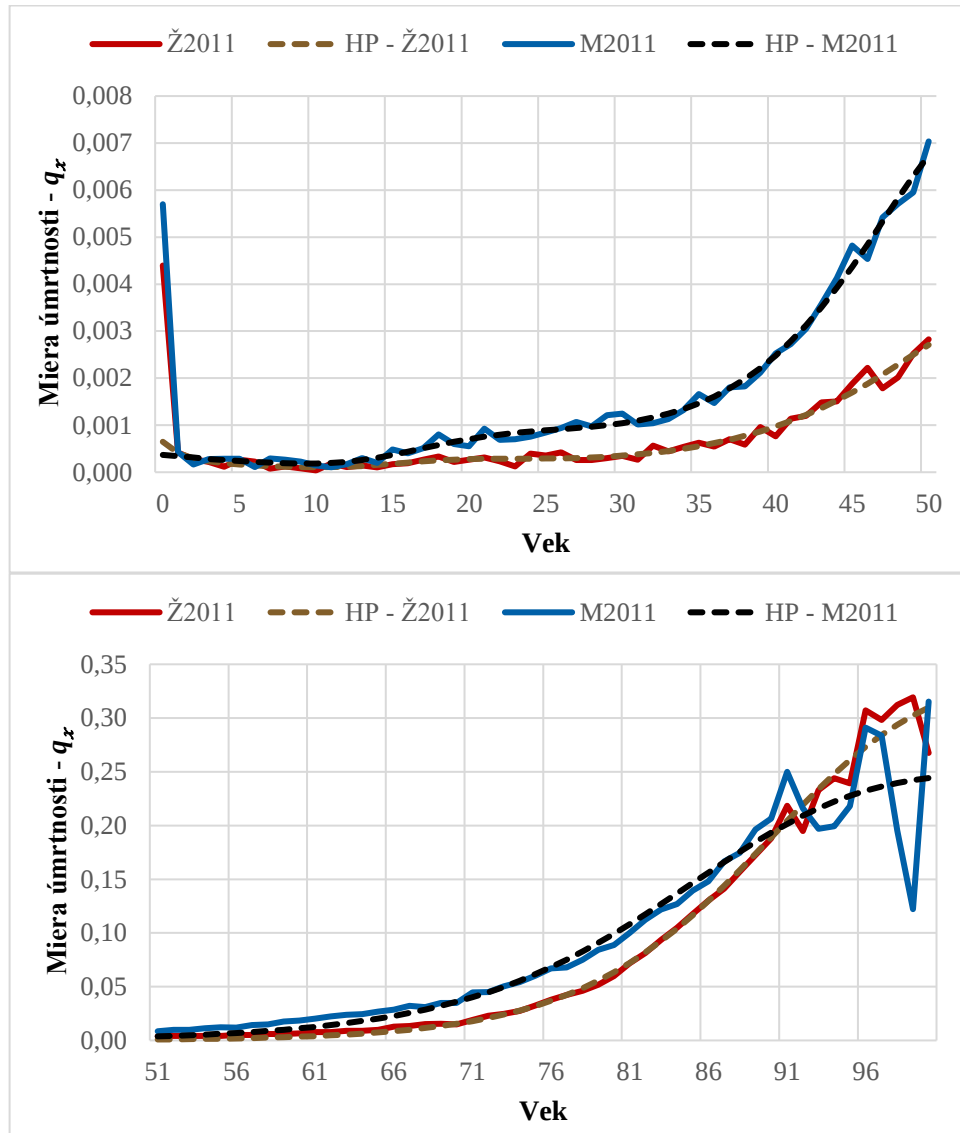
na posúdenie vhodnosti modelov a ich vzájomné porovnanie. V tabuľke č. 6 môžeme vidieť, že parametre *A* až *F* sú pre dva vekové intervaly rovnaké. Je to z toho dôvodu, že keďže sa tieto parametre vzťahujú k detskej úmrtnosti a úmrtnosti v havarijnom hrbe, nemajú vplyv na vekový interval nad 50 rokov. Parametre *G*, *H* a *K* sú rozdielne, pretože exponenciálny nárast úmrtnosti v závislosti od veku sa pre vekové intervaly líši. Ak by sme chceli mať len jednu súpravu parametrov, museli by sme zvoliť pri odhade parametrov iné vekové intervaly.

Tabuľka č. 6 - Odhad parametrov HP modelu pre roky 1993 a 2020

<i>Heligman-Pollard model</i>		1993			2020		
		Ženy	Muži	Spolu	Ženy	Muži	Spolu
<i>Parametre 50-</i>	A	0,00236	0,1432	0,00524	0,681	0,8257	0,00022
	B	1,5457	15,643	3,4783	12,6961	29,6209	-0,8973
	C	0,181	0,4828	0,223	1,1563	1,0833	0,036
	D	0,00023	0,00075	0,00045	0,00128	0,0181	0,0805
	E	28,2974	14,8924	12,9132	0,0456	0,2464	0,2185
	F	17,3594	21,0134	20,2703	14182,5	896,2	3333,1
	G	2,9E-05	5,3E-05	0,00004	1,6E-06	7,6E-09	1,1E-08
	H	1,1022	1,1144	1,1115	1,1586	1,3027	1,2831
	K	1,0559	1,0405	1,084	0,4455	-6E+15	1
<i>Parametre 50+</i>	A	0,00236	0,1432	0,00524	0,681	0,8257	0,00022
	B	1,5457	15,643	3,4783	12,6961	29,6209	-0,8973
	C	0,181	0,4828	0,223	1,1563	1,0833	0,036
	D	0,00023	0,00075	0,00045	0,00128	0,0181	0,0805
	E	28,2974	14,8924	12,9132	0,0456	0,2464	0,2185
	F	17,3594	21,0134	20,2703	14182,5	896,2	3333,1
	G	0,00033	5,1E-05	5,8E-05	5,3E-10	3,9E-08	8E-09
	H	1,0698	1,1054	1,0983	1,2623	1,2106	1,2261
	K	-959,7	1,0284	1,0221	1,0145	1,0193	1,0169
MSE		0,00024	0,0004	0,0001	0,00026	0,00041	0,00028
RMSE		0,01552	0,01999	0,01006	0,01598	0,0202	0,01664

Zdroj: vlastné spracovanie

Graf č. 15 - HP model úmrtnosti mužov a žien pre rok 2011

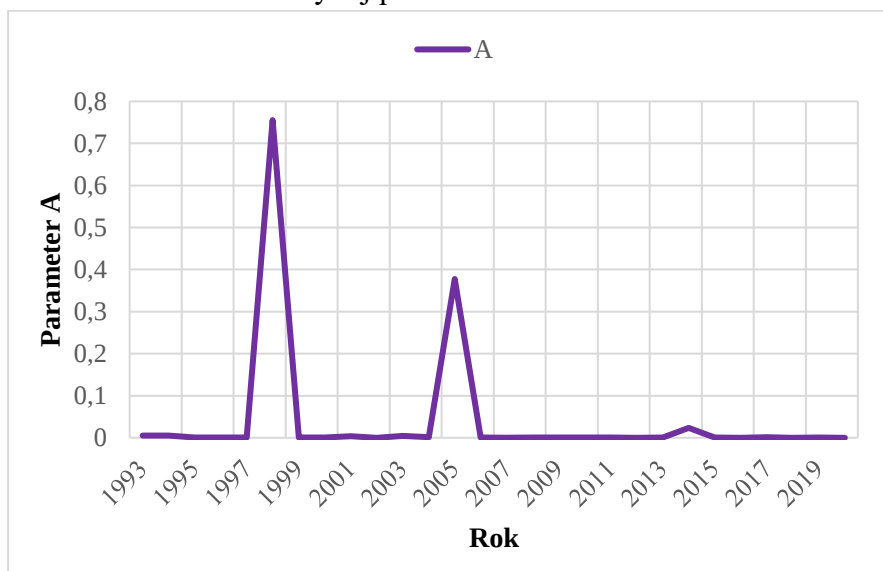


Zdroj: vlastné spracovanie

V grafickej analýze sa budeme sústrediť na rok 2011. Graf č. 15 obsahuje pozorované vekovo špecifické úmrtnosti mužov (M2011) a žien ($\check{Z}2011$) spolu s výsledkami modelovania pomocou Heligmanovho-Pollardovho zákona úmrtnosti (HP – M2011, GP – $\check{Z}2011$). Čo sa týka dojcenskej úmrtnosti, zatiaľ žiadny z modelov ju nebol schopný vyjadriť. Hoci pri Heligmanovom-Pollardovom zákone úmrtnosti ide o zložitejší model, nie pre všetky roky sa nám podarilo dospieť k správne mu odhadu dojcenskej úmrtnosti. Tak to bolo aj v znázornenom roku 2011. Okrem tohto nedostatku, na vekovom intervale (0; 50) dosahuje výborné výsledky u oboch pohlaví. Najväčšou výhodou modelu na tomto intervale je schopnosť prispôbiť sa úmrtnosti v oblasti havarijného hrbu, ktorý je pozorovateľný

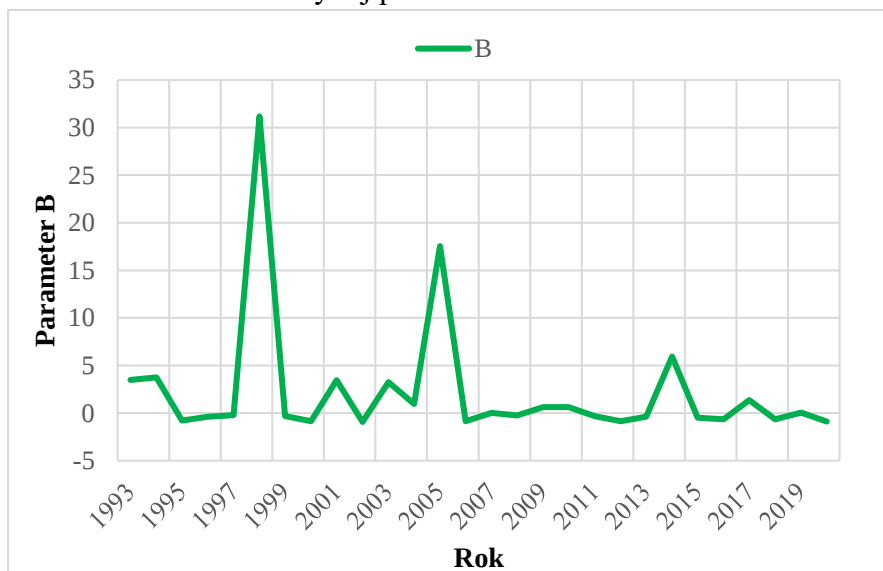
predovšetkým na mužskej populácii. V druhej časti vekového intervalu je to o niečo horšie. U mužov je viditeľné malé podhodnotenie úmrtnosti medzi 51 a 67 rokom života. Nadhodnotenie je naopak pozorovateľné medzi vekmi 76 až 87. Pre obe pohlavia je tiež možné si všimnúť výraznejšie podhodnotenie úmrtnosti vo veku nad 89 rokov, čo je však spôsobené najmä volatilitou údajov o úmrtnosti vo vysokých vekoch.

Graf č. 16 - Vývoj parametra A HP modelu v čase



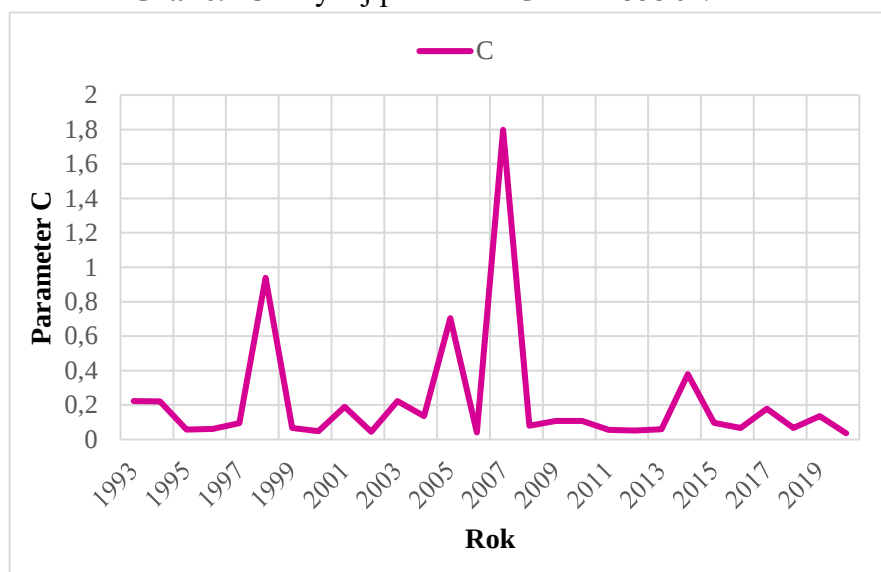
Zdroj: vlastné spracovanie

Graf č. 17 - Vývoj parametra B HP modelu v čase



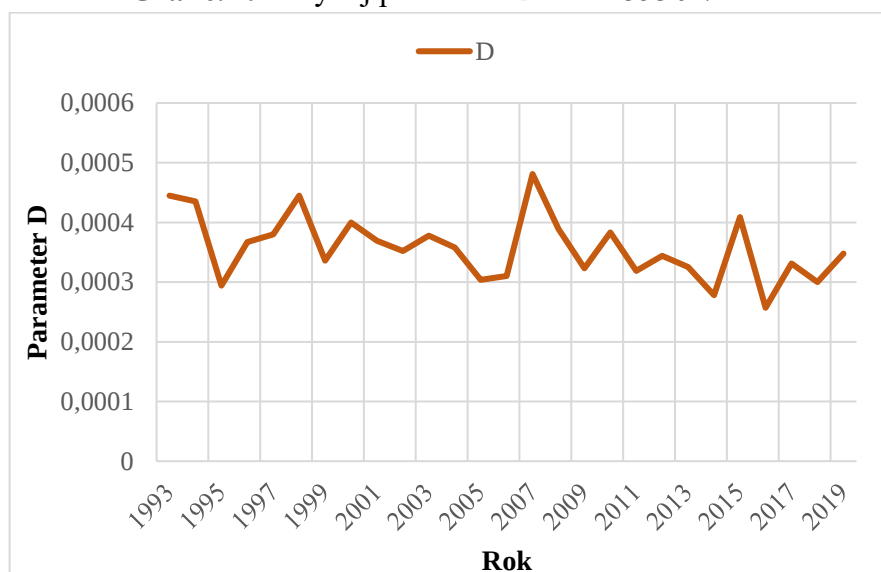
Zdroj: vlastné spracovanie

Graf č. 18 - Vývoj parametra C HP modelu v čase



Zdroj: vlastné spracovanie

Graf č. 19 - Vývoj parametra D HP modelu v čase



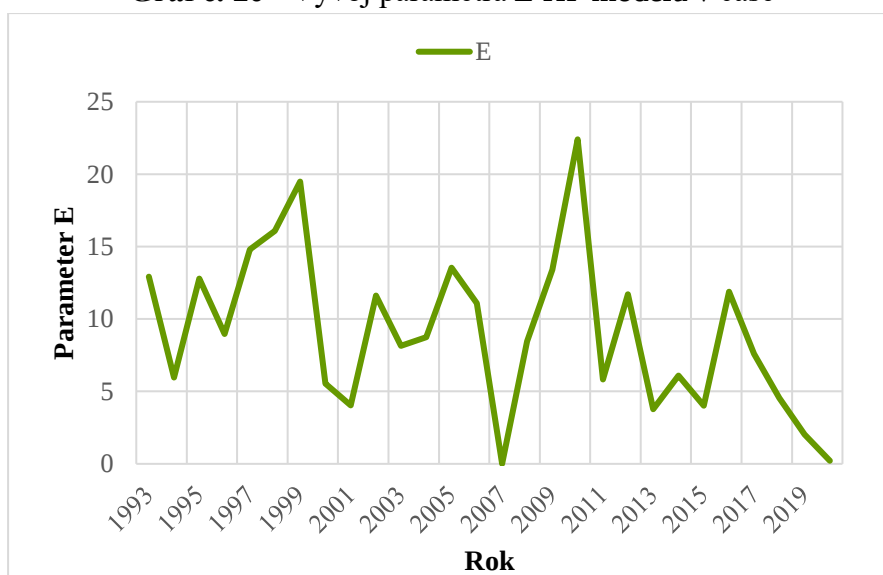
Zdroj: vlastné spracovanie

Na grafoch č. 16 až 24 je znázornený vývoj parametrov Heligmanovho-Pollardovho zákona úmrtnosti počas analyzovaného obdobia medzi rokmi 1993 a 2020. Parametre A, B a C tvoria prvý člen tohto zákona úmrtnosti zameraný na úmrtnosť detí. Parameter A ovplyvňuje úmrtnosť do 1 roka, parameter B vplyva práve na úmrtnosť v prvom roku života a napokon parameter C pôsobí na úmrtnosť až po začiatok dospelosti, teda približne do 18 roku života. Tieto parametre nadobúdali počas sledovaného obdobia malé hodnoty, najmä

z toho dôvodu, že pravdepodobnosť úmrtia v mladom veku je nízka. V ojedinelých prípadoch je na grafoch vidieť prudké odchýlenie od ostatných rokov. V týchto rokoch sme však neboli schopní spoľahlivo odhadnúť úmrtnosť do 1 roku života.

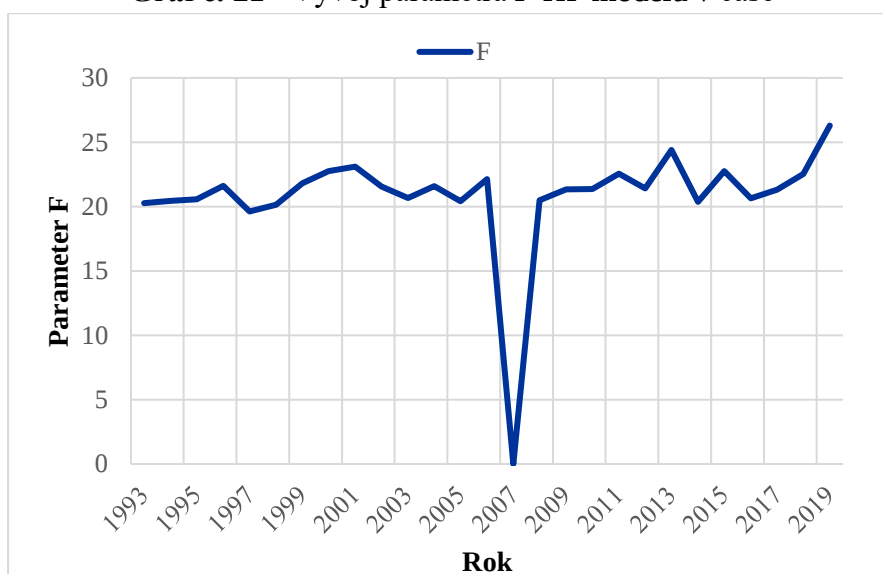
Vývoj úmrtnosti v oblasti havarijného hrbu ovplyvňuje parametre D , E a F , pričom parameter D zachytáva výšku hrbu, E meria jeho šírku a parameter F reprezentuje umiestnenie hrbu na vekovej osi. Na základe grafov č. 19, 20 a 21 môžeme povedať, že tieto parametre mali počas sledovaného obdobia volatilný ale vcelku stacionárny priebeh.

Graf č. 20 - Vývoj parametra E HP modelu v čase



Zdroj: vlastné spracovanie

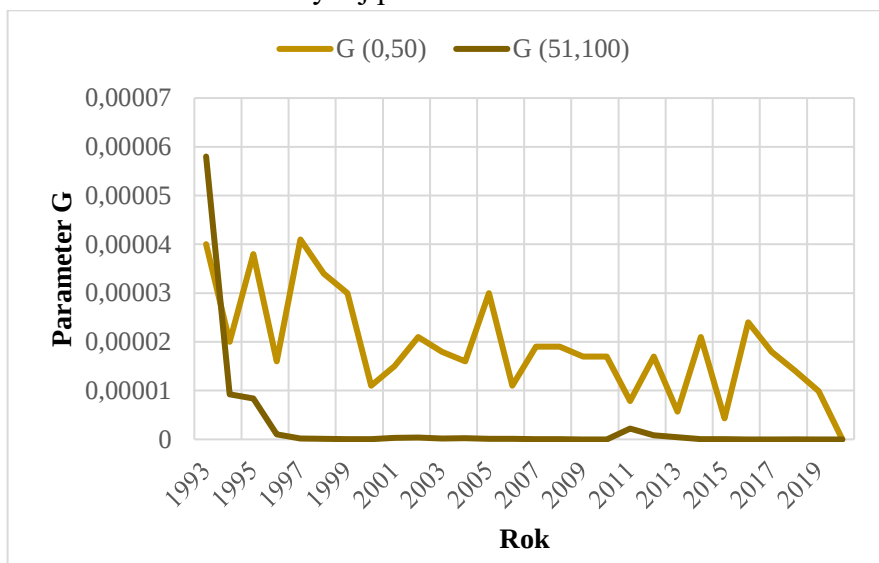
Graf č. 21 - Vývoj parametra F HP modelu v čase



Zdroj: vlastné spracovanie

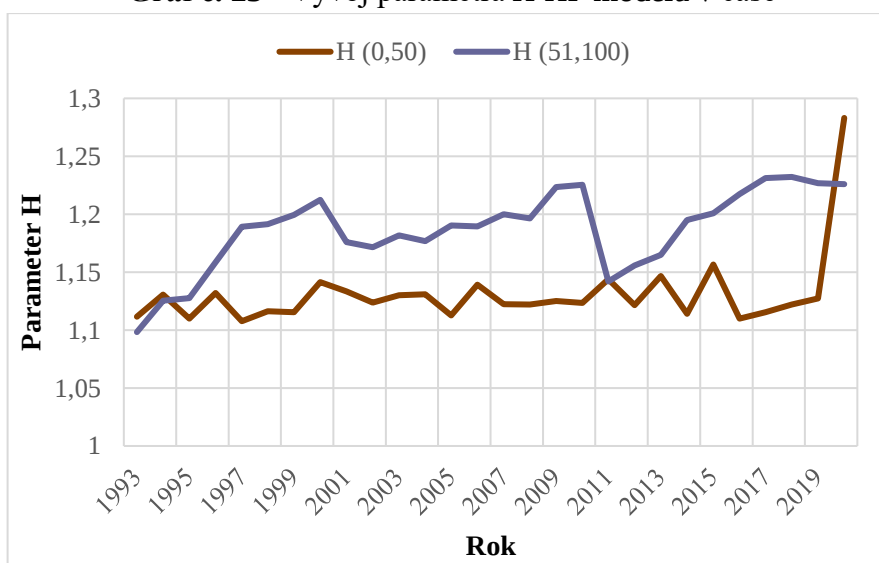
Posledný člen tvorený parametrami G , H a K reprezentuje úmrtnostné správanie dospelých ľudí až do konca vekového intervalu. Tieto parametre pritom ovplyvňujú základnú úroveň úmrtnosti a mieru nárastu, teda sklon krivky úmrtnosti. Vzhľadom na to, že vplývajú na úmrtnosť počas dospelosti, sú ako jediné rozdielne pre dva vekové intervaly $\langle 0; 50 \rangle$ a $\langle 50; 100 \rangle$. Ich hodnoty sú však podobné, takže nepotrebujeme osobitne osi pre jednotlivé intervaly vekov. Parameter G je blízky nule a má klesajúcu tendenciu. Zvyšné dva parametre sú vo všeobecnosti stacionárne a nadobúdajú hodnoty okolo 1.

Graf č. 22 - Vývoj parametra G HP modelu v čase



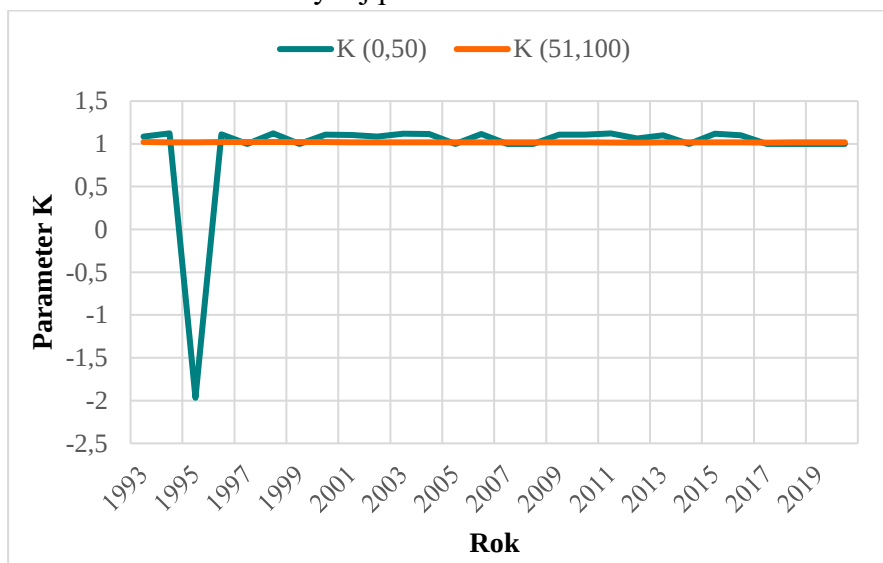
Zdroj: vlastné spracovanie

Graf č. 23 - Vývoj parametra H HP modelu v čase



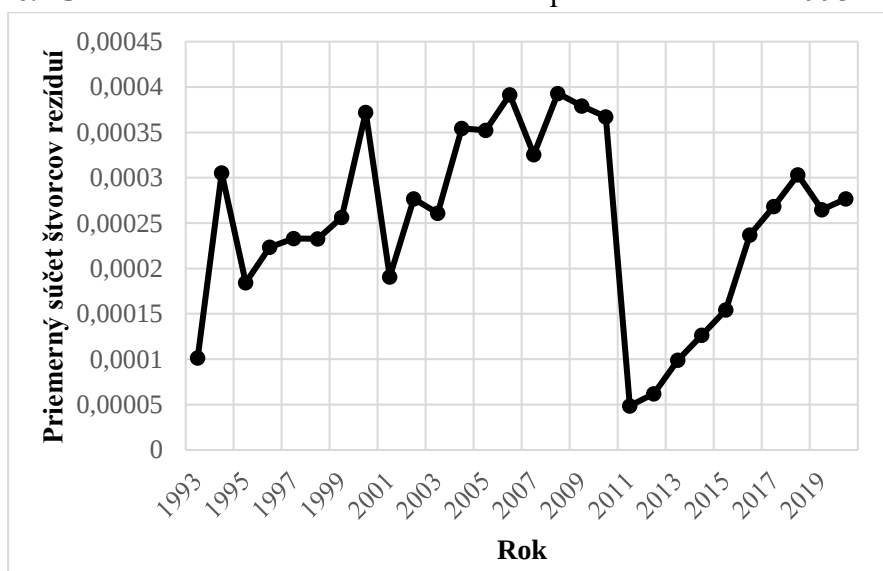
Zdroj: vlastné spracovanie

Graf č. 24 - Vývoj parametra K HP modelu v čase



Zdroj: vlastné spracovanie

Graf č. 25 - MSE HP modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020



Zdroj: vlastné spracovanie

Graf č. 25 zachytáva vývoj priemerného súčtu rezíduí MSE v jednotlivých rokoch pozorovaného obdobia. Tvar krivky je podobný s vývojom MSE prechádzajúcich modelov (Gompertzov-Makehamov zákon úmrtnosti, Kannistov-Makehamov zákon úmrtnosti), avšak nadobúda nižšie hodnoty, čo nasvedčuje tomu, že dosahuje lepšie výsledky. Podobnosť tvaru pripisujeme špecifickým črtám mier úmrtnosti v konkrétnych rokoch spôsobeným systematickými nepravidelnosťami reálnych údajov a ich volatilitou vo vysokých vekoch. Najlepšie hodnoty priemerného súčtu štvorcov rezíduí po vyrovnaní

krivky pozorovanej vekovo špecifickej miery úmrtnosti na základe Heligmanovho-Pollardovho zákona úmrtnosti boli namerané v rokoch 1993 a 2011. Na druhej strane, najhoršie sa nám podarilo priblížiť skutočným mieram úmrtnosti v rokoch 2006 a 2008.

Na záver môžeme povedať, že na základe výsledkov modelovania úmrtnosti Heligmanovým-Pollardovým zákonom úmrtnosti v rokoch od 1993 až po 2020, považujeme tento zákon vhodný na modelovanie úmrtnosti slovenskej populácie. Predovšetkým dobré výsledky dosahuje pri modelovaní havarijného hrbu. Naopak, pri dojčenskej úmrtnosti sme nie vždy dospeli k správnym hodnotám miery úmrtnosti, ani s takto zložitým modelom. Vo vysokých vekoch sme pozorovali mierne podhodnotenie úmrtnostného správania populácie Slovenska.

4.1.4 *CoDe model úmrtnosti*

CoDe model, podobne ako predchádzajúci zákon úmrtnosti, patrí ku komplikovanejším modelom, ktoré by sa svojim tvarom mali vedieť prispôbiť úmrtnostným trendom populácie. Skladá z piatich aditívnych členov reprezentujúcich úmrtnosť v za sebou idúcich fázach života (detská úmrtnosť, úmrtnosť mladistvých a úmrtnosť v období rannej dospelosti, neskorej dospelosti, ale aj v starobe), ktoré sú súčasťou matematického vyjadrenia (3.7). Aditívne členy obsahujú celkovo 10 parametrov, pričom šesť z nich sa v čase mení. Parametre, ktoré sú konštantné v čase, je potrebné v prvom rade nastaviť na základe potrieb populácie Slovenska.

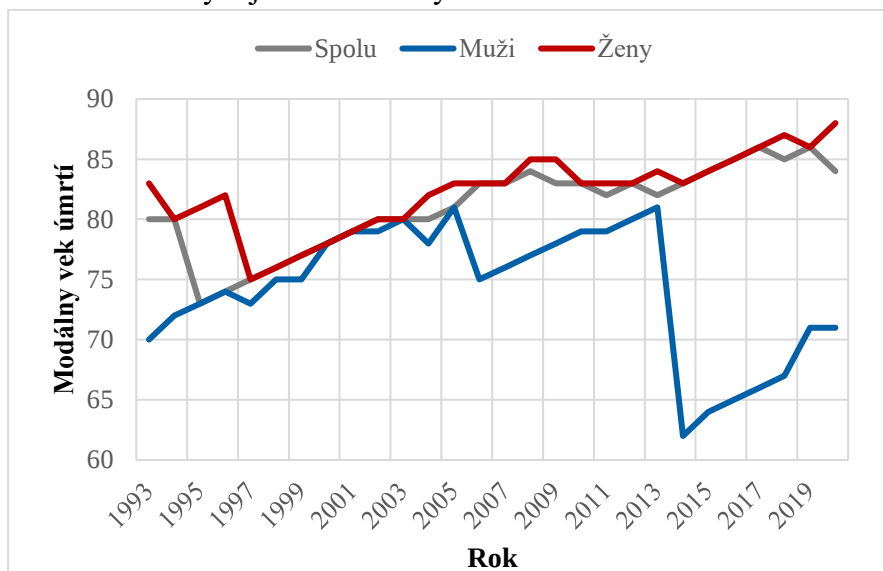
Parameter m predstavuje vek, v ktorom je nárast miery úmrtnosti najprudší. V pozorovanom období od 1993 po 2020 bol najprudší nárast mier úmrtnosti viditeľný medzi 14 až 18 rokom života, preto sme zvolili parameter $m = 16$ rokov.

Parameter M reprezentuje modus dĺžky života spomínaný v kapitole 1.2. Jeho hodnota je rozdielna v závislosti od pohlavia a takisto sa v čase vyvíja, hoci v modeli ju považujeme za konštantnú. Vývoj modusu dĺžky života zachytáva graf č. 26. Pri modelovaní sme vyskúšali rôzne hodnoty parametra M a usúdili sme, že je ideálne použiť $M = 80$ rokov, čo je približne priemerná hodnota modusu dĺžky života.

Parameter h môžeme považovať za pomocný parameter, ktorý nám umožňuje na základe modálneho veku celý vekový interval rozdeliť na približne rovnako veľké časti, ktoré budú korešpondovať so životnými fázami. Preto sme si zvolili $h = 30$.

Posledným časovo invariantným parametrom je parameter g , ktorý predstavuje hornú hranicu miery úmrtnosti, teda hodnotu, ku ktorej sa približuje pravdepodobnosť úmrtia osoby vo veku nad 110 rokov. Podľa rôznych zdrojov, ktorým sme sa venovali už v teoretickej časti 3.1.1.4, je táto hranica rôzna, avšak hodnota parametra g nemá veľký vplyv na výsledok modelu, preto sme použili $g = 0,7$ rovnako ako autori CoDe modelu. Okrem parametrov obsahuje vzťah pre CoDe model aj dve konštanty (3.8) a (3.9), ktorých hodnoty závisia od vypočítaných parametrov a slúžia na to, aby medzi úrovňami úmrtnosti v rôznych fázach života bol plynulý prechod. Na základe komplexnosti predpokladáme vhodnosť CoDe modelu na celom vekovom intervale.

Graf č. 26 - Vývoj modusu dĺžky života medzi rokmi 1993 až 2020



Zdroj: vlastné spracovanie

Tabuľka č. 7 zahŕňa odhadované parametre CoDe modelu spolu s vypočítanými konštantami zvlášť pre mužov aj ženy, ale tiež so zanedbaním pohlavia ako faktora úmrtnosti za roky 1993 a 2020 na celom vekovom intervale. Keďže v modeli sú oddelené rôzne fázy života pomocou indikátora $I(*)$, nie je potrebné rozdeľovať vekový interval na viac častí na zaručenie vyššej presnosti modelu, čo považujeme za veľkú výhodu. Súčasťou tabuľky č. 7 sú aj štatistické miery presnosti modelov MSE a RMSE na posúdenie vhodnosti modelov a ich vzájomné porovnanie.

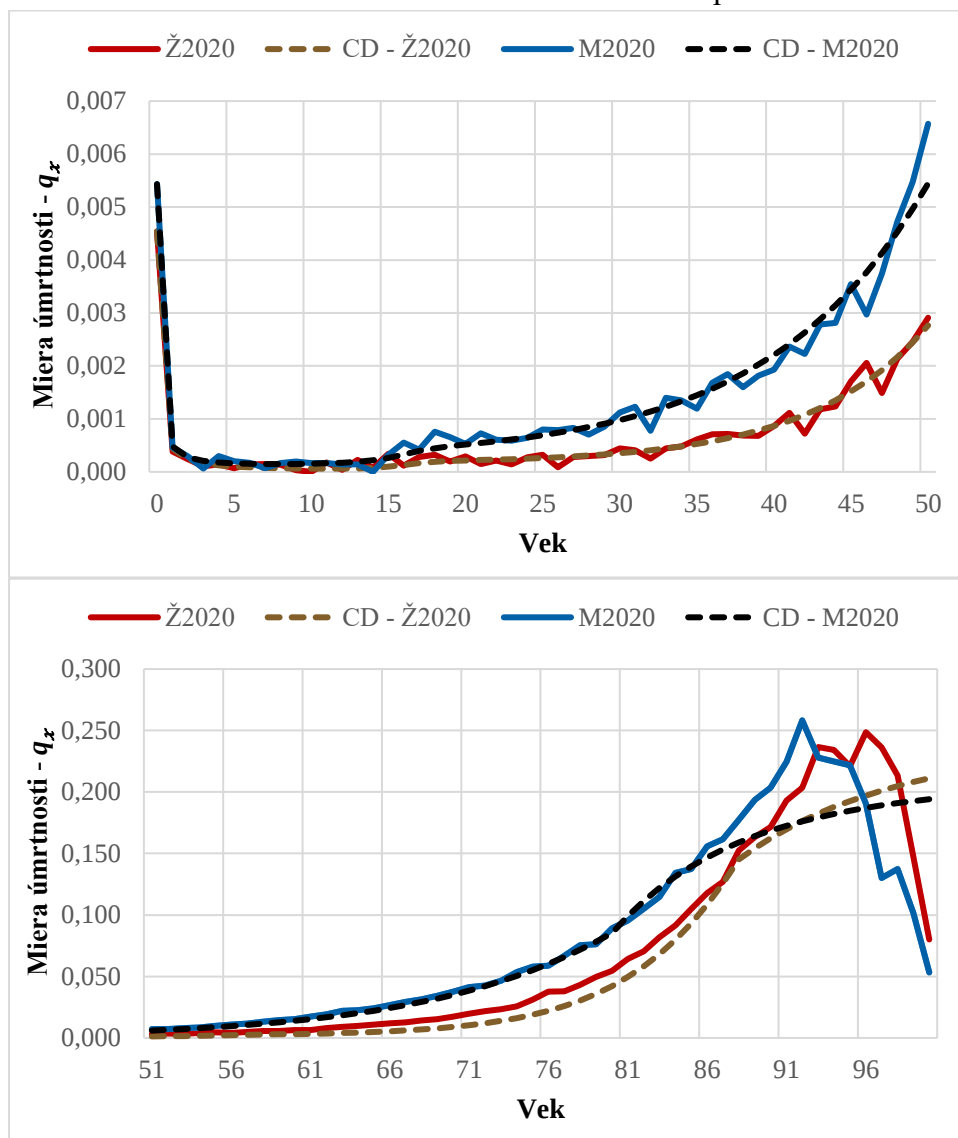
Tabuľka č. 7 - Odhad parametrov CD modelu pre roky 1993 a 2020

CoDe model		1993			2020		
		Ženy	Muži	Spolu	Ženy	Muži	Spolu
Parametre 0-100	A	0,00101	0,0006	0,0006	0,00048	0,00047	0,0005
	B	0,1174	0,04999	0,0583	0,105	0,0878	0,1005
	a	0,00019	-0,0002	-0,0002	0,00014	0,00021	0,00022
	m	16	16	16	16	16	16
	M	80	80	80	80	80	80
	h	30	30	30	30	30	30
	b1	0,1154	0,0721	0,0816	0,1302	0,0972	0,111
	g	0,7	0,7	0,7	0,7	0,7	0,7
	b2	0,0795	0,1119	0,0941	0,18	0,09684	0,0677
	b3	0,0864	0,0733	0,079	-0,086	-0,104	0,0928
	c1	-0,0036	0,00432	0,00144	0,0018	-4E-05	-0,0048
	c2	-0,0092	0,03444	0,0134	0,24303	0,20718	-0,025
MSE		0,00023	0,00054	0,00012	0,00035	0,00052	0,00032
RMSE		0,01504	0,02315	0,01072	0,01861	0,02277	0,0179

Zdroj: vlastné spracovanie

Na grafe č. 27 sme analyzovali schopnosť CoDe modelu opísať úmrtnostné správanie populácie Slovenska v roku 2020. Graf č. 27 pozostáva z pozorovaných vekovo špecifických mier úmrtnosti mužov (M2020) a žien (Ž2020) spolu s výsledkami modelovania pomocou CoDe modelu (CD – M2020, CD – Ž2020). Veľkou výhodou CoDe modelu je, že ako jediný nemal problém sa prispôbiť dojčenskej úmrtnosti počas celého sledovaného obdobia. Na intervale do 51 rokov CoDe model vedel opísať úmrtnostný trend v oblasti havarijného hrbu, hoci v roku 2020 tento hrb nie je taký viditeľný. Je to z toho dôvodu, že v roku 2020, aj v dôsledku pandémie, polícia zaznamenala najmenej dopravných nehôd za posledných desať rokov, ale aj celkovo od vzniku Slovenskej republiky. Na cestách zahynulo celkovo 224 osôb, čo je o 21 menej obetí ako v roku 2019 a zároveň druhý najnižší počet od vzniku samostatnej Slovenskej republiky (Ministerstvo vnútra SR, 2021). Na intervale od 51 rokov je to na základe grafu č. 27 o niečo horšie. U žien je možné pozorovať výrazné podhodnotenie úmrtnosti medzi 65 a 87 rokom života. U oboch pohlaví mal CoDe model problém priblížiť sa mieram úmrtnosti vo vysokých vekoch nad 87 rokov.

Graf č. 27 - CD model úmrtnosti mužov a žien pre rok 2020

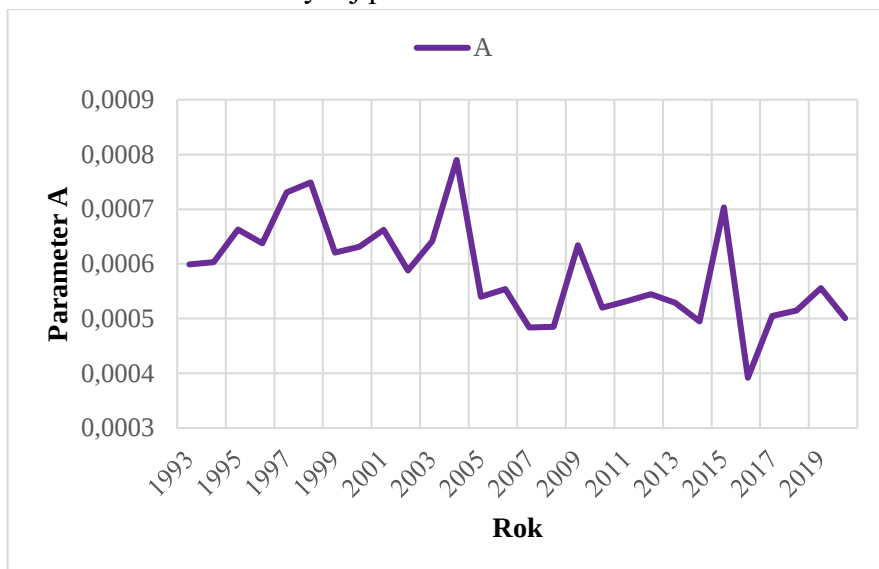


Zdroj: vlastné spracovanie

Zameriame sa teraz na vývoj jednotlivých parametrov CoDe modelu počas analyzovaného obdobia medzi rokmi 1993 a 2020, ktorý je znázornený na grafoch 28 až 33. Parametre A a B tvoria prvý člen tohto zákona úmrtnosti, ktorý sa venuje úmrtnosti v nízkom veku. Parameter A odráža dojčenskú úmrtnosť a z grafu č. 28 vidíme, že má klesajúci trend, čo nasvedčuje tomu, že úmrtnosť v rannom štádiu života sa znižuje. Parameter B reprezentuje detskú úmrtnosť. To, že na grafe č. 29 môžeme pozorovať rastúci trend, neznamená, že detská úmrtnosť rastie, pretože parameter B je ovplyvnený parametrom A . Efekt zmien v hodnote parametra B je v čase relatívne malý, preto zmeny úmrtnosti v detstve možno prisúdiť najmä zmenám parametra A . Parameter a slúži na opis úmrtnosti v oblasti

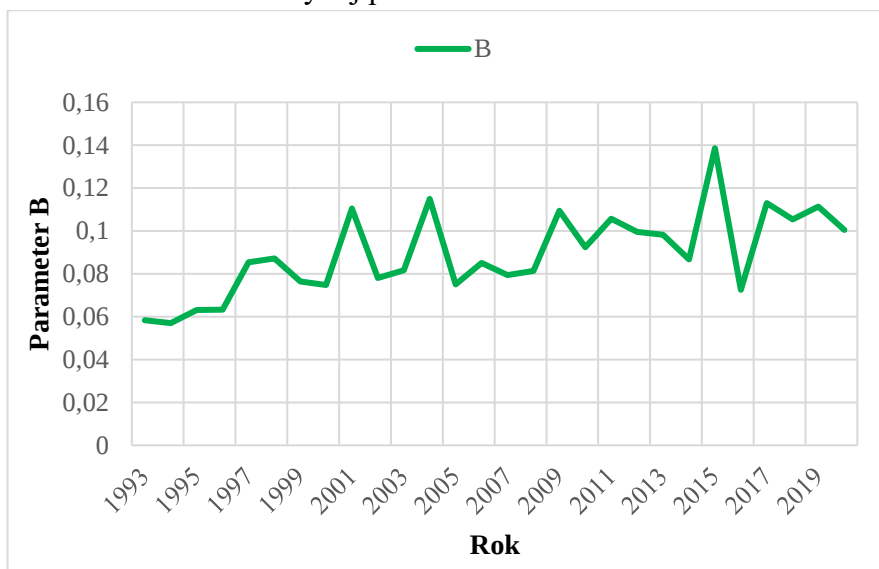
havarijného hrbu. Nadobúda nízke hodnoty blízke 0 a podľa grafu č. 30 vidíme, že má rastúci trend.

Graf č. 28 - Vývoj parametra A v CD modeli v čase



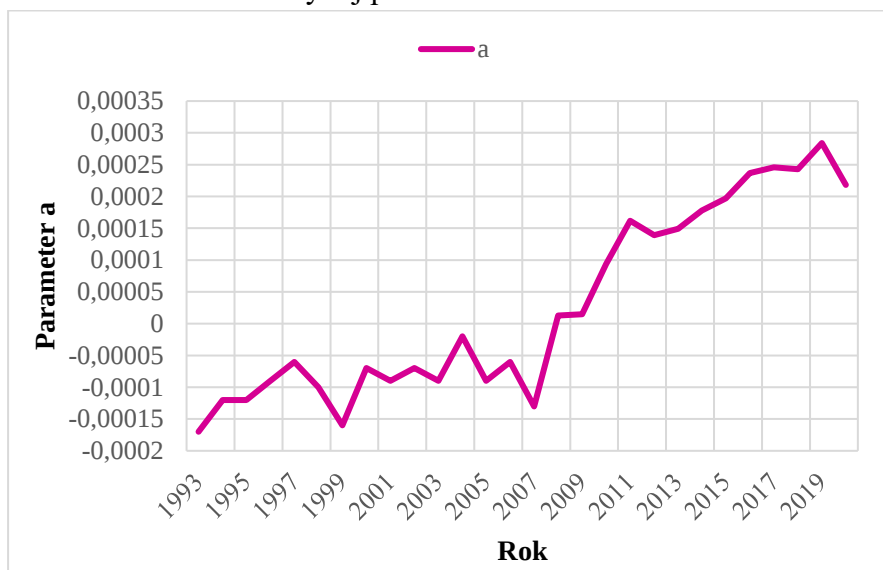
Zdroj: vlastné spracovanie

Graf č. 29 - Vývoj parametra B v CD modeli v čase



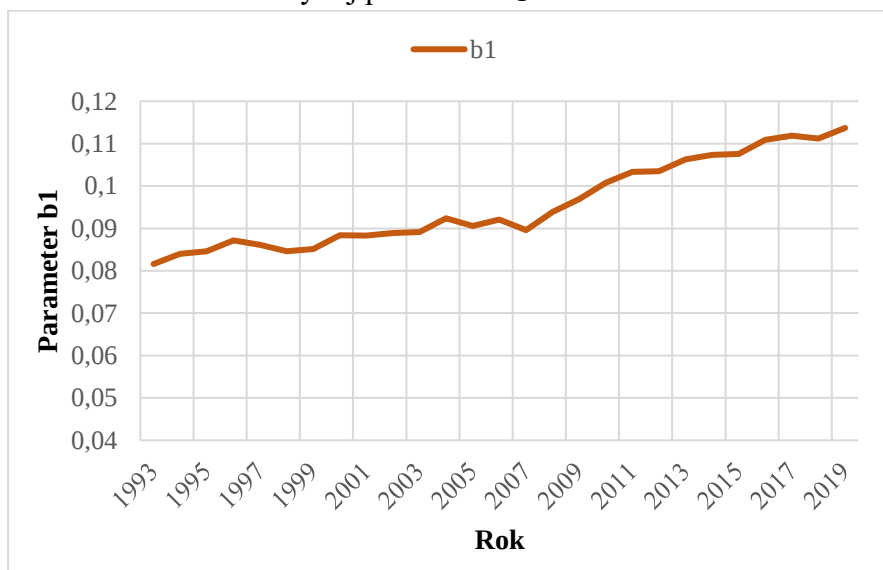
Zdroj: vlastné spracovanie

Graf č. 30 - Vývoj parametra a v CD modeli v čase



Zdroj: vlastné spracovanie

Graf č. 31 - Vývoj parametra b_1 v CD modeli v čase



Zdroj: vlastné spracovanie

Parametre b_1 , b_2 a b_3 potom opisujú úmrtnosť v rôznych fázach života v závislosti od zvolených konštantných parametrov modelu. Vzhľadom na hodnoty $M = 80$ a $h = 30$, parameter b_1 reprezentuje sklon úmrtnosti do veku 50 rokov (skorá dospelosť), b_2 opisuje úroveň úmrtnosti medzi 50 až 80 rokom života (neskorá dospelosť) a parameter b_3 vyjadruje nárast úmrtnosti v závislosti od veku pre osoby nad 80 rokov (staroba). Podľa grafu č. 31 má parameter b_1 mierne rastúci trend. Sklon úmrtnosti do veku 50 rokov však nie je ovplyvnený

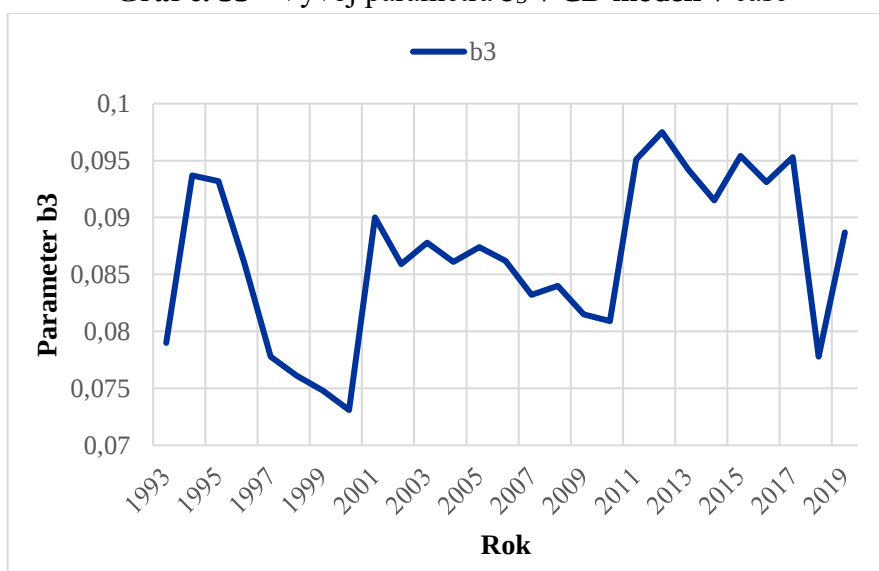
len týmto parametrom, ale závisí aj od úrovne parametrov predchádzajúcich členov. Úmrtnosť sa medzi 50 až 80 rokom života počas sledovaného obdobia podľa grafu č. 32 zlepšila, pričom v pandemickom roku 2020 je viditeľný výrazný nárast. Napokon vývoj parametra b_3 je veľmi kolísavý, čo je spôsobené volatilitou údajov o úmrtnosti vo vysokých vekoch.

Graf č. 32 - Vývoj parametra b_2 v CD modeli v čase



Zdroj: vlastné spracovanie

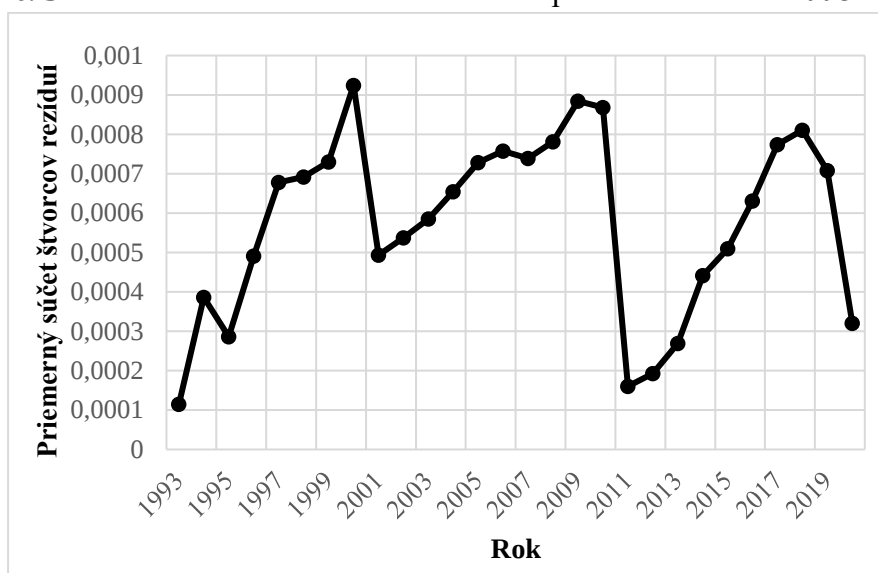
Graf č. 33 - Vývoj parametra b_3 v CD modeli v čase



Zdroj: vlastné spracovanie

Graf č. 34 zachytáva vývoj priemerného súčtu rezíduí MSE v jednotlivých rokoch pozorovaného obdobia. Samotný odhad parametrov bol v dôsledku komplexnosti modelu pomerne náročný. Na základe výsledkov MSE môžeme povedať, že CoDe model dosahuje lepšie výsledky než jednoduché modely (Gompertzov-Makehamov zákon úmrtnosti, Kannistov-Makehamov zákon úmrtnosti), avšak vo všeobecnosti horšie opisuje úmrtnostné správanie než Heligmanov-Pollardov zákon úmrtnosti. Najlepšie, a teda zároveň najnižšie hodnoty priemerného súčtu štvorcov rezíduí po vyrovnaní krivky pozorovanej vekovo špecifickej miery úmrtnosti na základe CoDe modelu, boli namerané v rokoch 1993 a 2011. Naopak, najmenej sa nám podarilo priblížiť skutočným mieram úmrtnosti v rokoch 2000 a 2009.

Graf č. 34 - MSE CD modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020



Zdroj: vlastné spracovanie

CoDe model považujeme za vhodný na modelovanie úmrtnosti v pozorovanom období. Jeho výhodou je schopnosť modelovať dojčenskú úmrtnosť a havarijný hrb. Vo vysokých vekoch nedosahoval až tak dobré výsledky, ako sme očakávali v súvislosti s jeho zložitosťou. Už samotné nastavenie konštantných parametrov si vyžadovalo dôslednú analýzu. Z toho dôvodu zložitnosť modelu považujeme za jeho najväčšiu nevýhodu. Existujú však už aj pokusy o zjednodušenie tohto modelu, najmä obmedzením počtu časovo závislých parametrov. Príkladom tohto pokusu je Bardoutsos a kol. (2018), ktorý si ale uvedomuje, že zníženie počtu parametrov vedie k robustnejším odhadom.

4.2 Stochastické modely úmrtnosti

4.2.1 Lee-Carterov model úmrtnosti

V tejto časti prejdeme na modelovanie úmrtnostného správania populácie Slovenskej republiky pomocou stochastických modelov úmrtnosti. Jedným z prvých a zároveň najznámejší je Lee-Carterov model úmrtnosti, pri ktorom pracujeme s logaritmiami vekovo špecifických centrálnych mier úmrtnosti, ktoré vyhovujú vzťahu (3.10). Tento model je vhodný na modelovanie celého vekového intervalu, pričom nemá problém sa prispôbiť anomáliám miery úmrtnosti, akými sú napríklad dojčenská úmrtnosť a havarijný hrb. Princíp modelovania pomocou stochastických modelov sa značne líši od tých deterministických. Lee-Carterov model obsahuje tri parametre, z ktorých dva závisia od veku a jeden sa vyvíja v závislosti od času. Na odhad parametrov sme použili metódu Habermana a Russolillo, ktorá vychádza z metódy najmenších štvorcov za použitia singulárneho rozkladu matice a pozornosť sme jej venovali v časti 3.3.1.

Tabuľka č. 8 - Odhad parametra $\kappa_t^{(2)}$ LC modelu

Rok	$\kappa_t^{(2)}$	Rok	$\kappa_t^{(2)}$
1993	16,646	2007	4,861
1994	20,793	2008	0,980
1995	22,138	2009	-2,016
1996	17,481	2010	-3,314
1997	17,542	2011	-9,337
1998	18,241	2012	-10,787
1999	14,966	2013	-14,556
2000	13,875	2014	-19,430
2001	13,340	2015	-16,407
2002	10,333	2016	-22,859
2003	10,052	2017	-21,953
2004	6,819	2018	-23,869
2005	8,292	2019	-29,535
2006	5,840	2020	-18,741

Zdroj: vlastné spracovanie

V tabuľkách č. 8 a 9 nájdeme výsledky odhadnutých parametrov Lee-Carterovho modelu pre populáciu Slovenskej republiky za časové obdobie od roku 1993 do 2020, pričom hodnoty parametrov závislých od veku sú uvedené v tabuľke č. 9 a hodnoty

parametra závislého od času sú uvedené v tabuľke č. 8. Z dôvodu veľkého množstva údajov obe tabuľky znázorňujú parametre modelu bez uvažovania pohlavia ako faktora, ktorý vplýva na úmrtnosť. Výstup modelovania úmrtnosti zvlášť pre mužov a ženy nájdeme v prílohe 3.

Tabuľka č. 9 - Odhad parametrov $\beta_x^{(1)}$ a $\beta_x^{(2)}$ LC modelu

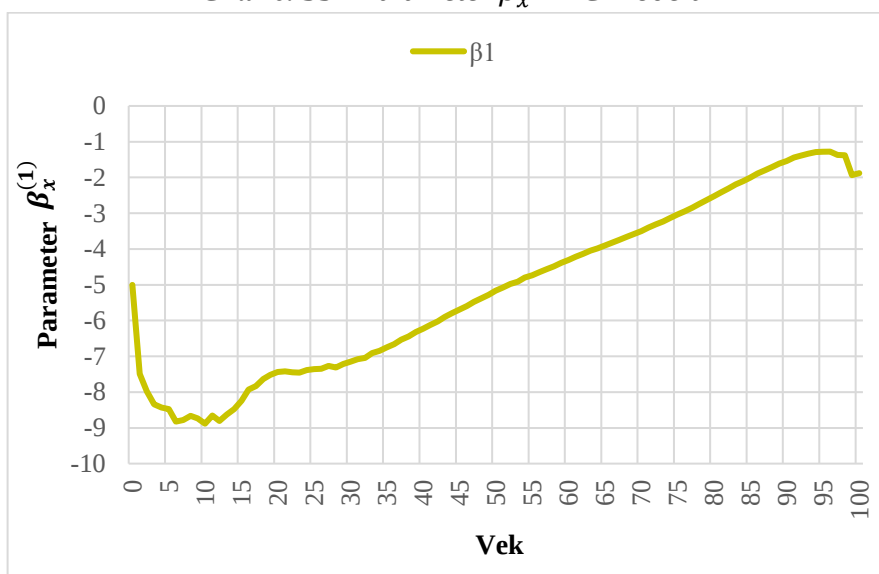
Vek	$\beta_x^{(1)}$	$\beta_x^{(2)}$	Vek	$\beta_x^{(1)}$	$\beta_x^{(2)}$	Vek	$\beta_x^{(1)}$	$\beta_x^{(2)}$	Vek	$\beta_x^{(1)}$	$\beta_x^{(2)}$	Vek	$\beta_x^{(1)}$	$\beta_x^{(2)}$
0	-5,006	0,015	20	-7,435	0,007	40	-6,234	0,015	60	-4,302	0,009	80	-2,526	0,009
1	-7,491	0,011	21	-7,420	0,011	41	-6,125	0,014	61	-4,214	0,008	81	-2,415	0,007
2	-7,972	0,016	22	-7,442	0,010	42	-6,020	0,014	62	-4,131	0,009	82	-2,307	0,007
3	-8,346	0,024	23	-7,458	0,011	43	-5,902	0,014	63	-4,049	0,008	83	-2,191	0,007
4	-8,427	0,016	24	-7,386	0,009	44	-5,789	0,014	64	-3,980	0,009	84	-2,101	0,006
5	-8,479	0,019	25	-7,357	0,011	45	-5,689	0,013	65	-3,904	0,010	85	-2,002	0,005
6	-8,824	0,026	26	-7,352	0,011	46	-5,591	0,013	66	-3,821	0,010	86	-1,887	0,005
7	-8,782	0,022	27	-7,268	0,010	47	-5,481	0,013	67	-3,742	0,010	87	-1,797	0,004
8	-8,667	0,014	28	-7,312	0,009	48	-5,378	0,013	68	-3,665	0,010	88	-1,714	0,004
9	-8,737	0,012	29	-7,210	0,010	49	-5,283	0,013	69	-3,579	0,010	89	-1,615	0,003
10	-8,883	0,020	30	-7,153	0,011	50	-5,164	0,011	70	-3,496	0,010	90	-1,536	0,002
11	-8,656	0,010	31	-7,077	0,009	51	-5,074	0,010	71	-3,404	0,010	91	-1,443	0,002
12	-8,809	0,015	32	-7,047	0,011	52	-4,978	0,011	72	-3,310	0,011	92	-1,391	0,001
13	-8,632	0,013	33	-6,909	0,011	53	-4,909	0,010	73	-3,229	0,011	93	-1,336	-0,001
14	-8,474	0,012	34	-6,848	0,011	54	-4,798	0,010	74	-3,129	0,011	94	-1,287	0,000
15	-8,248	0,012	35	-6,748	0,013	55	-4,737	0,010	75	-3,034	0,010	95	-1,281	0,001
16	-7,934	0,008	36	-6,662	0,010	56	-4,644	0,010	76	-2,942	0,010	96	-1,277	-0,001
17	-7,832	0,011	37	-6,538	0,010	57	-4,565	0,009	77	-2,849	0,009	97	-1,367	-0,003
18	-7,633	0,009	38	-6,446	0,014	58	-4,484	0,009	78	-2,741	0,009	98	-1,378	0,002
19	-7,521	0,008	39	-6,322	0,014	59	-4,388	0,009	79	-2,634	0,009	99	-1,929	-0,011

Zdroj: vlastné spracovanie

Pre lepšie pochopenie sú odhadované parametre Lee-Carterovho modelu za obdobie od roku 1993 po 2020 znázornené na grafoch č. 35, 36 a 37. Všeobecný tvar krivky logaritmu centrálnej miery úmrtnosti reprezentovaný parametrom $\beta_x^{(1)}$ vykazuje podľa očakávaní vysokú dojčenskú úmrtnosť, havarijný hrb medzi 18 až 24 rokom života a takmer lineárny nárast u starších vekových skupín. Parameter citlivosti úmrtnosti $\beta_x^{(2)}$ je kladný až do veku 92 rokov. To znamená, že do tohto veku sa úmrtnosť počas sledovaného obdobia zlepšovala. V porovnaní s inými vekmi sú hodnoty tohto parametra výrazne vyššie v rannom štádiu života okolo 5 až 15 roku života, čo naznačuje, že úmrtnosť na tomto vekovom intervale sa zlepšuje rýchlejšie ako v iných vekových skupinách. U ľudí nad 93 rokov je v niektorých

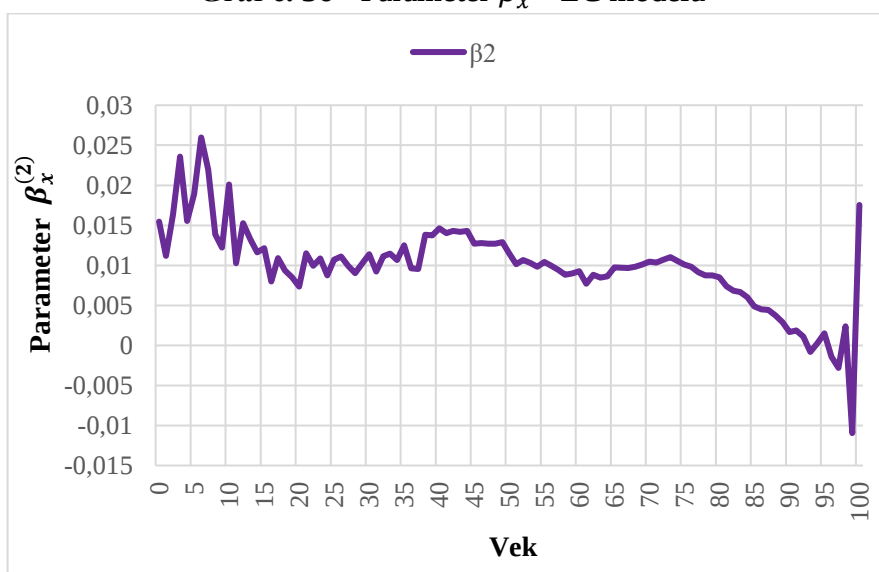
prípadoch parameter $\beta_x^{(2)}$ záporný alebo veľmi blízky 0, čo znamená, že úmrtnosť sa pre tieto veky zhoršuje alebo zostáva približne rovnaká. Index úmrtnosti $\kappa_t^{(2)}$ naznačuje vývoj úmrtnosti v čase. Na základe grafu č. 37 môžeme pozorovať, že má klesajúci trend, čo naznačuje zlepšovanie (pokles) úmrtnosti v čase. V poslednom pandemickom roku došlo k pomerne prudkému nárastu parametra $\kappa_t^{(2)}$, čo je v súlade s aktuálnou úmrtnostnou situáciou.

Graf č. 35 - Parameter $\beta_x^{(1)}$ LC modelu



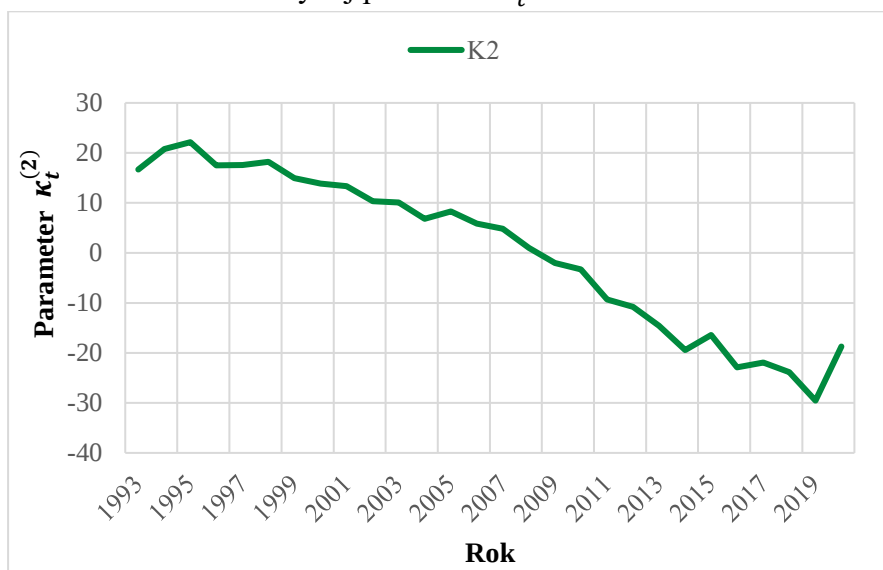
Zdroj: vlastné spracovanie

Graf č. 36 - Parameter $\beta_x^{(2)}$ LC modelu



Zdroj: vlastné spracovanie

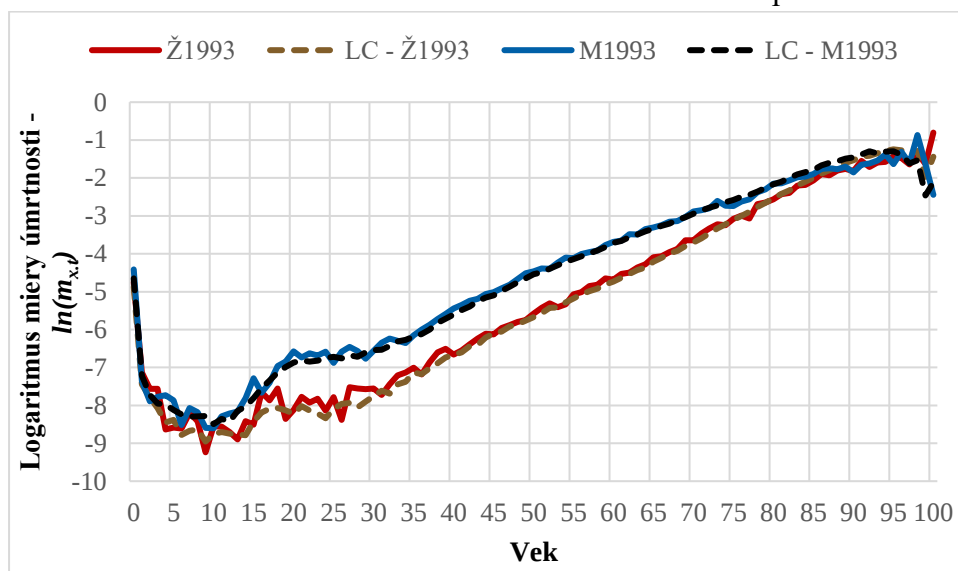
Graf č. 37 - Vývoj parametra $\kappa_t^{(2)}$ LC modelu v čase



Zdroj: vlastné spracovanie

Na grafe č. 38 sme znázornili pozorované centrálne miery úmrtnosti mužov (M1993) a žien (Ž1993) podľa veku prevedené do logaritmickej podoby na začiatku projekcie v roku 1993. Prerušovanými čiarami sú na grafe vykreslené vekovo špecifické centrálne miery úmrtnosti odhadované pomocou Lee-Carterovho stochastického modelu úmrtnosti.

Graf č. 38 - Ukážka LC modelu úmrtnosti mužov a žien pre rok 1993

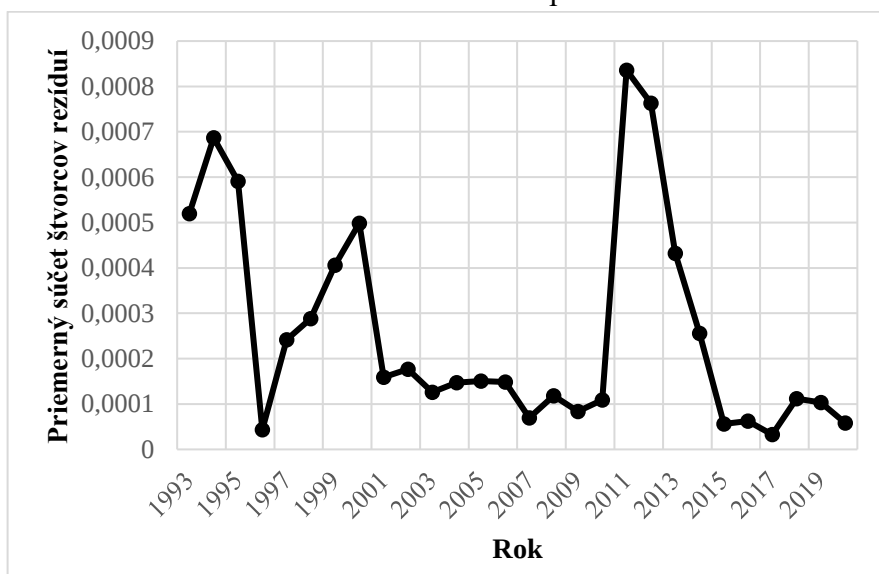


Zdroj: vlastné spracovanie

Zistili sme, že prispôsobené miery úmrtnosti sú veľmi blízke pozorovaným (skutočným) mieram úmrtnosti, okrem toho, že skutočné miery zahŕňajú nesystematickú volatilitu. Na grafe č. 38 nie sú viditeľné žiadne výrazné podhodnotenia respektíve nadhodnotenia úmrtnosti. Lee-Carterov model nemal problém sa prispôbiť dojčenskej úmrtnosti ako aj úmrtnosti v oblasti havarijného hrbu pozorovateľného najmä u mužskej časti populácie.

O kvalite Lee-Carterovho modelu svedčí vývoj priemerného súčtu rezíduí MSE zobrazeného na grafe č. 39. Čo je zaujímavé, vývoj MSE je protikladný ku deterministickým modelom. Kým u všetkých deterministických modelov bola úroveň priemerného súčtu rezíduí v roku 2011 najnižšia, a teda šlo o najlepší odhad, u Lee-Carterovho modelu to je presne opačne. Najspoľahlivejšie sa nám podarilo vyrovnať krivku logaritmu skutočnej vekovo špecifickej miery úmrtnosti pomocou Lee-Carterovho modelu v rokoch 1996 a 2017. Naopak najvyššie hodnoty priemerného súčtu štvorcov rezíduí sú namerané v rokoch 1994 a 2004.

Graf č. 39 - MSE LC modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020



Zdroj: vlastné spracovanie

Prognózovanie je jednou z najdôležitejších súčastí stochastického modelovania. Výhodou Lee-Carterovho modelu je, že keď už nájdeme hodnoty časovo invariantných parametrov $\beta_x^{(1)}$ a $\beta_x^{(2)}$, je potrebné predpovedať budúci vývoj len indexu úmrtnosti $\kappa_t^{(2)}$, ktorý je ako jediný závislý od času. Existuje niekoľko metód predikcie, napríklad použitie

regresných kriviek alebo jednou z najpopulárnejších metód, a to ARIMA modelovanie (autoregresný integrovaný kľzavý priemer), ktorému sme sa venovali v kapitole 3.3.2.

Tabuľka č. 10 - Prognóza $\kappa_t^{(2)}$ LC modelu s 95 % intervalom spoľahlivosti

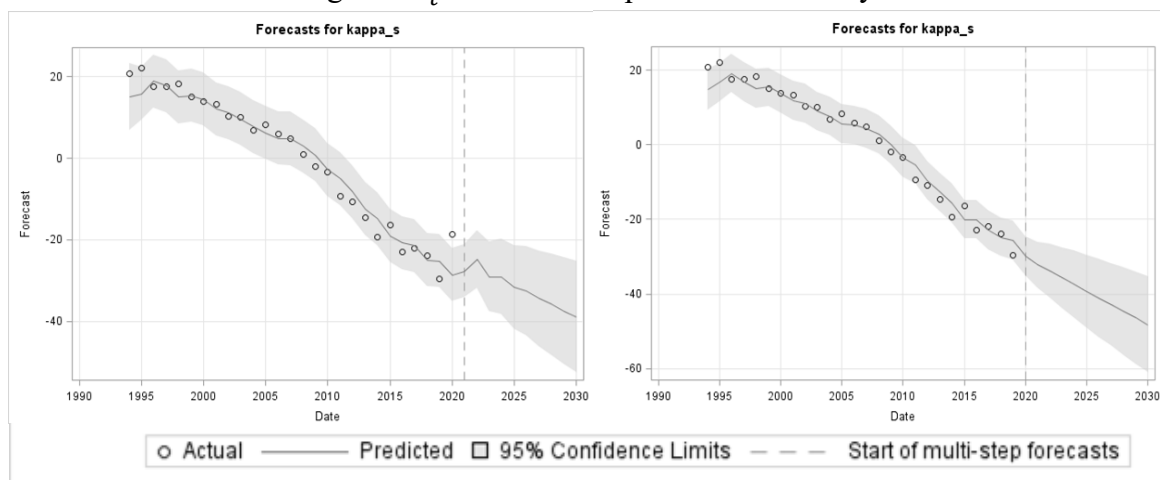
Rok	Prognóza	95 % Interval spoľahlivosti	
2021	-27,724	-34,222	-21,225
2022	-24,829	-31,826	-17,831
2023	-29,067	-37,632	-20,502
2024	-29,021	-38,286	-19,756
2025	-31,548	-41,795	-21,301
2026	-32,530	-43,483	-21,577
2027	-34,440	-46,162	-22,718
2028	-35,792	-48,175	-23,410
2029	-37,480	-50,524	-24,435
2030	-38,966	-52,620	-25,312
Rok	Prognóza Muži	95 % Interval spoľahlivosti	
2021	-27,636	-34,019	-21,252
2022	-26,434	-33,614	-19,253
2023	-29,411	-38,043	-20,780
2024	-30,363	-39,902	-20,824
2025	-32,297	-42,808	-21,786
2026	-33,754	-45,090	-22,419
2027	-35,443	-47,576	-23,310
2028	-37,019	-49,887	-24,152
2029	-38,650	-52,219	-25,081
2030	-40,255	-54,487	-26,022
Rok	Prognóza Ženy	95 % Interval spoľahlivosti	
2021	-25,708	-32,278	-19,138
2022	-22,773	-29,816	-15,729
2023	-26,823	-35,468	-18,178
2024	-26,588	-35,913	-17,263
2025	-28,981	-39,303	-18,660
2026	-29,763	-40,781	-18,744
2027	-31,533	-43,326	-19,739
2028	-32,696	-45,147	-20,246
2029	-34,232	-47,347	-21,117
2030	-35,540	-49,263	-21,816

Zdroj: vlastné spracovanie na základe výstupu zo SAS Enterprise Guide

Zvážili sme niekoľko možných modelov ARIMA na prognózovanie indexu úmrtnosti. Na identifikáciu najvhodnejšieho modelu sme použili Akaikovo informačné kritérium (AIC) opísané v časti 3.4. Všeobecne platí, že keď porovnávame modely, menšie AIC označuje lepšie vyhovujúci model. Podľa AIC je *ARIMA (1,1,0)* najlepšie vyhovujúcim modelom pre index úmrtnosti Lee-Carterovho modelu odhadovaného pre populáciu Slovenskej republiky bez uvažovania pohlavia ako faktora úmrtnosti s hodnotou AIC rovnou 143,7. Pri ARIMA modelovaní v závislosti od pohlavia zvlášť pre mužov a ženy sme dospeli k podobným výsledkom. Pre mužov bol najvhodnejší model takisto *ARIMA (1,1,0)* s AIC rovným 142,6. U žien bol odhad ťažší, ale vedúci model bol rovnako *ARIMA (1,1,0)* so zodpovedajúcim $AIC = 144,4$. Na predpovedanie budúcich hodnôt časovo závislého parametra sme použili najvhodnejšie ARIMA modely. V tabuľke č. 10 sú uvedené prognózované hodnoty indexu úmrtnosti $\kappa_t^{(2)}$ spolu s jeho 95 % intervalmi spoľahlivosti na najbližších 10 rokov.

Na základe tabuľky č. 10, očakávame, že do roku 2022 sa bude situácia zhoršovať, a že úmrtnosť populácie sa zvýši. Následne predpokladáme, že dôjde k poklesu úmrtnosti obyvateľstva Slovenska. Vyplýva to z klesajúceho trendu parametra $\kappa_t^{(2)}$. Počiatočný odhadovaný nárast úmrtnosti je spojený so zvýšenou úmrtnosťou počas posledného roku projekcie (2020), ktorý vysokou mierou zasiahla pandémia koronavírusu. Keďže ARIMA modely dávajú najvyššiu váhu práve najnovším pozorovaniam, malo to za následok nárast parametra $\kappa_t^{(2)}$ v prvých dvoch rokoch prognózy.

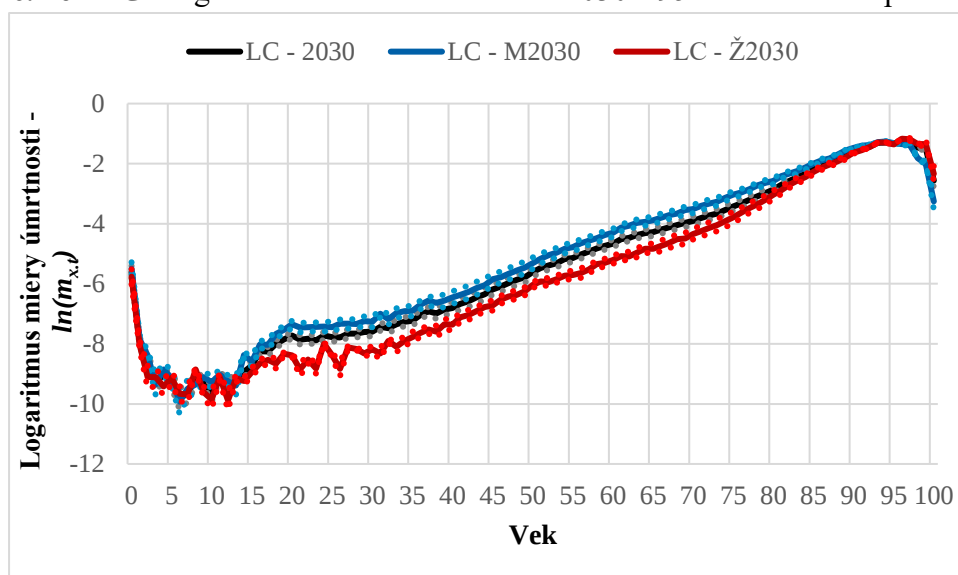
Obrázok č. 22 - Prognóza $\kappa_t^{(2)}$ LC modelu pri zohľadnení a vynechaní roku 2020



Zdroj: výstup zo SAS Enterprise Guide

ARIMA modelovanie sme skúsili zopakovať s vynechaním pandemického roku z analýzy. Grafický výstup prognóz zo softvéru SAS Enterprise Guide je zobrazený na obrázku č. 22. V prípade, že by sme vynechali pandemický rok, už v roku 2022 by sme boli na rovnakej očakávanej úmrtnosti ako v roku 2026 pri zohľadnení pandémie počas modelovania. Vzhľadom na aktuálnu situáciu vo svete spojenú s koronavírusom, považujeme za správne zohľadniť túto situáciu pri modelovaní úmrtnosti.

Graf č. 40 - LC Prognóza mier úmrtnosti na rok 2030 s 95 % intervalmi spoľahlivosti



Zdroj: vlastné spracovanie

Na predpovedanie hodnôt úmrtnosti podľa veku sme použili odhadnuté parametre $\beta_x^{(1)}$, $\beta_x^{(2)}$ a prognózované hodnoty indexu úmrtnosti $\kappa_t^{(2)}$. Na základe prognózy môžeme povedať, že v najbližších rokoch očakávame nárast úmrtnosti z dôvodu stále pretrvávajúcej pandémie koronavírusu. Tento stav by sa však mal neskôr ustáliť a úmrtnosť by mala mať naďalej klesajúci trend, ako to bolo do roku 2020. Na grafe č. 40 môžeme vidieť očakávané hodnoty logaritmu vekovo špecifických mier úmrtnosti v roku 2030 spolu s 95 % intervalom spoľahlivosti pre mužov a ženy samostatne, ako aj spolu. Časovo závislý parameter $\kappa_t^{(2)}$ môžeme odhadnúť aj na viac ako 10 rokov. Dĺžka prognózy závisí od požiadaviek poisťovne. Odporúčame však robiť predpoveď na kratšie obdobie. Jednak z toho dôvodu, že získame presnejšie a spoľahlivejšie odhady a tiež predpokladáme, že životné poisťovne prepočítavajú svoje predpoklady úmrtnosti minimálne raz za niekoľko rokov, čo sa odporúča

aj pri stochastickom Lee-Carterovom modeli z dôvodu možnosti zmeny štruktúry úmrtnosti populácie, ako to bolo aj medzi rokmi 1993 a 2020.

4.2.2 Cairns-Blake-Dowdov model úmrtnosti

Cairns-Blake-Dowdov (CBD) model sa riadi vzťahom (3.11). Vychádza z logistickej transformácie miery úmrtnosti, pričom využíva takmer lineárnosť krivky vo vyšších vekoch. Vzhľadom na tento fakt, nie je vhodný na modelovanie celého vekového intervalu. V prípade slovenskej populácie sme spozorovali, že *logit* krivka miery úmrtnosti začína mať lineárny trend vo veku od 30 rokov. Našu analýzu sme preto zamerali na vekový interval $\langle 30, 100 \rangle$, z čoho vyplýva, že priemerný vek na tomto vekovom intervale je $\bar{x} = 65$. Už na základe zvoleného intervalu vieme povedať, že CBD model úmrtnosti nie je vhodný na modelovanie dojčenskej úmrtnosti a ani havarijného hrbu. Využitie môže nájsť napríklad pri modelovaní penzijných schém. CBD model obsahuje dva parametre, pričom obidva sú závislé od času. Parametre modelu sme odhadli pomocou metódy najmenších štvorcov opísanej v kapitole 3.3.1 a na prognózovanie ďalšieho vývoja sme použili ARIMA modely.

Tabuľka č. 11 - Odhad parametra $\kappa_t^{(1)}$ CBD modelu

Rok	$\kappa_t^{(1)}$	$\kappa_t^{(1)}$ Muži	$\kappa_t^{(1)}$ Ženy	Rok	$\kappa_t^{(1)}$	$\kappa_t^{(1)}$ Muži	$\kappa_t^{(1)}$ Ženy
1993	-3,746	-3,476	-4,128	2007	-3,868	-3,587	-4,244
1994	-3,699	-3,429	-4,077	2008	-3,900	-3,618	-4,300
1995	-3,700	-3,422	-4,074	2009	-3,944	-3,669	-4,314
1996	-3,766	-3,522	-4,128	2010	-3,968	-3,688	-4,354
1997	-3,779	-3,522	-4,144	2011	-3,957	-3,691	-4,331
1998	-3,777	-3,514	-4,159	2012	-3,985	-3,717	-4,354
1999	-3,813	-3,538	-4,208	2013	-4,022	-3,766	-4,368
2000	-3,848	-3,577	-4,256	2014	-4,068	-3,791	-4,436
2001	-3,774	-3,495	-4,165	2015	-4,044	-3,768	-4,429
2002	-3,805	-3,515	-4,215	2016	-4,093	-3,862	-4,418
2003	-3,813	-3,533	-4,207	2017	-4,100	-3,846	-4,447
2004	-3,855	-3,578	-4,249	2018	-4,114	-3,858	-4,486
2005	-3,835	-3,565	-4,215	2019	-4,149	-3,907	-4,490
2006	-3,858	-3,564	-4,259	2020	-4,094	-3,838	-4,444

Zdroj: vlastné spracovanie

Tabuľky č. 11 a 12 pozostávajú z odhadnutých parametrov Cairns-Blake-Dowdovho modelu pre populáciu Slovenskej republiky za časové obdobie 1993 – 2020. V oboch tabuľkách sú odhadnuté indexy úmrtnosti zvlášť pre mužov a ženy, ako aj po nezohľadnení pohlavia ako faktora vplývajúceho na úmrtnosť.

Tabuľka č. 12 - Odhad parametra $\kappa_t^{(2)}$ CBD modelu

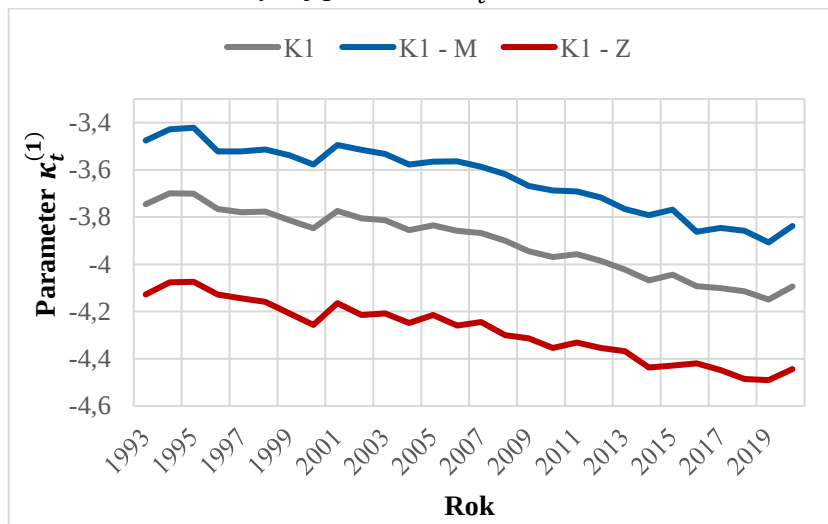
Rok	$\kappa_t^{(2)}$	$\kappa_t^{(2)}$ Muži	$\kappa_t^{(2)}$ Ženy	Rok	$\kappa_t^{(2)}$	$\kappa_t^{(2)}$ Muži	$\kappa_t^{(2)}$ Ženy
1993	0,0844	0,0766	0,0971	2007	0,0889	0,0823	0,1006
1994	0,0887	0,0810	0,1017	2008	0,0899	0,0821	0,1035
1995	0,0894	0,0828	0,1013	2009	0,0906	0,0839	0,1022
1996	0,0883	0,0789	0,1007	2010	0,0913	0,0847	0,1036
1997	0,0848	0,0764	0,0975	2011	0,0948	0,0876	0,1065
1998	0,0834	0,0747	0,0968	2012	0,0961	0,0893	0,1077
1999	0,0843	0,0761	0,0983	2013	0,0953	0,0889	0,1052
2000	0,0843	0,0750	0,0994	2014	0,0927	0,0874	0,1030
2001	0,0899	0,0820	0,1032	2015	0,0940	0,0875	0,1066
2002	0,0900	0,0823	0,1043	2016	0,0932	0,0862	0,1024
2003	0,0900	0,0821	0,1034	2017	0,0938	0,0876	0,1042
2004	0,0904	0,0826	0,1043	2018	0,0917	0,0843	0,1042
2005	0,0903	0,0822	0,1030	2019	0,0910	0,0841	0,1011
2006	0,0902	0,0835	0,1034	2020	0,0923	0,0861	0,1027

Zdroj: vlastné spracovanie

Z dôvodu lepšej prehľadnosti sú odhadované parametre Cairns-Blake-Dowdovho modelu za obdobie od roku 1993 po 2020 znázornené na grafoch č. 41 a 42. Úrovňová konštanta modelu $\kappa_t^{(1)}$ reprezentuje úroveň úmrtnosti v čase. Ako je znázornené na grafe č. 41, odhadovaný parameter $\kappa_t^{(1)}$ z roka na rok klesá, čo potvrdzuje myšlienku, že celková úroveň úmrtnosti sa v sledovanom období medzi rokmi 1993 a 2020 zlepšila. V roku 2000 vidíme na grafe č. 41 parametra $\kappa_t^{(1)}$ ostrý skok, čo môže súvisieť aj s tým, že v tomto roku sa nám podľa grafu č. 44 najmenej podarilo priblížiť ku skutočným mieram úmrtnosti. V poslednom roku môžeme vidieť prudší nárast tohto parametra, čo je spojené s pretrvávajúcou pandémiou koronavírusu. Na druhej strane však index úmrtnosti $\kappa_t^{(2)}$ má mierne rastúci trend. Nárast parametra je viditeľný viac u mužskej populácie, zatiaľ čo u ženskej populácie je tento parameter viac-menej stacionárny a môžeme ho považovať za stabilný počas celého analyzovaného obdobia. Zmena hodnôt parametra $\kappa_t^{(2)}$ spôsobuje

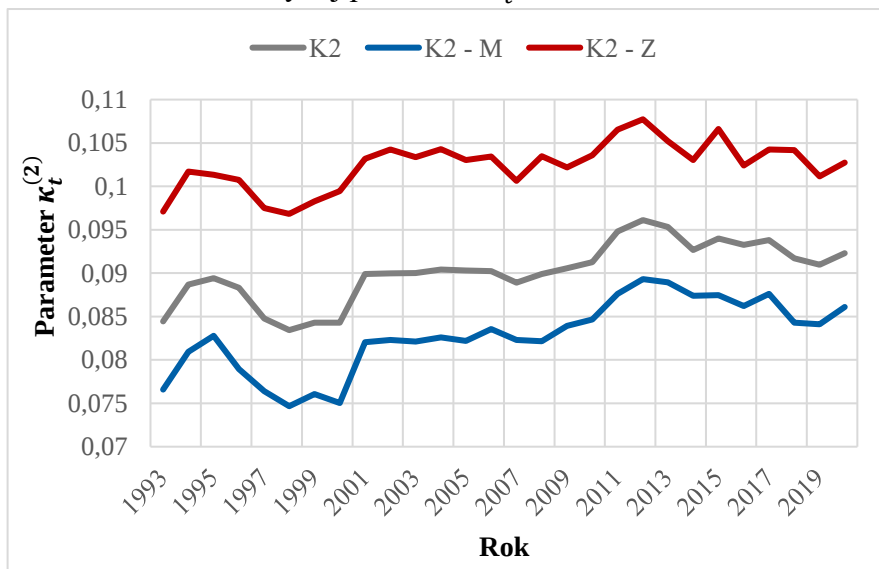
zmenu v strmosti logisticky transformovanej krivky úmrtnosti, čo znamená, že v prípade mužskej populácie sa úmrtnosť v mladšom veku (do priemerného veku 65 rokov) zlepšovala rýchlejšie ako vo vyšších vekoch nad 65 rokov. Naopak, v prípade populácie žien sa úmrtnosť zlepšovala rovnomerne na celom vekovom intervale.

Graf č. 41 - Vývoj parametra $\kappa_t^{(1)}$ CBD modelu v čase



Zdroj: vlastné spracovanie

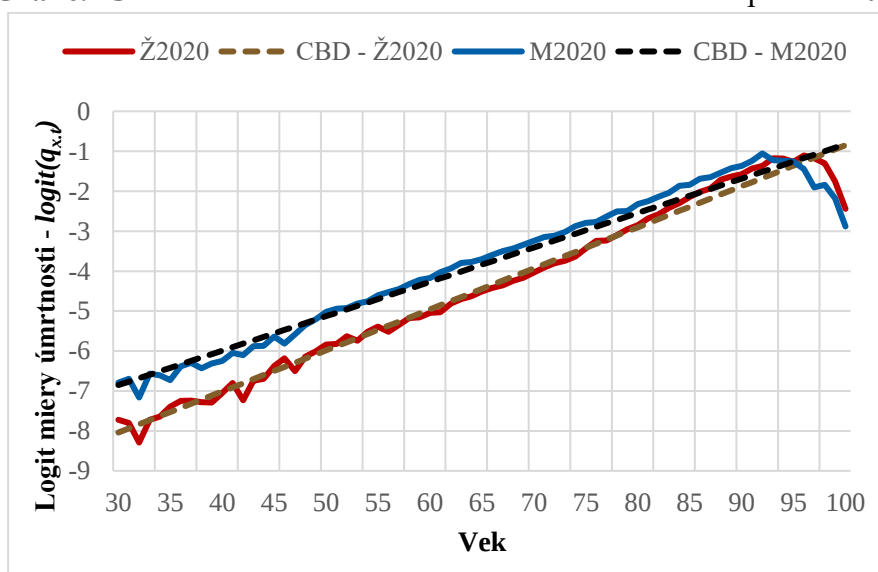
Graf č. 42 - Vývoj parametra $\kappa_t^{(2)}$ CBD modelu v čase



Zdroj: vlastné spracovanie

Odhadované hodnoty parametrov $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$ z tabuliek č. 11 a 12 sme dosadili do rovnice CBD modelu, aby sme získali očakávané hodnoty $\text{logit}(q_{x,t})$. Očakávané hodnoty $\text{logit}(q_{x,t})$ slúžia na porovnanie so skutočnými pozorovanými hodnotami miery úmrtnosti v logistickom tvare a na posúdenie vhodnosti modelu. Na grafe č. 43 sme znázornili pozorované pravdepodobnosti úmrtia mužov (M2020) a žien (Ž2020) vo veku od 30 do 100 rokov prevedené do logistickej podoby pre rok 2020. Prerušovanými čiarami sú na grafe č. 43 vykreslené hodnoty $\text{logit}(q_{x,t})$ odhadnuté pomocou Cairns-Blake-Dowdovho modelu. Vidíme, že očakávané a skutočné hodnoty sú veľmi podobné. Keďže miery úmrtnosti v logistickom tvare nie sú už lineárne v prípade ľudí starších ako 95 rokov, dochádza k tomu, že vyrovnané krivky sú menej strmé, a teda pri mužskej populácii je nadhodnotená úmrtnosť do 50 roku života a u oboch pohlaví dochádza k podhodnoteniu úmrtnosti od približne 80 roku života. Miera úmrtnosti je nadhodnotená aj u osôb starších ako 95 rokov.

Graf č. 43 - Ukážka CBD modelu úmrtnosti mužov a žien pre rok 2020

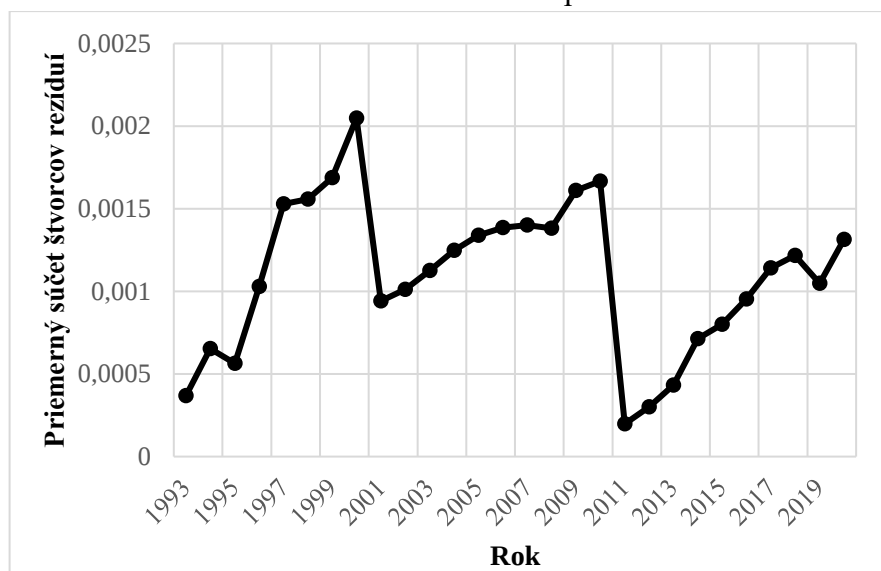


Zdroj: vlastné spracovanie

Na grafe č. 44 je zobrazený vývoj priemerného súčtu rezíduí MSE po vyrovnaní pomocou CBD modelu. Hodnoty MSE sú veľmi nízke, pričom závisia od lineárnosti logisticky transformovanej miery úmrtnosti vo vysokých vekoch. Je dôležité pripomenúť, že v nich nie je zohľadnený celý vekový interval, ale iba vek od 30 do 100 rokov. Najmenej spoľahlivo sme vyrovnali krivku pravdepodobnosti úmrtia v logistickom tvare pomocou

Cairns-Blake-Dowdovho modelu v rokoch 2000 a 2010. Na druhej strane, najnižšie hodnoty MSE sme dosiahli v rokoch 2011 a 2012.

Graf č. 43 - MSE CBD modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020



Zdroj: vlastné spracovanie

Model CBD pozostáva z dvoch časovo závislých parametrov, ktorých vývoj je potrebné predpovedať. V prognostických modeloch musí byť dodržaný implicitný predpoklad, že trend v budúcnosti nezmení smer. Rovnako ako pri predchádzajúcom Lee-Carterovom modeli sme zobrali do úvahy niekoľko možných ARIMA modelov, pričom na základe Akaikovho informačného kritéria sme vybrali tie, ktoré boli najvhodnejšie na prognózovanie indexov úmrtnosti $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$.

Začneme parametrom $\kappa_t^{(1)}$. Na základe $AIC = -108,7$ sa vie najlepšie prispôbiť doterajšiemu vývoju parametra, ktorý predstavuje úrovňovú konštantu, pre populáciu Slovenska bez zohľadnenia pohlavia model $ARIMA(1,1,0)$. Pre mužskú populáciu za najlepšie vyhovujúci model považujeme takisto $ARIMA(1,1,0)$ s hodnotou AIC zodpovedajúcou $-97,6$. V prípade žien sme dospeli k rozdielnemu výsledku. Najlepšie spĺňal požiadavky model $ARIMA(2,2,0)$ s AIC rovným $-98,8$.

Parameter $\kappa_t^{(2)}$ reprezentuje sklon CBD modelu. Pri prognózovaní bolo náročnejšie odhadnúť predpokladaný vývoj tohto parametra do budúcnosti. Ak neberieme do úvahy pohlavie, dospeli sme k záveru, že ďalší vývoj vieme najlepšie zachytiť pomocou modelu $ARIMA(0,1,3)$, pričom hodnota kritéria AIC predstavovala $-263,1$. Rovnako tomu bolo aj

pre mužskú populáciu avšak s odlišnou hodnotou AIC rovnajúcou sa $-251,6$. Napokon pre populáciu žien nám vyšiel za najvhodnejší model na predpovedanie parametra $\kappa_t^{(2)}$ $ARIMA(0,2,2)$ s $AIC = -235,1$. V tabuľkách č. 13 a 14 sú uvedené prognózované hodnoty indexov úmrtnosti $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$ Cairns-Blake-Dowdovho modelu spolu 95 % intervalmi spoľahlivosti na najbližších 10 rokov.

Tabuľka č. 13 - Prognóza parametra $\kappa_t^{(1)}$ CBD modelu s 95 % intervalom spoľahlivosti

Rok	Prognóza $\kappa_t^{(1)}$	95 % Interval spoľahlivosti	
2021	-4,142	-4,203	-4,082
2022	-4,158	-4,224	-4,092
2023	-4,173	-4,244	-4,102
2024	-4,188	-4,264	-4,113
2025	-4,204	-4,284	-4,123
2026	-4,219	-4,303	-4,135
2027	-4,234	-4,322	-4,146
2028	-4,249	-4,342	-4,157
2029	-4,265	-4,360	-4,169
2030	-4,280	-4,379	-4,181
Rok	Prognóza $\kappa_t^{(1)}$ Muži	95 % Interval spoľahlivosti	
2021	-3,890	-3,964	-3,815
2022	-3,906	-3,988	-3,823
2023	-3,921	-4,011	-3,832
2024	-3,937	-4,034	-3,841
2025	-3,953	-4,056	-3,850
2026	-3,969	-4,078	-3,860
2027	-3,985	-4,099	-3,870
2028	-4,001	-4,120	-3,881
2029	-4,016	-4,141	-3,892
2030	-4,032	-4,162	-3,903
Rok	Prognóza $\kappa_t^{(1)}$ Ženy	95 % Interval spoľahlivosti	
2021	-4,512	-4,579	-4,445
2022	-4,519	-4,586	-4,452
2023	-4,547	-4,618	-4,475
2024	-4,521	-4,592	-4,449
2025	-4,573	-4,662	-4,485
2026	-4,568	-4,656	-4,480
2027	-4,605	-4,700	-4,511
2028	-4,587	-4,682	-4,492
2029	-4,634	-4,739	-4,529
2030	-4,623	-4,728	-4,519

Zdroj: vlastné spracovanie na základe výstupu zo SAS Enterprise Guide

Tabuľka č. 14 - Prognóza parametra $\kappa_t^{(2)}$ CBD modelu s 95 % intervalom spoľahlivosti

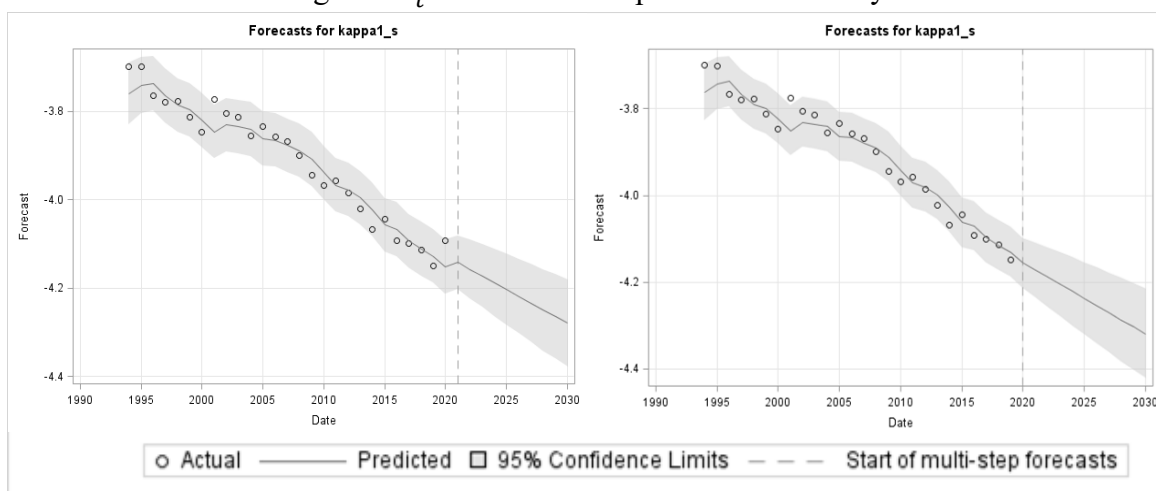
Rok	Prognóza $\kappa_t^{(2)}$	95 % Interval spoľahlivosti	
2021	0,0937	0,0903	0,0970
2022	0,0956	0,0909	0,1004
2023	0,0953	0,0895	0,1011
2024	0,0956	0,0898	0,1015
2025	0,0960	0,0900	0,1019
2026	0,0963	0,0903	0,1023
2027	0,0966	0,0906	0,1027
2028	0,0969	0,0908	0,1031
2029	0,0973	0,0911	0,1034
2030	0,0976	0,0914	0,1038
Rok	Prognóza $\kappa_t^{(2)}$ Muži	95 % Interval spoľahlivosti	
2021	0,0882	0,0839	0,0924
2022	0,0893	0,0833	0,0954
2023	0,0888	0,0814	0,0962
2024	0,0892	0,0815	0,0970
2025	0,0896	0,0816	0,0977
2026	0,0901	0,0817	0,0984
2027	0,0905	0,0819	0,0990
2028	0,0909	0,0820	0,0997
2029	0,0913	0,0821	0,1004
2030	0,0917	0,0823	0,1010
Rok	Prognóza $\kappa_t^{(2)}$ Ženy	95 % Interval spoľahlivosti	
2021	0,1033	0,0984	0,1082
2022	0,1043	0,0994	0,1092
2023	0,1050	0,0989	0,1111
2024	0,1059	0,0998	0,1120
2025	0,1054	0,0992	0,1116
2026	0,1063	0,1001	0,1125
2027	0,1058	0,0996	0,1120
2028	0,1067	0,1005	0,1129
2029	0,1062	0,0999	0,1125
2030	0,1071	0,1008	0,1134

Zdroj: vlastné spracovanie na základe výstupu zo SAS Enterprise Guide

Na základe tabuľky č. 13 môžeme v nasledujúcich rokoch očakávať pokles úmrtnosti obyvateľstva Slovenskej republiky. Toto tvrdenie je založené na klesajúcom trende parametra $\kappa_t^{(1)}$. Druhý parameter $\kappa_t^{(2)}$ má však naopak rastúci trend, preto predpokladáme, že v budúcnosti bude úmrtnosť v mladšom veku klesať rýchlejšie ako vo vyššom veku. Ak by sme vynechali rok 2020 z prognózy, index úmrtnosti $\kappa_t^{(1)}$ by klesal rovnakým tempom, avšak tento pokles začína v inom bode, preto sa prognóza javí akoby o dva roky posunutá. Hodnotu parametra, ktorú by sme očakávali v roku 2022 v prípade, že nezohľadníme

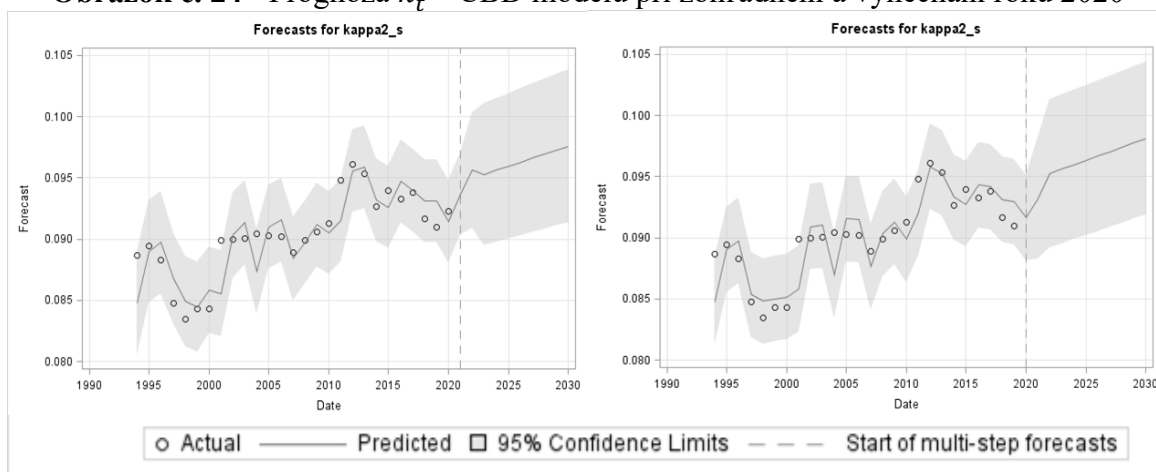
pandemický rok, dosiahneme až v roku 2024 pri zohľadnení tohto roku. Vývoj parametra $\kappa_t^{(2)}$ sa pri ignorovaní posledného roku veľmi nezmenil. Keďže koronavírus najviac zasiahol vekovú skupinu nad 65 rokov, stále platí, že očakávame rýchlejšie zlepšovanie úmrtnosti u mladších ľudí. Grafický výstup prognóz parametrov $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$ zo softvéru SAS Enterprise Guide pri zohľadnení a vynechaní pandemického roku je zobrazovaný na obrázkoch č. 23 a 24. Podľa nášho názoru je správne zobrať do úvahy aktuálnu situáciu pri modelovaní úmrtnosti.

Obrázok č. 23 - Prognóza $\kappa_t^{(1)}$ CBD modelu pri zohľadnení a vynechaní roku 2020



Zdroj: výstup zo SAS Enterprise Guide

Obrázok č. 24 - Prognóza $\kappa_t^{(2)}$ CBD modelu pri zohľadnení a vynechaní roku 2020

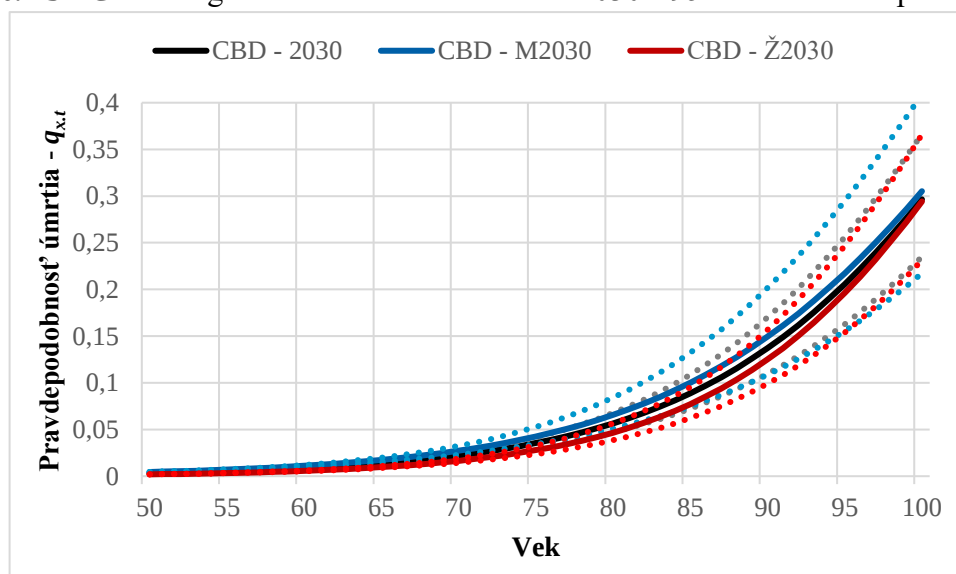


Zdroj: výstup zo SAS Enterprise Guide

Na predpovedanie pravdepodobnosti úmrtia podľa pohlavia a veku sme použili odhadnuté indexy úmrtnosti CBD modelu $\kappa_t^{(1)}$ a $\kappa_t^{(2)}$. Z prognózy vyplýva, že

v nasledujúcich rokoch očakávame pokles úmrtnosti, ktorý však bude v dôsledku pandémie koronavírusu pomalší. Na grafe č. 45 sú znázornené miery úmrtnosti mužov, žien, ako aj spolu v roku 2030 s 95 % intervalmi spoľahlivosti. Viditeľné je približovanie miery úmrtnosti mužov a žien, ktoré predpokladáme už na základe očakávanej dĺžky života pri narodení, ktorej sme sa venovali v kapitole 1.2.1. Za nevýhodu tohto modelu, v porovnaní s Lee-Carterovým modelom, považujeme, že intervaly spoľahlivosti pre výsledné miery úmrtnosti sú pomerne široké. Príliš široké intervaly spoľahlivosti nie sú pre praktické účely použiteľné a naznačujú, že odhad je nepresný. V našom prípade to môže byť spôsobené pravdepodobne parametrom $\kappa_t^{(2)}$, ktorého prognóza bola náročná a vykazovala väčšiu štandardnú odchýlku. Na šírku intervalu spoľahlivosti vplýva aj rozsah vstupných údajov, preto ak by sme chceli spresniť prognózu, museli by sme analyzovať dlhšie obdobie ako od roku 1993 po 2020, prípadne by sme museli skrátiť dĺžku prognózovanej doby.

Graf č. 45 - CBD Prognóza mier úmrtnosti na rok 2030 s 95 % intervalmi spoľahlivosti



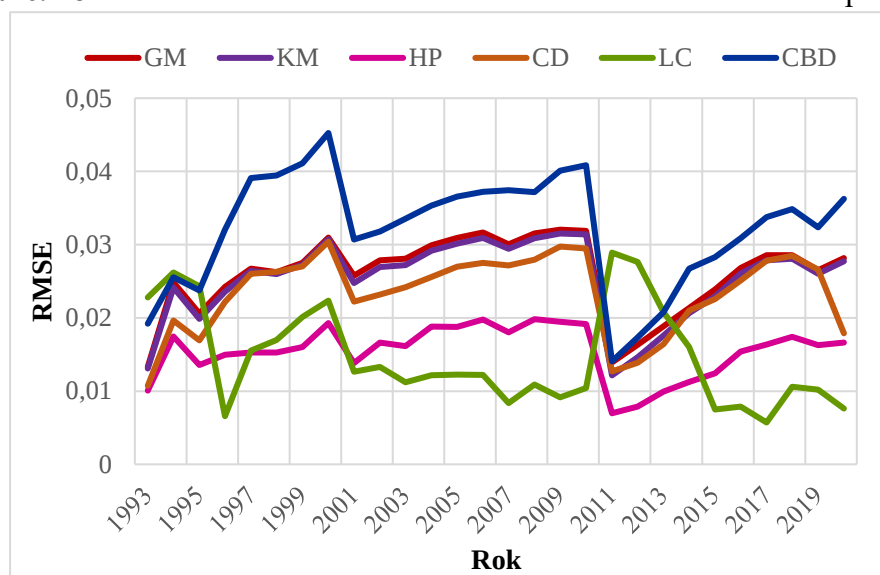
Zdroj: vlastné spracovanie

4.3 Porovnanie modelov úmrtnosti na základe RMSE

Jednou z najdôležitejších častí modelovania úmrtnosti je samotné vyhodnotenie a porovnanie modelov. Predmetom analýzy sú štyri deterministické a dva stochastické modely opísané v časti 3.1. Keďže poisťovňa sa snaží čo najpresnejšie odhadnúť úmrtnosť

svojich poistencov v závislosti od veku, snaží sa vybrať taký model, ktorý je schopný čo najlepšie opísať úmrtnostné správanie populácie. Existuje viacero ukazovateľov, pomocou ktorých je možné posúdiť kvalitu modelu. Pri našej analýze použijeme ukazovateľ RMSE (Štandardná odchýlka rezíduí), ktorý predstavuje druhú odmocninu MSE (Priemerný súčet štvorcov rezíduí), pričom nižšie hodnoty tohto indikátora znamenajú lepšiu schopnosť modelu prispôbiť sa pozorovaným hodnotám ukazovateľov úmrtnosti populácie Slovenska. Viac informácií o týchto ukazovateľoch nájdeme v kapitole 3.3.

Graf č. 46 - Porovnanie modelov na základe RMSE bez zohľadnenia pohlavia



Zdroj: vlastné spracovanie

Na grafe č. 46 sú zobrazené hodnoty ukazovateľa RMSE v jednotlivých rokoch pozorovaného obdobia pri zanedbaní pohlavia ako faktora vplývajúceho na úroveň úmrtnosti populácie Slovenskej republiky. Na základe výsledkov môžeme za modely, ktoré najviac vyhovujú dátam o úmrtnostnom správaní populácie Slovenska, považovať Heligmanov-Pollardov deterministický model úmrtnosti a Lee-Carterov stochastický model úmrtnosti. Počas sledovaného obdobia s rozsahom 28 rokov, až 17-krát z nich dosiahol najnižšie hodnoty RMSE práve Lee-Carterov model, ktorý sa veľmi často používa aj v praxi a zvyšných 11 rokov dosiahol najlepšie výsledky spomínaný Heligmanov-Pollardov zákon úmrtnosti. Veľkou výhodou týchto modelov je, že sú vhodné pre celé vekové rozpätie a vedia sa prispôbiť dojčenskej úmrtnosti, hoci pri Heligmanovom-Pollardovom modeli sa nám to nie vždy podarilo, ďalej havarijnému hrbu, ale aj úmrtnosti vo vysokých vekoch.

Za nevýhodu Heligmanovho-Pollardovho zákona úmrtnosti považujeme jeho zložitosť spojenú s pomerne veľkým počtom parametrov. Naopak Lee-Carterov model sa skladá len z troch parametrov a je relatívne jednoduchý, čo pravdepodobne súvisí s tým, prečo je tak populárny.

Na základe grafu č. 46 môžeme povedať, že priemerné výsledky dosahoval Gompertzov-Makehamov zákon úmrtnosti, Kannistov-Makehamov zákon úmrtnosti a takisto CoDe model. Hlavne CoDe model bol vzhľadom na svoju komplexnosť veľkým sklamaním najmä vo vysokých vekoch, kde sme očakávali, že bude pomerne presne opisovať krivku miery úmrtnosti. Vyzdvihnúť môžeme najmä jeho schopnosť modelovať dojčenskú úmrtnosť, s ktorou mala väčšina modelov problém.

Napokon za najmenej vyhovujúci považujeme Cairns-Blake-Dowdov model. Je to veľmi jednoduchý model, ktorý využíva takmer lineárnosť krivky miery úmrtnosti v logistickom tvare vo vyšších vekoch. Nie je teda vhodný na modelovanie celého vekového intervalu, čo predstavuje jeho najväčšiu nevýhodu. Hoci autori modelu tvrdia, že riziko úmrtnosti je spojené najmä s vyšším vekom, je pravda, že miery úmrtnosti v logistickom tvare nie sú už lineárne ani u ľudí starších ako 95 rokov, a teda jeho použitie je vhodné prakticky len na intervale medzi 30 a 95 rokov života. Keďže životné poisťovne majú aj mladších klientov, tento model nepovažujeme za použiteľný.

Tabuľka č. 15 - Hodnoty RMSE modelov úmrtnosti pre ženy

Ženy	1993	2002	2011	2020
RMSE GM	0,019976	0,048651	0,020433	0,052659
RMSE KM	0,019558	0,038423	0,014565	0,044433
RMSE HP	0,015516	0,01533	0,007266	0,01598
RMSE CD	0,015036	0,022745	0,010715	0,018611
RMSE LC	0,027106	0,01442	0,02941	0,00824
RMSE CBD	0,022367	0,031294	0,015445	0,034887

Zdroj: vlastné spracovanie

V závislosti od pohlavia sme vzhľadom na výsledky analyzovali v dizertačnej práci už len vybrané roky 1993, 2002, 2011 a 2020. Závery porovnania modelov boli totižto podobné ako v prípade zanedbania pohlavia pri analýze. Hodnoty RMSE jednotlivých modelov v konkrétnych rokoch v závislosti od pohlavia sú súčasťou tabuliek č. 15 a 16. Na základe tabuľky č. 15 vidíme, že u ženskej populácie dosahovali najlepšie výsledky takisto najmä Heligmanov-Pollardov model a Lee-Carterov model, pričom v roku 1993 vynikol aj CoDe

model, ktorý sa držal minimálne v prvej trojke najlepších modelov v každom z týchto rokov. Najhoršie dopadol v prípade žien Gompertzov-Makehamov model a Cairns-Blake-Dowdov model. U mužskej populácie analýza viedla k prakticky rovnakým záverom. Vo všeobecnosti sú však hodnoty RMSE v prípade mužov vyššie, čo je spôsobené najmä tým, že u mužskej populácie je viac viditeľný havarijný hrb a tiež preto, že vo vysokých vekoch dochádza k vyššej volatilité údajov.

Tabuľka č. 16 - Hodnoty RMSE modelov úmrtnosti pre mužov

Muži	1993	2002	2011	2020
RMSE GM	0,028813	0,066656	0,028263	0,057603
RMSE KM	0,028225	0,057888	0,026844	0,05394
RMSE HP	0,019993	0,021614	0,017844	0,020203
RMSE CD	0,023155	0,040423	0,025525	0,022767
RMSE LC	0,028835	0,017680	0,035193	0,012049
RMSE CBD	0,034737	0,04296	0,03086	0,046976

Zdroj: vlastné spracovanie

Štatistický úrad Slovenskej republiky (ŠÚ SR) každoročne zverejňuje úmrtnostné tabuľky slovenskej populácie. Na vyrovnanie nepravidelných výkyvov na krivke vypočítanej vekovo špecifickej miery úmrtnosti používa ŠÚ SR metódu kľzavých priemerov, ktorá je vhodná na odstránenie malých výkyvov, teda pre mladšie veku. U vyšších vekoch je to iné, tu metóda kľzavých priemerov nedokáže vyrovnať miery úmrtnosti tak, aby odstránila výkyvy spôsobené veľkými štatistickými odchýlkami vo vstupných dátach vyplývajúcich hlavne z nerovnomerného rozdelenia malého počtu zomrelých v starších vekoch populácie. V praxi sa uvedené veľké odchýlky vypočítaných mier úmrtnosti vo vyšších vekoch riešia extrapoláciou, a teda empirické hodnoty sa nahradia modelovými hodnotami. ŠÚ SR pri výpočte úmrtnostných tabuliek používa Gompertzov-Makehamov model. Mészáros (2016) v spolupráci s INFOSTAT vypracoval materiál, v ktorom testoval použiteľnosť viacerých deterministických modelov a dal ŠÚ SR odporúčanie na spracovanie úmrtnostných tabuliek v ďalších rokoch. Vo všeobecnosti dospel k záveru, že väčšina modelov je vyhovujúca, avšak navrhol nemeniť metodiku výpočtu úmrtnostných tabuliek, aby sa tým zabezpečila konzistentnosť a porovnateľnosť vypočítaných hodnôt s minulosťou. Pre porovnanie, napríklad Štatistický úrad Českej republiky aplikoval do roku 2017 rovnakú metodiku. V roku 2018 sa namiesto

Gompertzovho-Makehamovho modelu začal používať v úmrtnostných tabuľkách pre modelovanie úmrtnosti v najvyšších vekoch Kannistov-Makehamov model.

Čo sa týka využitia modelov úmrtnosti v oblasti poisťovníctva, životné poisťovne často zužitkujú priamo výsledky ŠÚ SR, ktoré upravujú tak, aby vyhovovali úmrtnostnému správaniu ich poisťného kmeňa. V prípade životných poisťovní, ktoré vytvárajú vlastné úmrtnostné modely, zvyčajne aplikujú podobné postupy ako ŠÚ SR. Stále viac ale sa v poisťnej praxi stáva populárnejší práve Lee-Carterov stochastický model úmrtnosti, ktorý pri našej analýze (graf č. 46) dosahoval najlepšie výsledky.

4.4 Analýza vplyvov na stanovenie poisťných sadzieb

V predchádzajúcej časti sme popísali niekoľko deterministických a stochastických modelov úmrtnosti, ktoré tvoria len malý výber zo všetkých existujúcich modelov. Snažili sme sa ale vybrať vzorku modelov tak, aby obsahovala prierez od najstarších po najnovšie a od tých najjednoduchších po najzložitejšie. V tejto časti sa preto zameriame na to, aký vplyv má výber modelu na poisťné sadzby. Okrem toho sa pozrieme aj na ďalšie vonkajšie vplyvy, akými sú zavedenie gender smernice a pandémie koronavírusu. Pri analýze budeme vychádzať zo základných a zároveň najviac kupovaných poisťných produktov, ktoré sme predstavili v časti 3.5. Na ich výpočet potrebujeme poznať ročné pravdepodobnosti úmrtia a prežitia v konkrétnom veku x . Ročná pravdepodobnosť úmrtia osoby vo veku x je priamo výstupom niektorých modelov úmrtnosti, pri ostatných sme použili predpoklad o konštantnej miere úmrtnosti medzi celočíselnými vekmi, podľa ktorého platí $q_{x,t} = 1 - e^{-\mu_{x,t}}$ a $\mu_{x,t} = m_{x,t}$, prípadne sme ich transformovali z logaritmickeho tvaru. Pravdepodobnosť ročnej miery dožitia osoby vo veku x vieme potom vyjadriť pomocou miery úmrtia ako $p_{x,t} = 1 - q_{x,t}$. Popri pravdepodobnostiach, ktoré bývajú súčasťou úmrtnostných tabuliek, potrebujeme na určenie poisťných sadzieb tiež poisťno-technickú úrokovú mieru, ktorá je platná v čase uzavretia zmluvy. My sme pri výpočtoch pre jednoduchosť predpokladali úrokovú mieru na úrovni 1 %.

4.4.1 Analýza vplyvu výberu modelu

Aby sme zistili, aký vplyv má výber modelu úmrtnosti na výšku poistných sadzieb, zobrali sme jednorôčné pravdepodobnosti úmrtia, ktoré sme dostali za použitia jednotlivých modelov v poslednom roku sledovaného obdobia, teda v roku 2020. Na základe nich sme potom vypočítali poistné sadzby pre dočasné poistenie na úmrtie a dočasný predlehotný dôchodok pre rôzne veku od 10 do 80 rokov s dobou trvania poistenia do 90-teho roku života osoby. Pri stochastických modeloch sme brali do úvahy aj prognózu vývoja parametrov daných modelov do budúcnosti. Najmä vďaka tomu, že stochastické modely zohľadňujú zlepšovanie úmrtnostného správania populácie, platilo, že čím bola poistná doba dlhšia, tým viac sa poistné sadzby medzi sebou odlišovali.

Ako príklad uvidíme dva prípady, konkrétne to je poistenie pre 30 ročnú osobu na dobu trvania 60 rokov a pre 70 ročnú osobu s poistnou dobou 20 rokov. Vypočítané súčasné hodnoty týchto dvoch poistení, čiže hodnoty jednorazového netto poistného zaplateného v prípade jednorazovej výplaty poistnej sumy pri úmrtí, resp. výplaty dôchodku počas života poistenej osoby vo výške 1 p. j., sú zaznamenané v tabuľke č. 17.

Tabuľka č. 17 - Vplyv výberu modelu na poistné sadzby

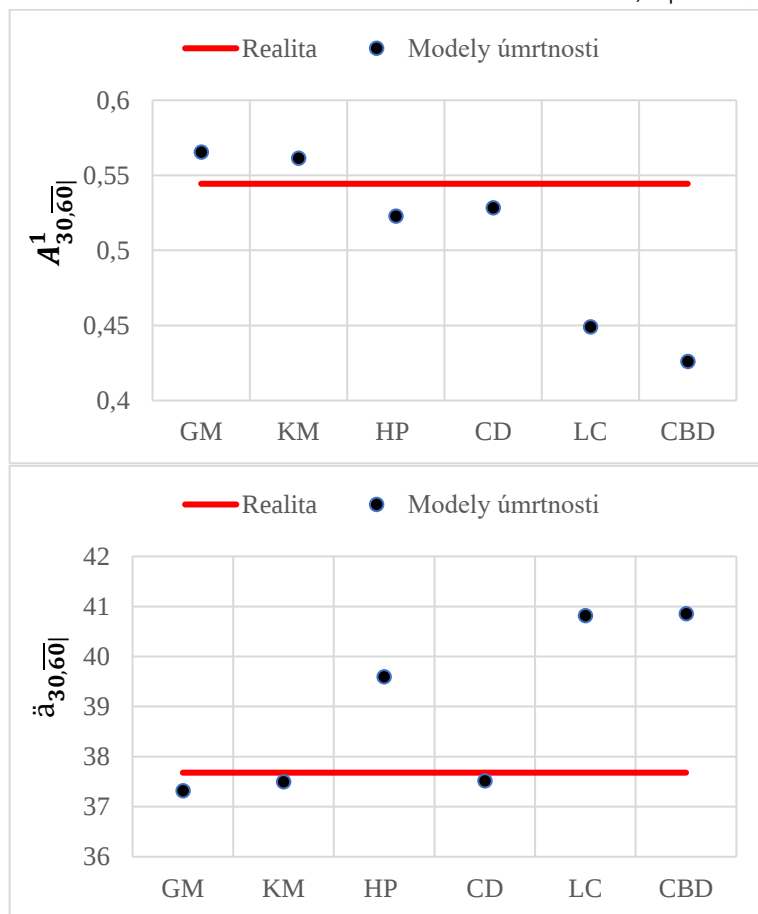
Model	$A_{30,\overline{60} }^1$	$\ddot{a}_{30,\overline{60} }$	Model	$A_{70,\overline{20} }^1$	$\ddot{a}_{70,\overline{20} }$
GM	0,56558	37,31921	GM	0,76905	10,51556
KM	0,56144	37,49934	KM	0,76507	10,64854
HP	0,52286	39,59510	HP	0,73256	12,48671
CD	0,52852	37,51868	CD	0,67640	12,40714
LC	0,44918	40,82240	LC	0,69607	13,05442
CBD	0,42611	40,86078	CBD	0,65858	12,94262

Zdroj: vlastné spracovanie

Aby sme lepšie ukázali, čo to pre poisťovňu znamená, budeme predpokladať, že poistná suma nebude jednotková, ale napríklad vo výške 1 000 €. V tomto prípade by mohol byť rozdiel medzi jednorazovým netto poistným, ktoré si životná poisťovňa vypýta za krytie rizika úmrtnosti, až 140 € v závislosti od toho, aký model na modelovanie úmrtnosti použije. Tento rozdiel nie je až taký výrazný, no pri dôchodkovom poistení môže byť rozdiel v prípade $\ddot{a}_{30,\overline{60}|}$ až 3 542 €.

Vzhľadom na to, že výsledky sa viac odlišujú pri dlhšom trvaní poistenia, v grafickej analýze znázorníme výšky poistných sadzieb pre 30 ročnú osobu s poistnou dobou 60 rokov. Na grafe č. 47 sú znázornené výšky jednotlivých poistných sadzieb vypočítaných na základe pravdepodobností úmrtia získaných pomocou rôznych modelov úmrtnosti. Červenou čiarou je na grafe č. 47 zobrazená úroveň poistnej sadzby vypočítaná pomocou reálnych údajov pred aplikovaním úmrtnostných modelov. Keďže vypočítané hodnoty sú vysoko volatilné a neberú do úvahy vývoj úmrtnosti v čase, čiara má len informatívny charakter. Na základe grafu č. 47 môžeme povedať, že najväčšie rozdiely sú práve medzi deterministickými a stochastickými modelmi úmrtnosti a životná poisťovňa by mala zvážiť, či bude brať do úvahy očakávané zlepšovanie úmrtnosti.

Graf č. 47 - Porovnanie poistných sadzieb pre $A_{30:\overline{60}|}^1$ a $\ddot{a}_{30:\overline{60}|}$



Zdroj: vlastné spracovanie

4.4.2 Analýza vplyvu gender smernice

Analýzu vplyvu gender smernice o rodovej rovnosti na výšku jednorazového netto poistného rovnako ukážeme na dočasnom predlehotnom dôchodku a na dočasnom poistení na úmrtie. Pri výpočtoch poistných sadzieb použijeme pravdepodobnosti získané pomocou Heligmanovho-Pollardovho zákona úmrtnosti pre rok 2020, ktorý na základe RMSE patrí k modelom, ktoré najviac vyhovujú úmrtnostnému správaniu populácie Slovenska. Konkrétne poistné sadzby sme vypočítali pre veku 10, 20, 30, 40, 50, 60, 70 a 80, pričom uvažujeme u oboch produktov s dĺžkou trvania poistenia 20 rokov.

Jednotlivé výšky poistných sadzieb pre rôzne veku a dobu trvania poistenia 20 rokov pri dočasnom poistení na úmrtie sú zaznamenané v tabuľke č. 18 a graficky znázornené na grafe č. 48. Tabuľka č. 18 okrem toho obsahuje aj zmenu poistných sadzieb mužov a žien v percentách oproti spoločnej sadzbe, ako aj rozdiel poistných sadzieb mužov a žien v absolútnom a relatívnom vyjadrení. Síce sú sadzby poistného veľmi malé vzhľadom na použitie jednotkovej poistnej sumy pre tento typ poistenia, rozdiely medzi poistnými sadzbami pred a po zavedení rovnosti žien a mužov sú pomerne veľké.

Tabuľka č. 18 - Vplyv gender smernice na sadzby pre dočasné poistenie na úmrtie

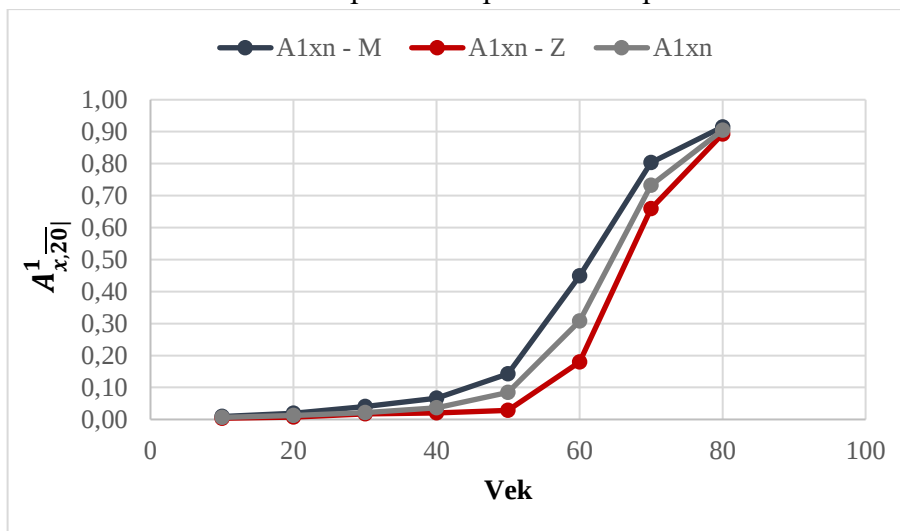
Vek	$A^1_{x,20 M}$	$A^1_{x,20 Z}$	$A^1_{x,20 }$	Zmena M	Zmena Z	Absolútny rozdiel (M-Z)	Percentuálny rozdiel (M-Z)
10	0,00917	0,00379	0,00640	-30,17 %	69,02 %	0,00538	84,03 %
20	0,01903	0,00697	0,01251	-34,25 %	79,55 %	0,01206	96,39 %
30	0,04010	0,01758	0,02149	-46,41 %	22,20 %	0,02252	104,78 %
40	0,06653	0,02059	0,03657	-45,04 %	77,56 %	0,04594	125,63 %
50	0,14239	0,02811	0,08481	-40,44 %	201,67 %	0,11428	134,74 %
60	0,44876	0,17947	0,30726	-31,53 %	71,21 %	0,26929	87,64 %
70	0,80373	0,65899	0,73256	-8,85 %	11,16 %	0,14474	19,76 %
80	0,91481	0,89261	0,90441	-1,14 %	1,32 %	0,02221	2,46 %

Zdroj: vlastné spracovanie

Pri dočasnom poistení na úmrtie hrá úmrtnosť veľmi dôležitú úlohu. Z grafu č. 48 je zrejmý vplyv zavedenia gender smernice na výšku poistného pre tento typ poistenia. Do veku 30 rokov sú rozdiely menšie a vznikajú najmä v dôsledku toho, že havarijný hrb je viditeľný viac u mužskej časti populácie. Avšak potom začnú rásť a pre vyššie veku sú už

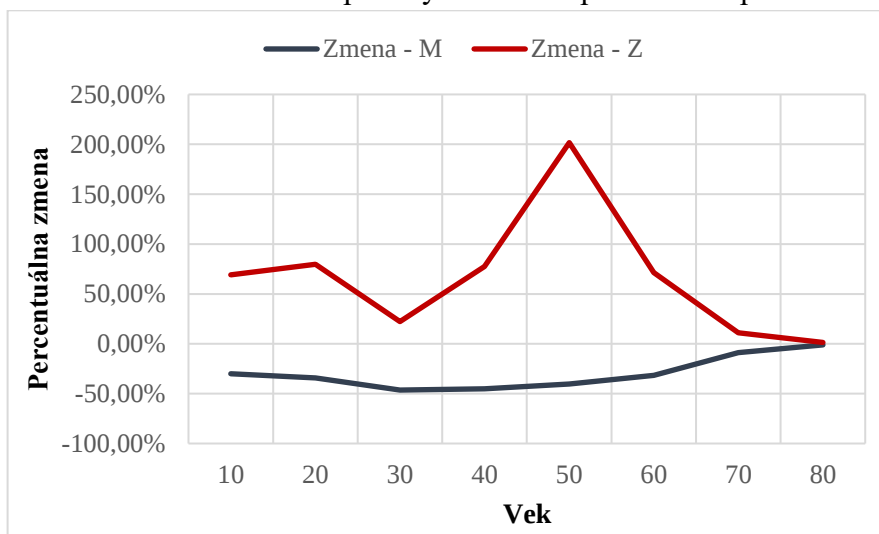
rozdiely neprehliadnuteľné. Napokon pre osoby nad 80 rokov sa poistné sadzby znovu približujú.

Graf č. 48 - Porovnanie poistného pre dočasné poistenie na úmrtie



Zdroj: vlastné spracovanie

Graf č. 49 - Percentuálna zmena poistných sadzieb pre dočasné poistenie na úmrtie



Zdroj: vlastné spracovanie

Môžeme však urobiť aj iné závery. Napríklad, kým vo veku 20 rokov je nárast poistného pri unisex úmrtnostných tabuľkách o 79,55 % vyšší než pre ženy, vo veku 70 rokov je tento nárast „len“ 11,16 %. Muži zase platia vo veku 20 rokov o 34,25 % menej, ako by platili pred zavedením rodovej rovnosti pohlaví. Vo veku 60 rokov je to „iba“ o 8,85 %. Najvyššia zmena je viditeľná u 50 ročných žien, ktoré by za dočasné poistenie

na úmrtie na dobu 20 rokov zaplatili o 201,67 % vyššie jednorazové netto poistné. Naše tvrdenia sú podložené grafom č. 49, kde je zaznamenaná zmena poistnej sadzby v percentách po zavedení gender smernice na základe tabuľky č. 18.

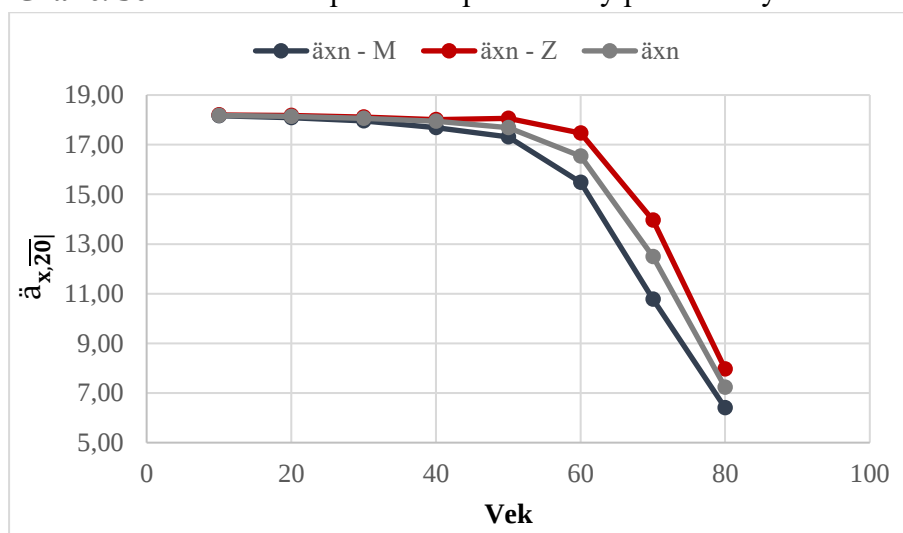
Teraz prevedieme analýzu vplyvu rodovej rovnosti mužov a žien na sadzby poistného pre dočasný predlehotný dôchodok. V tabuľke č. 19 a na grafe č. 50 sú znázornené jednotlivé výšky poistných sadzieb pre rôzne veku a dobu trvania poistenia 20 rokov.

Tabuľka č. 19 - Vplyv gender smernice na sadzby pre dočasný predlehotný dôchodok

Vek	$\ddot{a}_{x,20 }$ M	$\ddot{a}_{x,20 }$ Z	$\ddot{a}_{x,20 }$	Zmena M	Zmena Z	Absolútny rozdiel (M-Z)	Percentuálny rozdiel (M-Z)
10	18,16581	18,19673	18,18126	0,09 %	-0,08 %	-0,03091	-0,17 %
20	18,08545	18,17587	18,13293	0,26 %	-0,24 %	-0,09042	-0,50 %
30	17,95477	18,11709	18,05883	0,58 %	-0,32 %	-0,16232	-0,90 %
40	17,68599	18,01163	17,94880	1,49 %	-0,35 %	-0,32565	-1,81 %
50	17,31784	18,05785	17,68928	2,14 %	-2,04 %	-0,74001	-4,18 %
60	15,48741	17,47102	16,53390	6,76 %	-5,36 %	-1,98361	-12,00 %
70	10,77700	13,96145	12,48671	15,86 %	-10,56 %	-3,18445	-25,50 %
80	6,39894	7,96442	7,22730	12,95 %	-9,26 %	-1,56549	-21,66 %

Zdroj: vlastné spracovanie

Graf č. 50 - Porovnanie poistného pre dočasný predlehotný dôchodok

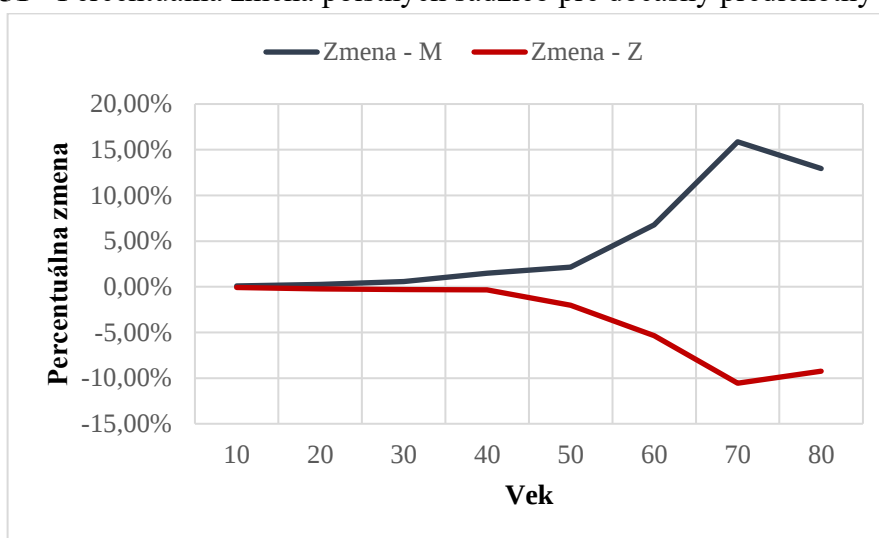


Zdroj: vlastné spracovanie

Na rozdiel od dočasného poistenia na úmrtie, po zavedení rovnakých sadzieb poistného pre mužov a ženy, muži platia vyššie poistné a ženy nižšie. Pred zavedením gender smernice

do praxe, vzhľadom na to, že ženy sa dožívali vyššieho veku, a teda pravdepodobnosť vyplatenia dôchodku bola vyššia, ženy platili vyššie poistné. U mužov je to samozrejme naopak. Ako môžeme vidieť na grafe č. 50, rozdiely opäť narastajú až vo vyšších vekoch. Voči dočasnému poisteniu na úmrtie je však nárast, resp. pokles poistných sadzieb oproti používaniu rôznych úmrtností žien a mužov nižší. Významné percentuálne zmeny sú u osôb nad 60 rokov. Napríklad muž vo veku 70 rokov, ktorý by chcel dočasný predľhotný dôchodok na 20 rokov by musel zaplatiť jednorazové netto poistné vyššie o 15,86 %, zatiaľ čo 70 ročná žena by platila o 10,56 % menej. Toto tvrdenie je zrejmé z grafu č. 51.

Graf č. 51 - Percentuálna zmena poistných sadzieb pre dočasný predľhotný dôchodok



Zdroj: vlastné spracovanie

Na záver môžeme skonštatovať, že hoci sa rozdiely v úmrtnosti mužov a žien neustále znižujú, ako sme si ukázali v kapitole 1.2.1, pri určitých typoch poistenia má zavedenie gender smernice stále významný vplyv na výpočet jednorazového netto poistného. Pri dočasnom poistení na úmrtie ženy zaplatia na poistnom v priemere až o 67 % viac zatiaľ čo u mužov je to v priemere o 30 % menej, ako to bolo pred zavedením nariadenia do praxe. Na druhej strane, pri dôchodkovom poistení naopak sú ženy zvýhodnené a muži zaplatia viac, avšak v priemere to je do 5 %.

Keď však začneme porovnávať reálne brutto alebo tarifné poistné, rozdiely už závisia aj od typu poisťovne, akým spôsobom oceňuje jednotlivé produkty, aké využíva faktory a postupy pri odhadoch a výpočtoch sadzieb poistného.

4.4.3 Analýza vplyvu koronavírusu

Ako sme už spomenuli v kapitole 1.2.1, pandémie Covid-19 mala negatívny vplyv na vývoj úmrtnosti populácie Slovenska, čo sa môže odzrkadliť aj na výške poistného produktov životného poistenia. Dôsledky pandémie na životné poistenie ukážeme na základných produktoch životného poistenia, ktorými sú dočasné poistenie na úmrtie (tzv. úverové poistenie) a dočasný predľehotný dôchodok. V analýze sa zameriame na dva úmrtnostné modely, deterministický Heligmanov-Pollardov model a stochastický Lee-Carterov model. Pri deterministickom modeli porovnáme poistné sadzby vypočítané pomocou pravdepodobností úmrtia a prežitia získaných za roky 2019 a 2020. Pri analýze so stochastickým modelom úmrtnosti sa zameriame na to, ako ovplyvní poistné sadzby vynechanie pandemického roku 2020 pri prognózovaní vývoja parametrov LC modelu.

Jednotlivé poistné sadzby pre rôzne veku s dobou trvania poistenia 20 rokov pri dočasnom poistení na úmrtie a dočasnom predľehotnom dôchodku za roky 2019 a 2020 vypočítané za použitia Heligmanovho-Pollardovho modelu úmrtnosti sú znázornené v tabuľke č. 20 a na grafe č. 52. Tabuľka č. 20 obsahuje aj percentuálnu zmenu, ktorá vyjadruje o koľko vzrástla, respektíve poklesla, poistná sadzba v dôsledku zmeny úmrtnosti medzi rokmi 2019 a 2020. Na základe hodnôt v tabuľke č. 20, vidíme že došlo k poklesu poistných sadzieb pre $A_{x:\overline{20}|}^1$ a nárastu v prípade $\ddot{a}_{x:\overline{20}|}$ pre osoby do 30 rokov. Je to najmä z toho dôvodu, že v 2020 došlo k menšiemu počtu havárii, a teda úmrtnosť v oblasti havarijného hrbu bola nižšia.

Tabuľka č. 20 - Vplyv koronavírusu na sadzby pri deterministickom modeli úmrtnosti

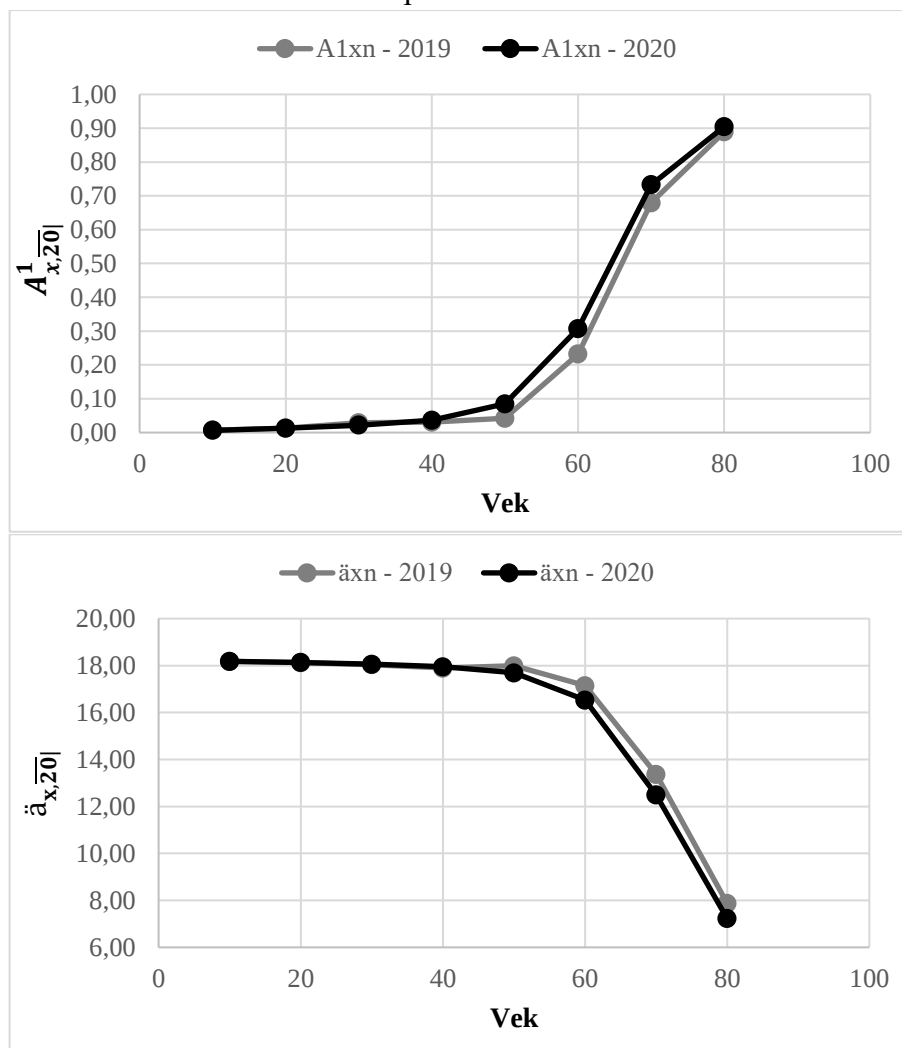
Deterministický Heligmanov-Pollardov model úmrtnosti						
Vek	$A_{x:\overline{20} }^1$ - 2019	$A_{x:\overline{20} }^1$ - 2020	Zmena	$\ddot{a}_{x:\overline{20} }$ - 2019	$\ddot{a}_{x:\overline{20} }$ - 2020	Zmena
10	0,00693	0,00640	-7,67 %	18,17767	18,18126	0,02 %
20	0,01320	0,01251	-5,20 %	18,12680	18,13293	0,03 %
30	0,02881	0,02149	-25,41 %	18,03559	18,05883	0,13 %
40	0,03060	0,03657	19,48 %	17,89555	17,94880	0,30 %
50	0,04217	0,08481	101,10 %	17,98912	17,68928	-1,67 %
60	0,23171	0,30726	32,61 %	17,14464	16,53390	-3,56 %
70	0,67915	0,73256	7,86 %	13,36587	12,48671	-6,58 %
80	0,88875	0,90441	1,76 %	7,87246	7,22730	-8,20 %

Zdroj: vlastné spracovanie

Pri dôchodkových poisteniach predpokladáme, že zhoršenie úmrtnostnej situácie, bude mať za následok pokles poistného, keďže rizikom pre dôchodkové poistenia je dožívanie a nie úmrtie, čo je vidieť aj na grafe č. 52. Keďže pandémia ovplyvnila najviac vyššie vekové kategórie, pozorovateľné je to až od veku 50 rokov. Ak by poistná suma nebola 1 €, ale napríklad 1 000 €, 60 - ročná osoba by v roku 2020 zaplatila jednorazové netto poistné nižšie o 611 €, čo predstavuje pokles o 3,56 %.

Na rozdiel od dôchodkového poistenia, pri zhoršení úmrtnostnej situácie pri poistení na úmrtie by došlo k nárastu poistného pre osoby staršie ako 40 rokov. Ak by si teda 60 ročná osoba chcela zaobstarať dočasné poistenie na úmrtie na 20 rokov s poistnou sumou 10 000 € v roku 2020 namiesto 2019, zaplatila by zaň jednorazové netto poistné vyššie o 756 €, čo znamená nárast o 32,61 %.

Graf č. 52 - Porovnanie sadzieb za použitia deterministického modelu úmrtnosti



Zdroj: vlastné spracovanie

Ďalej ukážeme, ako by táto analýza vyzerala, ak by sme použili pri modelovaní úmrtnosti stochastický Lee-Cartevov model. Tomu, ako sa zmenila prognóza parametra $\kappa_t^{(2)}$ LC modelu pri zohľadnení a vynechaní roku 2020, sme sa venovali v časti 4.2.1. Teraz sa pozrieme na to, aký by to malo vplyv na poistné sadzby dočasného poistenia na úmrtie a dočasného predlehotného dôchodku. Tabuľku č. 21 tvoria poistné sadzby pre rôzne veku s rovnakou dobou trvania 20 rokov pri dočasnom poistení na úmrtie a dočasnom predlehotnom dôchodku pred a po vynechaní roku 2020 v procese prognózovania parametra Lee-Carterovho modelu. Porovnanie poistných sadzieb je tiež graficky znázornené na grafe č. 53.

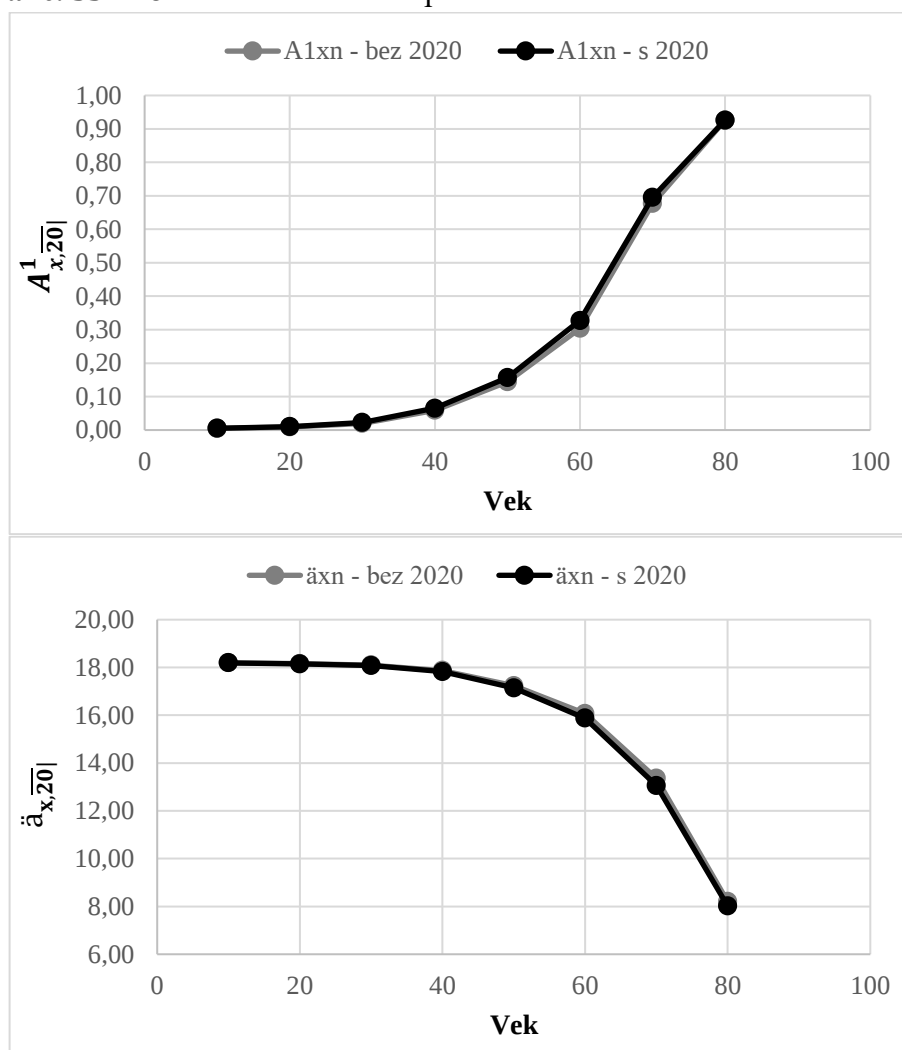
Tabuľka č. 21 - Vplyv koronavírusu na sadzby pri stochastickom modeli úmrtnosti

Stochastický Lee-Carterov model úmrtnosti						
Vek	$A_{x,\overline{20} }^1$ - bez 2020	$A_{x,\overline{20} }^1$ - s 2020	Zmena	$\ddot{a}_{x,\overline{20} }$ - bez 2020	$\ddot{a}_{x,\overline{20} }$ - s 2020	Zmena
10	0,00486	0,00536	10,21 %	18,19352	18,19000	-0,02 %
20	0,00928	0,01027	10,65 %	18,15669	18,14954	-0,04 %
30	0,02004	0,02266	13,05 %	18,09862	18,08242	-0,09 %
40	0,05891	0,06520	10,69 %	17,87174	17,82771	-0,25 %
50	0,14405	0,15655	8,68 %	17,23382	17,14109	-0,54 %
60	0,30410	0,32762	7,73 %	16,06778	15,87888	-1,18 %
70	0,67725	0,69607	2,78 %	13,36030	13,05442	-2,29 %
80	0,92642	0,92650	0,01 %	8,20282	8,01898	-2,24 %

Zdroj: vlastné spracovanie

V porovnaní s deterministickým modelom úmrtnosti, ponechanie pandemického roku pri prognózovaní parametra, ovplyvnilo všetky vekové kategórie. Tento dopad však nie je ani zďaleka taký výrazný ako v prípade deterministického modelu, čo nám potvrdzuje aj graf č. 53. Napríklad, ak by sme zobrali do úvahy rovnaké parametre ako v predchádzajúcom prípade, teda vek 60 rokov, doba trvania poistenia 20 rokov a poistná suma 1 000 €, jednorazové netto poistné pre dôchodkové poistenie by sa znížilo o 189 €, čo predstavuje pokles o 1,18 %. Pri poistení na úmrtie s poistnou sumou 10 000 € by jednorazové netto poistné s rovnakými parametrami vzrástlo o 7,73 %, čo je v absolútnej hodnote 235 €.

Graf č. 53 - Porovnanie sadzieb za použitia stochastického modelu úmrtnosti



Zdroj: vlastné spracovanie

Zvýšená úmrtnosť v dôsledku koronavírusu má taktiež vplyv aj na poistné sadzby produktov životného poistenia. To, v akej miere, závisí aj od druhu modelu, ktorý si poisťovňa vyberie na modelovanie úmrtnosti. Pri dôchodkových produktoch by došlo v dôsledku koronavírusu k zníženiu poistných sadzieb, a teda k zlacneniu týchto poistných produktov, pričom najviac by sa táto situácia odzrkadlila pri osobách starších ako 50 rokov. Pri použití deterministického modelu úmrtnosti by došlo k poklesu poistnej sadzby dočasného predlehotného dôchodku v priemere o 2,4 %, zatiaľ čo pri použití stochastického modelu úmrtnosti by to bolo len o 0,8 %. Naopak pri poistení na úmrtie zvýšená úmrtnosť viedla k vyšším poistným sadzbám. Na sadzby dočasného poistenia na úmrtie mal koronavírus väčší vplyv ako na dočasný predlehotný dôchodok. Pri deterministickom modeli

došlo k nárastu sadzby v priemere o 16 % a ovplyvnilo to najmä vyššie vekové kategórie. V prípade stochastického modelu úmrtnosti došlo k nárastu sadzieb pre všetky vekové kategórie, ale v priemere to bolo len o menej ako polovicu, teda 8 %.

Táto situácia však reálne v praxi nenastala, keďže v životnom poistení sa uzatvárajú dlhodobé poistné produkty a jedno-, resp. dvojročný výraznejší výkyv v úmrtnosti populácie neovplyvní kalkuláciu poistného. Poist'ovne sa z dlhodobého hľadiska chránia pred zvýšenou úmrtnosťou ako rizikom istými formami zaistenia, ktoré znižujú riziko strát poist'ovne. Okrem toho sa v prípade životného poistenia pri oceňovaní produktov, za účelom preventívnej ochrany pred následkami zlého výberu, zvyčajne volia pesimistické predpoklady. Každá poist'ovňa sa už v minulosti snažila počítať s pandémiou, a preto si nemôže dovoliť nebyť pripravená, ak také niečo nastane. Pandémia koronavírusu naplno zasiahla fungovanie krajiny a poist'ovne aktuálne evidujú zvýšený záujem klientov o informácie, či a ako životné poistenie kryje aj ochorenie na COVID-19. Mnohé poist'ovne sa rozhodli aj nad rámec poistných podmienok neuplatňovať žiadnu výluky v priamej súvislosti s koronavírusom, t. j. z dôvodu pandémie alebo infekčnej choroby. Platí to pre riziká práceneschopnosti, hospitalizácie, vážnych chorôb, invalidity a smrti vo všetkých produktoch životného poistenia, dokonca často aj tých, ktoré už nie sú v aktívnom predaji.

Odpoveď na to, aké následky malo krytie rizika spojeného s koronavírusom na jednotlivé poist'ovne, a tiež ako celkovo ovplyvnil COVID-19 poistný obchod, je možné nájsť v Správe o solventnosti a finančnom stave poist'ovne (SFCR). Na základe SFCR jednotlivých poist'ovní sme dospeli k záveru, že v dôsledku pandémie došlo k poklesu záujmu o sporiace a investičné druhy produktov životného poistenia. Naopak výrazne sa zvýšil dopyt po rizikovom životnom poistení a zdravotnom poistení. Vzhľadom na to, že sa poist'ovne rozhodli kryť riziko spojené s koronavírusom, prejavilo sa to aj na náraste poistných plnení a poklesu ziskovosti súvisiacich produktov životného poistenia. Okrem toho pandémia ovplyvnila aj finančné trhy, čo sa takisto odzrkadlilo na číslach uverejnených v SFCR. Veľkou výzvou v čase pandémie bolo pre poist'ovne tiež skvalitnenie a zefektívnenie bezkontaktného predaja poistení, alebo aj elektronického hlásenia poistných udalostí.

Záver

Úmrtnosť je významným predpokladom aktuárskej bázy. Životná poisťovňa na základe pravdepodobnosti úmrtia v závislosti od veku vykonáva mnohé poisťno-matematické výpočty, medzi ktoré patrí napríklad kalkulácia poistného a technických rezerv. Dramatická zmena úmrtnosti má u životnej poisťovne za následok vážne finančné straty. Očakávania o úmrtnostnom správaní poistného kmeňa úzko súvisia s typom produktu životnej poisťovne. Podhodnotenie úmrtnosti má negatívny vplyv na dočasné poistenie na úmrtie a naopak, v prípade dôchodkového poistenia môže práve nadhodnotenie úmrtnosti viesť k vysokým stratám poisťovne.

Predmetom dizertačnej práce bolo modelovanie úmrtnosti pomocou deterministických a stochastických modelov, porovnanie ich aplikovateľnosti na úmrtnostné správanie populácie Slovenska, a tiež ich využitie v oblasti aktuárstva.

V úvode dizertačnej práce sme predstavili problematiku týkajúcu sa úmrtnosti v oblasti poisťovníctva. Oboznámili sme sa so základnými pojmami úmrtnosti ako súčasti aktuárskej bázy. Ďalej sme bližšie opísali úmrtnosť, jej ukazovatele a analyzovali sme vývoj úmrtnosti od vzniku Slovenskej republiky až po súčasnosť. Štatistiky úmrtnosti v posledných rokoch výrazne ovplyvnila pandémia koronavírusu. Infekcia COVID-19 mala za následok 6,8 % úmrtí, pričom najviac sa prejavila vo vekovej skupine seniorov (vo veku nad 65 rokov). Z dlhodobého hľadiska sa úmrtnosť neustále zlepšuje, za čo môžeme vďačiť najmä rozvoju medicíny a ďalších vied súvisiacich so smrťou. V celom sledovanom období pretrvávala nadúmrtnosť mužskej časti populácie, avšak rozdiely v úmrtnosti medzi pohlaviami klesajú. Na konci prvej kapitoly sme priblížili legislatívnu úpravu poisťovníctva v Európskej únii so zameraním na gender smernicu, ktorá sa zásadne dotýka práve úmrtnosti, keďže sa zaoberá problematikou vylúčenia pohlavia ako faktora vstupujúceho do poisťno-matematických výpočtov.

Druhá kapitola vymedzuje ciele záverečnej práce. Hlavným cieľom dizertačnej práce bolo teoreticky a aj prakticky preskúmať problematiku modelovania úmrtnosti pomocou rôznych typov parametrických modelov z aktuárskeho hľadiska. Následne aplikovať nadobudnuté poznatky na modelovanie úmrtnosti populácie Slovenska, a taktiež analyzovať vplyvy, akými sú gender smernica a koronavírus, ktoré majú dopad na samotné modelovanie úmrtnosti. Tieto ciele sme naplnili po teoretickej aj praktickej stránke.

Na začiatku 3. kapitoly sme opísali vybrané deterministické a stochastické modely úmrtnosti. Ukázali sme ich základné vzťahy, prednosti, ale aj nedostatky, ktoré môžu mať vplyv na presnosť modelov v praktickej časti. Venovali sme sa dôkladnej charakteristike metód odhadu parametrov modelov úmrtnosti. Pri deterministických modeloch sme použili metódy nelineárnej regresie. Vybrali sme si Gaussovu-Newtonovu metódu, Gradientnú metódu a Levenbergovu-Marquardtovu metódu, ktoré sú súčasťou analytického softvéru SAS Enterprise Guide. Za najefektívnejšiu metódu môžeme považovať Levenbergovu-Marquardtovu metódu, ktorá spája prednosti ostatných dvoch spomínaných metód. Na odhad parametrov stochastických modelov sa používa niekoľko metód. My sme si zvolili metódu najmenších štvorcov za pomoci singulárneho rozkladu matice a Habermanovho-Russolillovho prístupu. Na prognózovanie vývoja parametrov stochastických modelov sme vybrali modely ARIMA. ARIMA vytvárajú veľmi presné prognostické modely pre rôzne druhy údajov, pričom platí, že využívajú najnovšie pozorovania pri analyzovaní minulých chýb prognóz pre správne odhadnutie budúceho vývoja. Ďalej sme sa v tejto kapitole zaoberali mierami kvality modelov a presnosti ich prognóz – priemerný súčet štvorcov rezíduí, štandardná odchýlka rezíduí, Akaikovo informačné kritérium a Bayesovské informačné kritérium, ktoré nám poslúžili ako nástroj na vzájomné porovnanie modelov v praktickej časti ako aj na porovnanie prognostických schopností stochastických modelov. V závere kapitoly sme priblížili softvérové rozhranie SAS Enterprise Guide so zameraním na procedúry nelineárnej regresie a ARIMA modelovania.

Praktická časť záverečnej práce je pre čitateľov najviac prínosná, pretože je priamo zameraná na aplikáciu deterministických a stochastických modelov úmrtnosti na reálne dáta o úmrtnostnom správaní populácie Slovenskej republiky v čase od jej vzniku v roku 1993 až po súčasnosť. Predmetom modelovania sú vekovo špecifické miery úmrtnosti zvlášť mužov a žien, a tiež spoločné miery úmrtnosti v rozpätí od 0 do 100 rokov vypočítané pomocou údajov o úmrtnosti získaných zo Štatistického úradu Slovenskej republiky. V prípade deterministických modelov sme vekový interval rozdelili na dve časti (od 0 do 50 rokov a od 51 do 100 rokov), čo prispelo k lepšej interpretovateľnosti výsledkov. Najväčším nedostatkom starších modelov bolo zanedbanie dojčenskej úmrtnosti. Z deterministických bol CoDe model ako jediný schopný spoľahlivo odhadnúť hodnotu úmrtnosti do 1 roku života. Pri Heligmanovom-Pollardovom modeli sa nám to podarilo len v niektorých rokoch.

Čo sa týka stochastických modelov, tak Lee-Carterov model nemal s odhadom doľčenskej úmrtnosti takisto problém. Jednoduchšie modely oproti tým zložitejším mali ešte jednu slabinu, ktorou bol tvar krivky. Vzhľadom na nenáročnosť mali jednoduchý exponenciálny tvar, ktorý sa nevedel prispôbiť úmrtnostnému správaniu populácie v tzv. havarijnom hrbe, viditeľnom v období okolo 20 roku života. Preto dosahovali lepšie výsledky na dátach v roku 1993, kedy tento hrb nebol taký výrazný. Za výhodu stochastických modelov považujeme to, že zohľadňujú očakávané zlepšovanie úmrtnosti v čase. CBD model je však vhodný iba na modelovanie úmrtnosti populácie nad 30 rokov, čo nie je pre životnú poisťovňu veľmi praktické. Okrem toho, v porovnaní s Lee-Carterovým modelom, intervaly spoľahlivosti pre výsledné miery úmrtnosti boli pomerne široké, čo naznačuje, že odhad je nepresný. Najnáročnejší na prácu bol CoDe model a hoci výborne opisuje doľčenskú úmrtnosť a oblasť havarijného hrbu, vo vysokých vekoch nedosahoval až tak dobré výsledky, ako sme očakávali v súvislosti s jeho zložitou. Už samotné nastavenie konštantných parametrov si vyžadovalo dôslednú analýzu.

Keďže poisťovňa sa snaží čo najpresnejšie odhadnúť úmrtnosť svojich poistencov v závislosti od veku, snaží sa vybrať taký model, ktorý je schopný čo najlepšie opísať úmrtnostné správanie populácie. Po individuálnej analýze vybraných modelov, sme ich preto porovnali na základe miery presnosti, konkrétne štandardnej odchýlke rezíduí RMSE. Na základe výsledkov, modely, ktoré najviac vyhovujú dátam o úmrtnostnom správaní populácie Slovenska môžeme považovať Heligmanov-Pollardov deterministický model úmrtnosti a Lee-Carterov stochastický model úmrtnosti. V závislosti od pohlavia sme analyzovali len vybrané roky 1993, 2002, 2011 a 2020. Výsledky porovnania modelov boli podobné ako v prípade zanedbania pohlavia pri analýze, veľmi dobré výsledky však dosahoval aj CoDe model.

Nakoniec sme analyzovali, aký dopad má výber konkrétneho modelu na poisťné sadzby jednorazového netto poisťného, ale aj vplyv zavedenia rodovej rovnosti a efekt zhoršenej úmrtnosti v dôsledku koronavírusu na konci sledovaného obdobia na deterministický a stochastický model. Vychádzali sme zo základných poisťných produktov, ktorými sú dočasné poistenie na úmrtie a dočasný predľhotný dôchodok. Poisťné sadzby za použitie vybraných modelov úmrtnosti sa výrazne odlišovali. V prípade dočasného poistenia na úmrtie bol rozdiel medzi najvyšším a najnižším vypočítaným jednorazovým netto poisťným až 33 % a pri dočasnom predľhotnom dôchodku to bolo okolo 9 %. Najväčšie

rozdiely boli práve medzi deterministickými a stochastickými modelmi úmrtnosti a životná poisťovňa by mala zvážiť, či bude brať do úvahy očakávané zlepšovanie úmrtnosti. Čo sa týka zavedenia rodovej rovnosti, hoci sa rozdiely v úmrtnosti mužov a žien neustále zmenšujú, pri určitých typoch poistenia má zavedenie gender smernice stále významný vplyv pri výpočte jednorazového netto poistného. Zvýšená úmrtnosť v dôsledku koronavírusu má taktiež vplyv aj na poistné sadzby produktov životného poistenia. To, v akej miere, závisí aj od druhu modelu, ktorý si poisťovňa vyberie na modelovanie úmrtnosti.

Výskum sme realizovali najmä v spomínanom prostredí SAS Enterprise Guide. Ako pomocné nástroje nám poslúžili aj softvérové rozhranie R a Microsoft Excel. Výpočtovej realizácii praktickej časti práce predchádzalo množstvo času štúdia jednotlivých modelov a metód odhadu ich parametrov, ako aj štúdium práce s použitým softvérovým rozhraním. Vzhľadom na to, že sme analyzovali obdobie medzi rokmi 1993 až 2020 a zároveň úmrtnosť zvlášť v závislosti od pohlavia, ale aj pri zanedbaní pohlavia ako faktora, ktorý vplýva na úmrtnosť, pričom sme sa tiež snažili zachytiť vplyv koronavírusu na parametre stochastických modelov, samotný proces modelovania spočíval v odhade nesmierne veľkého množstva modelov úmrtnosti. Preto veríme, že dizertačná práca prinesie čitateľom nové informácie a poznatky z modelovania úmrtnosti a poukáže na dôležitosť štatistiky v aktuárskych vedách.

Zoznam obrázkov, tabuliek a grafov

Tabuľka č. 1 - Úmrtnosť podľa príčiny na Slovensku v roku 2020.....	23
Tabuľka č. 2 - Zákony úmrtnosti.....	37
Tabuľka č. 3 - Základné stochastické modely ľudskej úmrtnosti.....	44
Tabuľka č. 4 - Odhad parametrov GM modelu pre roky 1993 a 2020.....	86
Tabuľka č. 5 - Odhad parametrov KM modelu pre roky 1993 a 2020.....	91
Tabuľka č. 6 - Odhad parametrov HP modelu pre roky 1993 a 2020.....	96
Tabuľka č. 7 - Odhad parametrov CD modelu pre roky 1993 a 2020.....	105
Tabuľka č. 8 - Odhad parametra $\kappa_t^{(2)}$ LC modelu.....	111
Tabuľka č. 9 - Odhad parametrov $\beta_x^{(1)}$ a $\beta_x^{(2)}$ LC modelu.....	112
Tabuľka č. 10 - Prognóza $\kappa_t^{(2)}$ LC modelu s 95 % intervalom spoľahlivosti.....	116
Tabuľka č. 11 - Odhad parametra $\kappa_t^{(1)}$ CBD modelu.....	119
Tabuľka č. 12 - Odhad parametra $\kappa_t^{(2)}$ CBD modelu.....	120
Tabuľka č. 13 - Prognóza parametra $\kappa_t^{(1)}$ CBD modelu s 95 % intervalom spoľahlivosti..	124
Tabuľka č. 14 - Prognóza parametra $\kappa_t^{(2)}$ CBD modelu s 95 % intervalom spoľahlivosti..	125
Tabuľka č. 15 - Hodnoty RMSE modelov úmrtnosti pre ženy.....	129
Tabuľka č. 16 - Hodnoty RMSE modelov úmrtnosti pre mužov.....	130
Tabuľka č. 17 - Vplyv výberu modelu na poisťné sadzby.....	132
Tabuľka č. 18 - Vplyv gender smernice na sadzby pre dočasné poistenie na úmrtie.....	134
Tabuľka č. 19 - Vplyv gender smernice na sadzby pre dočasný predľehotný dôchodok..	136
Tabuľka č. 20 - Vplyv koronavírusu na sadzby pri deterministickom modeli úmrtnosti..	138
Tabuľka č. 21 - Vplyv koronavírusu na sadzby pri stochastickom modeli úmrtnosti.....	140
Obrázok č. 1 - Princíp režimu Solventnosť II.....	27
Obrázok č. 2 - Dôsledky odstránenia pohlavia ako faktora z výpočtu poisťného.....	29
Obrázok č. 3 - Nelineárna regresná analýza v SAS Enterprise Guide.....	70
Obrázok č. 4 - Fázy iteračného postupu.....	71
Obrázok č. 5 - Sumár procesu odhadu parametrov.....	72
Obrázok č. 6 - Analýza najmenších štvorcov.....	72
Obrázok č. 7 - Odhad parametrov.....	73

Obrázok č. 8 - Korelačná matica.....	73
Obrázok č. 9 - Graf závislosti miery úmrtnosti od veku.....	73
Obrázok č. 10 - ARIMA modelovanie v SAS Enterprise Guide.....	74
Obrázok č. 11 - Opisné štatistiky.....	75
Obrázok č. 12 - Grafy autokorelačných funkcií pred a po pridaní integračného člena.....	76
Obrázok č. 13 - Test bieleho šumu.....	77
Obrázok č. 14 - Odhad parametrov ARIMA modelu.....	78
Obrázok č. 15 - Štatistiky goodness of fit.....	78
Obrázok č. 16 - Korelačná matica.....	78
Obrázok č. 17 - Testy autokorelácie rezíduí.....	79
Obrázok č. 18 - Grafická analýza rezíduí.....	79
Obrázok č. 19 - Matematické vyjadrenie modelu.....	80
Obrázok č. 20 - Prognóza.....	80
Obrázok č. 21 - Grafické znázornenie prognózy.....	81
Obrázok č. 22 - Prognóza $\kappa_t^{(2)}$ LC modelu pri zohľadnení a vynechaní roku 2020.....	117
Obrázok č. 23 - Prognóza $\kappa_t^{(1)}$ CBD modelu pri zohľadnení a vynechaní roku 2020.....	126
Obrázok č. 24 - Prognóza $\kappa_t^{(2)}$ CBD modelu pri zohľadnení a vynechaní roku 2020.....	126
Graf č. 1 - Vývoj hrubej miery úmrtnosti na Slovensku.....	24
Graf č. 2 - Očakávaná dĺžka života pri narodení a jej rozdiel medzi mužmi a ženami.....	25
Graf č. 3 - Úmrtnosť na Slovensku podľa pohlavia a veku v rokoch 1993 a 2020.....	26
Graf č. 4 - Miery úmrtnosti mužov a žien v roku 2020.....	83
Graf č. 5 - GM model úmrtnosti mužov a žien pre rok 1993.....	87
Graf č. 6 - Vývoj parametra A GM modelu v čase.....	88
Graf č. 7 - Vývoj parametra B GM modelu v čase.....	89
Graf č. 8 - Vývoj parametra c GM modelu v čase.....	89
Graf č. 9 - MSE GM modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020.....	90
Graf č. 10 - KM model úmrtnosti mužov a žien pre rok 2002.....	92
Graf č. 11 - Vývoj parametra A KM modelu v čase.....	93
Graf č. 12 - Vývoj parametra B KM modelu v čase.....	93
Graf č. 13 - Vývoj parametra C KM modelu v čase.....	94
Graf č. 14 - MSE KM modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020.....	95

Graf č. 15 - HP model úmrtnosti mužov a žien pre rok 2011.....	97
Graf č. 16 - Vývoj parametra A HP modelu v čase.....	98
Graf č. 17 - Vývoj parametra B HP modelu v čase.....	98
Graf č. 18 - Vývoj parametra C HP modelu v čase.....	99
Graf č. 19 - Vývoj parametra D HP modelu v čase.....	99
Graf č. 20 - Vývoj parametra E HP modelu v čase.....	100
Graf č. 21 - Vývoj parametra F HP modelu v čase.....	100
Graf č. 22 - Vývoj parametra G HP modelu v čase.....	101
Graf č. 23 - Vývoj parametra H HP modelu v čase.....	101
Graf č. 24 - Vývoj parametra K HP modelu v čase.....	102
Graf č. 25 - MSE HP modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020.....	102
Graf č. 26 - Vývoj modálneho veku medzi rokmi 1993 až 2020.....	104
Graf č. 27 - CD model úmrtnosti mužov a žien pre rok 2020.....	106
Graf č. 28 - Vývoj parametra A v CD modeli v čase.....	107
Graf č. 29 - Vývoj parametra B v CD modeli v čase.....	107
Graf č. 30 - Vývoj parametra a v CD modeli v čase.....	108
Graf č. 31 - Vývoj parametra b_1 v CD modeli v čase.....	108
Graf č. 32 - Vývoj parametra b_2 v CD modeli v čase.....	109
Graf č. 33 - Vývoj parametra b_3 v CD modeli v čase.....	109
Graf č. 34 - MSE CD modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020.....	110
Graf č. 35 - Parameter $\beta_x^{(1)}$ LC modelu.....	113
Graf č. 36 - Parameter $\beta_x^{(2)}$ LC modelu.....	113
Graf č. 37 - Vývoj parametra $\kappa_t^{(2)}$ LC modelu v čase.....	114
Graf č. 38 - Ukážka LC modelu úmrtnosti mužov a žien pre rok 1993.....	114
Graf č. 39 - MSE LC modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020.....	115
Graf č. 40 - LC Prognóza mier úmrtnosti na rok 2030 s 95 % intervalmi spoľahlivosti..	118
Graf č. 41 - Vývoj parametra $\kappa_t^{(1)}$ CBD modelu v čase.....	121
Graf č. 42 - Vývoj parametra $\kappa_t^{(2)}$ CBD modelu v čase.....	121
Graf č. 43 - Ukážka CBD modelu úmrtnosti mužov a žien pre rok 2020.....	122
Graf č. 44 - MSE CBD modelu bez zohľadnenia pohlavia v rokoch 1993 až 2020.....	123
Graf č. 45 - CBD Prognóza mier úmrtnosti na rok 2030 s 95 % intervalmi spoľahlivosti...	127

Graf č. 46 - Porovnanie modelov na základe RMSE bez zohľadnenia pohlavia.....	128
Graf č. 47 - Porovnanie poistných sadzieb pre $A_{30,\overline{60} }^1$ a $\ddot{a}_{30,\overline{60} }$	133
Graf č. 48 - Porovnanie poistného pre dočasné poistenie na úmrtie.....	135
Graf č. 49 - Percentuálna zmena poistných sadzieb pre dočasné poistenie na úmrtie.....	135
Graf č. 50 - Porovnanie poistného pre dočasný predlehotný dôchodok.....	136
Graf č. 51 - Percentuálna zmena poistných sadzieb pre dočasný predlehotný dôchodok...	137
Graf č. 52 - Porovnanie sadzieb za použitia deterministického modelu úmrtnosti.....	138
Graf č. 53 - Porovnanie sadzieb za použitia stochastického modelu úmrtnosti.....	141

Zoznam použitej literatúry

- [1] ARLT, J., ARTLOVA, M. (2009). *Ekonomické časové rady*. 1. vyd. Praha: Professional Publishing.
- [2] BARDOUTSOS, A., de BEER, J., JANSSEN, F. (2018). *Projecting delay and compression of mortality*. Genus, 74(1), 1-28.
- [3] BEARD, R. E. (1959). *Note on some mathematical mortality models*. In Ciba Foundation Symposium-The Lifespan of Animals (Colloquia on Ageing) (Vol. 5, pp. 302-311). Chichester, UK: John Wiley & Sons.
- [4] BILÍKOVÁ, M. (2003). *Spojité metódy v poisťnej matematike*. Bratislava: EKONÓM.
- [5] BILÍKOVÁ, M., JOHANESOVÁ, M. (2008). *Aktuárske výpočty pre rôzne druhy poistenia m-tice osôb*. Bratislava: EKONÓM.
- [6] BOOTH, H., HYNDMAN, R. J., TICKLE, L., DE JONG, P. (2006). *Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions*. Demographic research, 15, 289-310.
- [7] BOX, G. E., JENKINS, G. M., REINSEL, G. C. (1994). *Time series analysis, forecasting and control*. Englewood Cliffs.
- [8] CAIRNS, A. J., BLAKE, D., DOWD, K., COUGHLAN, G. D., EPSTEIN, D., ONG, A., BALEVICH, I. (2009). *A quantitative comparison of stochastic mortality models using data from England and Wales and the United States*. North American Actuarial Journal, 13(1), 1-35.
- [9] CAIRNS, A. J., BLAKE, D., DOWD, K., COUGHLAN, G. D., EPSTEIN, D., KHALAF-ALLAH, M. (2011). *Mortality density forecasts: An analysis of six stochastic mortality models*. Insurance: Mathematics and Economics, 48(3), 355-367.
- [10] CAIRNS, A. J., BLAKE, D., DOWD, K., (2006). *A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration*. Journal of Risk and Insurance, 73(4), 687-718.
- [11] CHAN, W. S., LI, J. S. H., LI, J. (2014). *The CBD mortality indexes: modeling and applications*. North American Actuarial Journal, 18(1), 38-58.
- [12] CIPRA, T. (1986). *Analýza časových řad s aplikacemi v ekonomii*. 1. vyd. Praha: Alfa, Státní nakladatelství technické literatury
- [13] CIPRA, T. (2006). *Pojistná matematika - teorie a praxe*. Praha: EKOPRESS.

- [14] COALE, A., KISKER, E. E. (1990). *Defects in data on old age mortality in the United States: New procedures for calculating approximately accurate mortality schedules and life tables at the highest ages*. In Asian and Pacific Population Forum, Vol. 4, No. 1, pp. 1-31.
- [15] CURRIE, I. D. (2006). *Smoothing and forecasting mortality rates with P-splines*. Talk given at the Institute of Actuaries.
- [16] DE BEER, J., JANSSEN, F. (2016). *A new parametric model to assess delay and compression of mortality*. Population health metrics.
- [17] DEBÓN, A. M., MONTES, F., SALA, R. (2005). *A comparison of parametric models for mortality graduation. Application to mortality data of the Valencia region (Spain)*. SORT: statistics and operations research transactions, 29(2), pp. 269-288.
- [18] DENNIS Jr, J. E. (1977). *Nonlinear least squares and equations*. New York: Academic Press, 269-312.
- [19] DICSON, D. C. M., HARDY, M. R., WATERS, H. R. (2009). *Actuarial mathematics for life Contingent Risks*. New York: Cambridge University Press.
- [20] EUROSTAT (2021). Data Browser [online databáza]. Tab. Life expectancy at birth by sex. Citované dňa 28. 8. 2021. Dostupné na:
<https://ec.europa.eu/eurostat/databrowser/view/tps00205/default/table?lang=en>
- [21] FILADELFIOVÁ, J. (2005). *Demografická situácia a správanie rodín vz. verejná politika v SR*. Sociológia-Slovak Sociological Review, (5), 387-418.
- [22] FORFAR, D. O. (2014). *Mortality laws*. Wiley StatsRef: Statistics Reference Online, 1-11.
- [23] GAMPE, J. (2010). *Human mortality beyond age 110*. In Supercentenarians (pp. 219-230). Springer, Berlin, Heidelberg.
- [24] GAVRILOVA, N. S., GAVRILOV, L. A. (2011). *Stárnutí a dlouhověkost: Zákony a prognózy úmrtnosti pro stárnoucí populace*. Demografie, 53(2), 109–128.
- [25] GOMPERTZ, B. (1825). XXIV. *On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies*. In a letter to Francis Baily, Esq. FRS &c. Philosophical transactions of the Royal Society of London, (115), 513-583.
- [26] HABERMAN, S., RUSSOLILLO, M. (2005). *Lee Carter mortality forecasting: application to the Italian population*.

- [27] HAINAUT, D. (2012). *Multidimensional Lee–Carter model with switching mortality processes*. Insurance: Mathematics and Economics, 50(2), 236-246.
- [28] HELIGMAN, L., POLLARD, J. H. (1980). *The age pattern of mortality*. Journal of the Institute of Actuaries 107.1
- [29] HUMAN MORTALITY DATABASE. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). [dátum stiahnuté 2.2.2019] Dostupné na: www.mortality.org
- [30] HYNDMAN, R. J., KHANDAKAR, Y. (2008). *Automatic time series forecasting: The forecast package for R*. Journal of Statistical Software, 27(1), 1–22.
- [31] IBRAHIM, R. I. et al. (2017). *Forecasting the mortality rates of Malaysian population using Heligman-Pollard model*. In: AIP Conference Proceedings. AIP Publishing.
- [32] JIANG, L. (1995). *Changing kinship structure and its implications for old-age support in urban and rural China*. Population Studies, 49(1), 127-145.
- [33] KANNISTO, V. (1994). *Development of oldest-old mortality, 1950-1990: Evidence from 28 developed countries* (No. 1). University Press of Southern Denmark.
- [34] KOISSI, M. C., SHAPIRO, A. F., HÖGNÄS, G. (2006). *Evaluating and extending the Lee–Carter model for mortality forecasting: Bootstrap confidence interval*. Insurance: Mathematics and Economics, 38(1), 1-20.
- [35] KOLLER, M. (2012). *Stochastic models in life insurance*. Springer Science & Business Media.
- [36] KRČOVÁ, I. O., SAKALOVÁ, K. (2018). *Analýza hodnoty poisťného v životnom poistení vzhľadom na nariadenie 2004/113/ES*. Economics And Informatics, 16(1).
- [37] KÚTIKOVÁ, J. (2019). *Gompertzov-Makehamov model úmrtnosti mužov a žien v rokoch 1993, 2005 a 2017*. 100 rokov štatistiky v Československu: aktuálne otázky a odpovede, Bratislava, Letra Edu
- [38] KÚTIKOVÁ, J. (2019). *Helligmanov-Pollardov model úmrtnosti mužov a žien pre rok 2017*. 100 rokov štatistiky v Československu: aktuálne otázky a odpovede, Bratislava, Letra Edu
- [39] KÚTIKOVÁ, J. (2020). *Slovak Population Mortality Modelling and Forecasting Using Lee-Carter Model*. EDAMBA 2020: International Scientific Conference for Doctoral Students and Post-Doctoral Scholars. Bratislava: Vydavateľstvo EKONÓM

- [40] KÚTIKOVÁ, J. (2021). *Porovnanie stochastických modelov úmrtnosti na údajoch Slovenska*. Ekonomika a informatika: vedecký časopis FHI EU v Bratislave a SSHI. 19 (1).
- [41] KÚTIKOVÁ, J. (2021). *Slovak Population Mortality Modelling and Forecasting Using Cairns-Blake-Dowd Model*. DOKBAT 2021: 17th Annual International Bata Conference for Ph.D. Students and Young Researchers. Zlín: Tomas Bata University in Zlín, Faculty of Management and Economics
- [42] KÚTIKOVÁ, J. (2021). *Analýza úmrtnosti v Slovenskej republike vzhľadom na životné poistenie*. Mladá veda AIESA 2021: participácia doktorandov a mladých vedeckých pracovníkov na budovaní spoločnosti založenej na vedomostiach. Bratislava: Letra Edu
- [43] KÚTIKOVÁ, J. (2021). *ARIMA modelovanie v SAS Enterprise Guide*. Trendy vo vzdelávaní študentov študijného programu Aktuárstvo: recenzovaný monografický zborník vedeckých prác. České Budějovice: Vysoká škola evropských a regionálnych študií.
- [44] LEE, R. D., CARTER, L. R. (1992). *Modeling and forecasting US mortality*. Journal of the American statistical association 87.419
- [45] LEE, R. (2000). *The Lee-Carter method for forecasting mortality, with various extensions and applications*. North American actuarial journal 4.1
- [46] LEVENBERG, K. (1944). *A method for the solution of certain non-linear problems in least squares*. Quarterly of Applied Mathematics.
- [47] MARQUARDT, D. (1963). *An Algorithm for Least Squares Estimation of Nonlinear Parameters*. Journal of the Society for Industrial and Applied Mathematics.
- [48] MCNOWN, R., ROGERS, A. (1989). *Forecasting mortality: A parameterized time series approach*. Demography, 26(4), 645-660.
- [49] MÉSZÁROS, J. (2016). *Požiadavky modernizácie demografickej štatistiky vo vzťahu k novým metódam graduácie pravdepodobností úmrtia. Testovanie, aplikácia a zhodnotenie nových metód graduácie*. Bratislava: Výskumné demografické centrum INFOSTAT. 35 s. (Pracovný materiál)
- [50] NAU, R., (2020). *Summary of rules for identifying ARIMA models*. Statistical forecasting. Dostupné na: <https://people.duke.edu/~rnau/411home.htm>
- [51] NÁRODNÁ BANKA SLOVENSKA. *Solventnosť II*, Dostupné na:

http://www.nbs.sk/_img/Documents/_Dohlad/ORM/Poistovnictvo/Solventnost_II.pdf

- [52] NOCITO, S. (2015). *Stochastic Mortality Projections: A Comparison of the Lee-Carter and the Cairns-Blake-Dowd models Using Italian Data*. University of Studies of Turin
- [53] O'HARE, C., LI, Y. (2017). *Modelling mortality: are we heading in the right direction?*. Applied Economics, 49(2), 170-187.
- [54] O'HARE, C., LI, Y. (2012). *Explaining young mortality*. Insurance: Mathematics and Economics, 50(1), 12-25.
- [55] PERKS, W. (1932). *On some experiments in the graduation of mortality statistics*. Journal of the Institute of Actuaries, 63(1), 12-57.
- [56] PITACCO, E. (2016). *High age mortality and frailty. Some remarks and hints for actuarial modeling*. CEPAR Working Paper 2016/19. Available at: <http://www.cepar.edu.au/working-papers/working-papers-2016.aspx>
- [57] PITACCO, E., DENUIT, M., HABERMAN, S., OLIVIERI, A. (2009). *Modelling longevity dynamics for pensions and annuity business*. Oxford University Press.
- [58] PISCH, T. (2021). *Aká bola najčastejšia príčina úmrtí na Slovensku v roku 2020? Covid prepisoval štatistiky*. Dostupné na: <https://spravy.rtv.slovensko.sk/2021/03/aka-bola-najcastejsia-pricina-umrti-na-slovensku-v-roku-2020-covid-prepisoval-statistiky-ale-nedominoval/>
- [59] PODMANICKÁ, Z. a kol. (2018) *Štatistika v súvislostiach: Hlavné trendy populačného vývoja v SR v roku 2017*. Bratislava: Štatistický úrad SR, 125 s.
- [60] PREZÍDIUM POLICAJNÉHO ZBORU, *Dopravná nehodovosť v Slovenskej republike*, navštívené dňa 20.12.2021. Dostupné na: <https://www.minv.sk/?statisticke-ukazovatele-sluzby-dopravnej-policie>
- [61] PROMISLOW, S. D. (2015). *Fundamentals of actuarial mathematics (3rd ed.)*. John Wiley & Sons.
- [62] REMUND, A., CAMARDA, C. G., RIFFE, T. (2018). *A cause-of-death decomposition of young adult excess mortality*. Demography, 55(3), 957-978.
- [63] RENSHAW, A. E., HABERMAN, S. (2006). *A cohort-based extension to the Lee-Carter model for mortality reduction factors*. Insurance: Mathematics and economics, 38(3), 556-570.

- [64] ROTAR, V. I. (2014). *Actuarial models: the mathematics of insurance (2nd ed.)*. Chapman and Hall/CRC.
- [65] RUBLÍKOVÁ, E., PRÍHODOVÁ, I. (2008). *Analýza vybraných časových radov–ARIMA modely*. Vydavateľstvo EKONÓM.
- [66] SAIKIA, P., BORAH, M. (2014). *A comparative study of parametric models of old-age mortality*. International Journal of Science and Research, 3(5).
- [67] SAKÁLOVÁ, K. (2001). *Oceňovanie produktov v životnom poistení*. Bratislava: Vydavateľstvo EKONÓM, 2001, 156 s. ISBN 80-225-1350-4.
- [68] SAKÁLOVÁ, K. (2006). *Aktuárske analýzy*. Bratislava: Vydavateľstvo EKONÓM, 2006, 113 s., ISBN 80-225-2155-9.
- [69] SAKÁLOVÁ, K., KOTLÁROVÁ, A. (2015). *Vplyv úrokovej miery na ziskovosť produktov životného poistenia*. Economics And Informatics, 13(2).
- [70] SAS Documentation for Nonlinear regression procedure. [online] Dostupné na: <https://support.sas.com/documentation/onlinedoc/stat/141/nlin.pdf>
- [71] SAS Documentation. [online] Dostupné na: <https://support.sas.com/documentation/onlinedoc/guide/>
- [72] SAS Documentation for ARIMA procedure. [online] Dostupné na: <https://support.sas.com/documentation/onlinedoc/ets/132/arma.pdf>
- [73] SEBER, G. A. F., WILD, C. J. (2003). *Nonlinear Regression*. New Jersey: John Wiley & Sons, Inc.
- [74] SEKEROVÁ, V., BILÍKOVÁ, M. (2007). *Poistná matematika*. Bratislava: Vydavateľstvo EKONÓM, 180 s. ISBN 978-80-225-2302-2.
- [75] SIVAŠOVÁ, D. (2008). *Aktuárska demografia v prostredí konkurenčného poistného trhu*. Bratislava: Vydavateľstvo EKONÓM.
- [76] SMERNICA RADY 2004/113/ES z 13. decembra 2004 o vykonávaní zásady rovnakého zaobchádzania medzi mužmi a ženami v prístupe k tovaru a službám a k ich poskytovaní. [online] Dostupné na: <https://eur-lex.europa.eu/legal-content/SK/TXT/PDF/?uri=CELEX:32004L0113&from=NL>
- [77] SLASPO (2011). *Stanovisko SLASPO k vplyvu antidiskriminačnej agendy Európskej únie na poisťovací priemysel*. Poistné rozhľady. Bratislava: Slovenská asociácia poisťovní, č. 5.

- [78] ŠOLTÉS, E. (2019). *Regresná a korelačná analýza: s aplikáciami v softvéri SAS*. Bratislava: Letra Edu, 2019. 238 s.
- [79] ŠOLTÉSOVÁ, T. (2019). *Aktuárske modelovanie v životnom poistení*. Bratislava: Vydavateľstvo Letra Edu, 148 s
- [80] ŠOLTÉSOVÁ, T., KÚTIKOVÁ, J. (2020). *Využitie regresnej analýzy pri modelovaní úmrtnosti v životnom poistení*. Slovenská štatistika a demografia: vedecký časopis. Bratislava: Štatistický úrad Slovenskej republiky, 30 (3).
- [81] ŠOLTÉSOVÁ, T., KÚTIKOVÁ, J. (2021). *Application of Mortality Laws on the Population in Slovakia Using Nonlinear Regression*. 75th International Scientific Conference on Economic and Social Development Development. Croatia: Varazdin Development and Entrepreneurship Agency
- [82] ŠPIRKOVÁ, J., URBANÍKOVÁ, M. (2012). *Aktuárska matematika - životné poistenie*. Bratislava: Iura Edition.
- [83] ŠPROCHA, B. (2021). *Pandémia ochorenia COVID-19 v roku 2020 a niektoré charakteristiky úmrtnosti na Slovensku*. Slovenská štatistika a demografia, č. 4/2021, s. 59-74.
- [84] THATCHER, A. R., KANNISTO, V., VAUPEL, J. W. (1998). *The force of mortality at ages 80 to 120*. Denmark: Odense University Press.
- [85] WILMOTH, J. R. (1996). *13 Mortality Projections for Japan*. Health and mortality among elderly populations.
- [86] WILMOTH, J.R., ANDREEV, K., JDANOV, D., GLEI, D.A., RIFFE, T. (2017) *Methods protocol for the human mortality database*. Navštívené v decembri 2021. Dostupné na:
<http://www.mortality.org/Public/Docs/MethodsProtocol.pdf>
- [87] WOLF, D. A. (2004). *Another variation on the Lee-Carter model*. In annual meeting of the Population Association of America.
- [88] YADAV, A., YADAV, S., KESARWANI, R. (2012). *Decelerating mortality rates in older ages and its prospects through Lee-Carter approach*. PloS one, 7(12), e50941.
- [89] ZELINOVÁ, S. (2021). *Medzinárodné štandardy finančného výkazníctva s aplikáciami pre prácu aktuára v životnom poistení*. (Dizertačná práca) Ekonomická univerzita v Bratislave.

