

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

Evidenčné číslo: 103006/I/2019/36086129769085956

Zákony úmrtnosti a ich použiteľnosť
pre populáciu na Slovensku

Diplomová práca

Bratislava 2019

Bc. Jana Kútiková

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

**Zákony úmrtnosti a ich použiteľnosť
pre populáciu na Slovensku**

Diplomová práca

Študijný program: Aktuárstvo

Študijný odbor: Kvantitatívne metódy v ekonómii

Školiace pracovisko: Katedra matematiky a aktuárstva

Vedúci diplomovej práce: Mgr. Tatiana Šoltésová, PhD.

Bratislava 2019

Bc. Jana Kútiková



Ekonomická univerzita v Bratislave
Fakulta hospodárskej informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Jana Kútiková
Študijný program: aktuárstvo (Jednoodborové štúdium, inžiniersky II. st., denná forma)
Študijný odbor: 3.3.24 Kvantitatívne metódy v ekonómii
Typ záverečnej práce: Inžinierska záverečná práca
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Zákony úmrtnosti a ich použiteľnosť pre populáciu na Slovensku

Anotácia: Aktuárske výpočty v životnom poistení ovplyvňuje úmrtnostné správanie populácie v danej krajine. V záverečnej práci sa budeme zaoberať úmrtnostným správaním populácie na Slovensku, modelovaním úmrtnosti a zákonmi úmrtnosti. Opíšeme najznámejšie zákony úmrtnosti a pokúsime sa ich aplikovať na skutočné údaje o úmrtnosti populácie na Slovensku.

Vedúci: Mgr. Tatiana Šoltésová, PhD.

Katedra: KMA FHI - Katedra matematiky a aktuárstva FHI

Dátum zadania: 30.10.2017

Dátum schválenia: 03.11.2017

doc. RNDr. Jozef Fecenko, CSc.
vedúci katedry

Čestné vyhlásenie

Čestne vyhlasujem, že záverečnú prácu som vypracovala samostatne, a že som uviedla všetku použitú literatúru.

Dátum:

.....
(podpis študenta)

Pod'akovanie

Touto cestou si dovoľujem poďakovať vedúcej diplomovej práce Mgr. Tatiane Šoltésovej, PhD. za usmernenie a cenné rady pri vypracovávaní diplomovej práce, a ochotu pomôcť pri akýchkoľvek problémoch a otázkach.

ABSTRAKT

KÚTIKOVÁ, Jana: *Zákony úmrtnosti a ich použiteľnosť pre populáciu na Slovensku* – Ekonomická univerzita v Bratislave. Fakulta hospodárskej informatiky; Katedra matematiky a aktuárstva. – Vedúci záverečnej práce: Mgr. Tatiana Šoltésová, PhD. – Bratislava: FHI EU, 2019, 83 strán.

Cieľom práce je nahliadnuť do problematiky modelovania úmrtnosti pomocou vybraných zákonov úmrtnosti a následné využitie poznatkov pri modelovaní úmrtnostného správania sa populácie Slovenskej republiky vo vekovom rozpätí od 0 do 100 rokov života použitím softvéru SAS Enterprise Guide. Práca je rozdelená na štyri kapitoly – súčasný stav problematiky, cieľ práce, metodika práce a metódy skúmania a nakoniec výsledky práce a diskusia. Začiatok práce je venovaný základným pojmom súvisiacim so životným poistením so zameraním na jeden z hlavných predpokladov aktuárskej bázy a to úmrtnosť. Okrem teoretického opisu úmrtnosti a jej ukazovateľov, táto kapitola obsahuje aj vývoj úmrtnosti v období od 1993 do 2017. V ďalšej kapitole sú vymedzené hlavné a vedľajšie ciele záverečnej práce. Tretia kapitola obsahuje opis vybraných zákonov úmrtnosti a použitých metód a prostriedkov na ich modelovanie a porovnávanie. Medzi modely, ktoré sú predmetom nášho záujmu patria Gompertzov-Makehamov model, Coaleov-Kiskerov model, logistický model Kannista, Helligmanov-Pollardov model a najmladší CoDe model. Na ich modelovanie sme použili metódy nelineárnej regresie dostupné v softvéri SAS Enterprise Guide. Štvrtá a zároveň posledná kapitola pozostáva z aplikácií daných modelov na populáciu Slovenska pre roky 1993, 2005 a 2017. Najskôr využijeme pohlavie ako faktor úmrtnosti, aby sme poukázali na rozdiely medzi úmrtnosťou mužov a žien a potom zanedbáme tento faktor z dôvodu legislatívnej úpravy poisťovníctva – Gender smernice. Na porovnanie kvality modelov sme využili priemerný súčet štvorcov rezíduí *MSE*.

Kľúčové slová: životné poistenie, úmrtnosť, Gender smernica, zákony úmrtnosti, nelineárna regresia, SAS Enterprise Guide, priemerný súčet štvorcov rezíduí

ABSTRACT

KÚTIKOVÁ, Jana: *Mortality laws and their usability for the population in Slovakia* – University of Economics, Bratislava. Faculty of Economic Informatics; Department of Mathematics and Actuarial Science – Thesis supervisor: Mgr. Tatiana Šoltésová, PhD. – Bratislava: FHI EU, 2019, 83 pages.

The aim of this thesis is to look into the problematic of mortality modelling using selected mortality laws and the subsequent use of knowledge for modelling mortality pattern of the population of Slovak Republic from age 0 to 100 years using the SAS Enterprise Guide software. The thesis is divided into four chapters – present status of solved problem, the main objective of thesis, methodology of work, results and discussion. The beginning of thesis is dedicated to basic terms related to life insurance, focusing on one of the main components of actuarial base, mortality. In addition to the theoretical description of mortality and its indicators, this chapter also includes the development of mortality in period from 1993 to 2017. The next chapter specify primary and secondary goals of thesis. The third chapter contains a description of selected mortality laws and used methods and means for their modelling and comparison. Among the discussed models are the Gompertz-Makeham model, the Coale-Kisker model, the Kannisto logistic model, the Helligman-Pollard model, and the youngest CoDe model. We used nonlinear regression methods available in SAS Enterprise Guide to model. The fourth and last chapter consists of the application of these models to the population of Slovakia in years 1993, 2005 and 2017. First we use gender as a factor of mortality, to prove the existence of differences between the mortality of men and women, and subsequently we neglect this factor due to legislative regulation of insurance – Gender directive. To compare quality of the models we used mean squared error *MSE*.

Key words: life insurance, mortality, Gender directive, mortality laws, nonlinear regression, SAS Enterprise Guide, mean squared error

Obsah

Úvod.....	9
1 Súčasný stav riešenej problematiky doma a v zahraničí.....	11
1.1 Životné poistenie	11
1.2 Úmrtnosť a jej ukazovatele	12
1.2.1 Úmrtnosť na Slovensku v roku 2017	15
1.3 Legislatívna úprava poisťovníctva EÚ – Gender smernica	17
2 Cieľ práce.....	23
3 Metodika práce a metódy skúmania	25
3.1 Zákony úmrtnosti	25
3.1.1 Konštantná intenzita úmrtnosti	28
3.1.2 Gompertzov zákon úmrtnosti.....	28
3.1.3 Makehamov zákon úmrtnosti.....	30
3.1.4 Coaleov - Kiskerov zákon úmrtnosti	31
3.1.5 Kannistov zákon úmrtnosti	32
3.1.6 Helligmanov - Pollardov zákon úmrtnosti.....	33
3.1.7 CoDe zákon úmrtnosti	34
3.2 Odhad parametrov nelineárnych regresných modelov	36
3.2.1 Gaussova - Newtonova metóda	37
3.2.2 Gradientná metóda	39
3.2.3 Levenbergova - Marquardtova metóda.....	40
3.3 Meranie kvality modelov a presnosti ich prognóz	42
3.4 Softvér SAS Enterprise Guide	44
4 Výsledky práce a diskusia.....	52
4.1 Modely úmrtnosti pri zohľadnení pohlavia.....	53
4.1.1 Gompertzov - Makehamov model úmrtnosti.....	54

4.1.2	Coaleov - Kiskerov model úmrtnosti	58
4.1.3	Kannistov model úmrtnosti	60
4.1.4	Helligmanov - Pollardov model úmrtnosti.....	61
4.1.5	CoDe model úmrtnosti	63
4.2	Porovnanie modelov úmrtnosti populácie Slovenska	65
4.2.1	Analýza v roku 1993	66
4.2.2	Analýza v roku 2005	69
4.2.3	Analýza v roku 2017	73
Záver		77
Zoznam obrázkov, tabuliek a grafov.....		79
Zoznam použitej literatúry		81

Úvod

Veková štruktúra úmrtnosti má vo väčšine krajín Európskej únie podobný tvar. Dojčenská úmrtnosť je značne vysoká, výrazne klesá v rannom detstve, prudko začína narastať v rokoch dospievania, ďalej sa exponenciálne zvyšuje a miera nárastu v starobe postupne klesá. Parametrické modely majú za cieľ reprezentovať toto úmrtnostné správanie pomocou matematického modelu s obmedzeným počtom parametrov. Používajú sa na vyhladenie vekovo špecifických mier úmrtnosti a redukciu štatistických chýb, na konštrukciu úmrtnostných tabuliek, na porovnanie úmrtnosti v jednotlivých krajinách a na prognózy vývoja úmrtnosti populácie založené na analýze zmien hodnôt parametrov.

Disponovať poznatkom o úmrtnosti populácie v závislosti od veku je v životnom poistení nesmierne dôležité. Práve na základe úmrtnostného správania sa populácie realizujú poisťovne mnohé aktuárske výpočty, ako napríklad kalkuláciu poisťného a technických rezerv. Poisťovne sa opierajú o dlhodobé pozorovania, ktorých zdrojom sú úmrtnostné tabuľky. Tieto úmrtnostné tabuľky si môže poisťovňa vytvárať sama na základe informácií o úmrtnosti vo svojom poisťnom portfóliu alebo pre presnejšie hodnoty môže použiť informácie o úmrtnostnom správaní populácie v danej krajine.

Táto diplomová práca sa venuje problematike modelovania úmrtnosti pomocou zákonov úmrtnosti a ich použiteľnosti pre úmrtnostné správanie populácie Slovenskej republiky za pomoci analytického softvéru SAS Enterprise Guide. Predmetom modelovania sú vekovo špecifické miery úmrtnosti mužov a žien v rozpätí od 0 do 100 rokov.

Práca je obsahovo rozdelená na teoretickú a praktickú časť. Teoretická časť je obsiahnutá v prvých troch kapitolách. Prvá kapitola je venovaná základným pojmom súvisiacim so životným poistením so zameraním na jeden z hlavných predpokladov aktuárskej bázy, úmrtnosť. Okrem teoretického opisu úmrtnosti a jej základných ukazovateľov, táto kapitola obsahuje aj vývoj úmrtnosti v období od roku 1993 do 2017. V sledovanom období došlo k zmierneniu rozdielov medzi úmrtnosťou mužov a žien, a teda k poklesu mužskej nadúmrtnosti. Bez ohľadu na stále existujúce rozdiely však začala od decembra 2012 platiť už aj pre oblasť poisťovníctva Gender smernica, v dôsledku ktorej bolo používanie faktora pohlavia pri aktuárskych výpočtoch zakázané. Časť prvej kapitoly je preto zameraná aj na legislatívnu úpravu poisťovníctva a dopad Gender smernice na poisťovne. Druhá kapitola vymedzuje hlavné a vedľajšie ciele záverečnej práce. Tretia kapitola nás uvedie do metodiky diplomovej práce. V rámci nej sa na jednej strane

zoznámime s najznámejšími a najpoužívanějšími zákonmi úmrtnosti, medzi ktoré patria Gompertzov-Makehamov zákon úmrtnosti, Coaleov-Kiskerov zákon úmrtnosti, či Kannistov logistický zákon úmrtnosti. Na druhej strane si predstavíme aj zložitejšie zákony úmrtnosti, konkrétne Helligmanov-Pollardov zákon úmrtnosti a najmladší CoDe model, ktorý pochádza z roku 2016. Našou základnou úlohou je odhadnúť parametre vybraných modelov, pri čom nám pomôže analytický softvér SAS Enterprise Guide. V tejto kapitole sa preto okrem toho venujeme metódam odhadu parametrov nelineárnych regresných modelov, ktoré sú v softvéri dostupné. Sú nimi Gaussova-Newtonova metóda, Gradientná metóda a v neposlednom rade Levenbergova-Marquardtova metóda. Za najefektívnejšiu z nášho pohľadu môžeme považovať Levenbergovu-Marquardtovu metódu, ktorá spája prednosti predchádzajúcich metód. Ďalej sa v tejto kapitole zaoberáme mierami kvality modelov a presnosti ich prognóz, priemerný súčet štvorcov rezíduí a štandardná odchýlka rezíduí, ktoré slúžia pre nás ako nástroj na vzájomné porovnanie modelov v praktickej časti. V závere kapitoly predstavíme softvérové rozhranie SAS Enterprise Guide spolu s procedúrou nelineárnej regresie.

Posledná kapitola tvorí praktickú časť tejto záverečnej práce a je priamo zameraná na aplikáciu zákonov úmrtnosti. Hlavným cieľom je porovnanie projekčných vlastností jednotlivých úmrtnostných modelov pre dáta Slovenskej republiky v rokoch 1993, 2005, 2017 na vekovom intervale od 0 do 100 rokov. Tento interval sme rozdelili na 2 časti (od 0 do 50 rokov, od 51 do 100 rokov), čo prispieva k presnosti modelu. Pre dosiahnutie ešte spoľahlivejších výsledkov by bolo vhodné krivku rozdeliť na viac častí, kde prvá časť modeluje kojeneckú úmrtnosť, druhá časť úmrtnosť do 20 rokov života, tretia časť takzvanú normálnu úmrtnosť pracujúcej populácie a štvrtá časť by sa zaoberala úmrtnosťou vo vyšších rokoch. Praktická časť je bližšie rozdelená na dve podkapitoly. V prvej porovnávame mieru úmrtnosti v závislosti od pohlavia. Pokúšame sa opísať miery úmrtnosti pomocou vybraných modelov, v ktorých berieme do úvahy pohlavie ako faktor úmrtnosti, vďaka čomu sme schopný posúdiť dopad Gender smernice pri odhade úmrtnosti. Na záver, v druhej časti kapitoly zanedbáme pohlavie ako faktor pri modelovaní úmrtnosti a pokúsime sa nájsť model, ktorý najlepšie zodpovedá reálnym dátam úmrtnosti Slovenskej republiky. Pre lepšie pochopenie sme sa snažili všetko znázorniť na grafoch.

Dúfame, že diplomová práca čitateľovi poslúži na rozšírenie vedomostí z oblasti zákonov úmrtnosti a preukáže, 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 kľúčové poznať pravdepodobnosti úmrtia populácie v rôznych vekových skupinách, pretože ich nesprávne určenie môže mať negatívny vplyv na distribúciu financií v rámci poisťovne.

V tejto kapitole sa venujeme základným pojmom životného poistenia so zameraním na úmrtnosť a jej ukazovatele. Analyzujeme vývoj úmrtnosti od roku 1993 do 2017. Ďalej sa zaoberáme otázkou Gender smernice a jej vplyvu na oblasť poisťovníctva. Pri spracovaní kapitoly sme vychádzali hlavne zo zdrojov [4], [5], [11], [13], [22] a [26].

1.1 Životné poistenie

Život je charakteristický množstvom neistôt a náhod, ktoré môžu mať na človeka pozitívne aj negatívne dôsledky. V súvislosti s týmito neistotami sa používa v aktuárstve a v ďalších vedách pojem riziko. Tento pojem je tiež úzko spojený s pojmom poistenie. Poistenie predstavuje prenesenie rizika z poistenca na špecializovanú inštitúciu – poisťovňu, ktorá sa stará o kolektívnu tvorbu finančnej rezervy na náhradu škôd vzniknutých realizáciou rizík. Na to, aby sme riziko mohli označiť za poistné, musí spĺňať niektoré kritériá:

- *identifikovateľnosť* – musí byť jednoznačne možné určenie príčiny udalosti, ktorej výsledkom bude strata krytá poistením,
- *vyčísliteľnosť straty* – je možné poistiť len také udalosti, ktoré spôsobia konkrétne straty dohodnuté v poistnej zmluve a takto vzniknuté straty možno vypočítať,
- *náhodnosť poistnej udalosti* – spočíva v neurčitosti nastatia rizika, ktoré je kryté poistením,
- *ekonomická únosnosť* – poisťovňa môže prevziať do poistenia len také riziko, ktoré je pre ňu únosné.

Poistenie plní funkciu stabilizátora životnej aj ekonomickej úrovne obyvateľstva a pomáha riešiť problémy vzniknuté v súvislosti s náhodnými nepriaznivými udalosťami. V krajinách s vyspelou trhovou ekonomikou, do ktorých spadajú aj krajiny Európskej únie, sa poistenie vo všeobecnosti delí na životné a neživotné. Diplomová práca, v ktorej sa venujeme úmrtnostnému správaniu populácie na Slovensku, súvisí so životným poistením. Jeho význam rastie v nadväznosti na rast životnej úrovne, ale tiež z dôvodu starnutia obyvateľstva. Starnutie obyvateľstva je dlhodobo diskutovanou témou a v Slovenskej

republike prebieha zdola. To znamená, že je spôsobené poklesom pôrodnosti. Viditeľne sa znižuje podiel detskej zložky populácie a rovnako nastáva rast podielu produktívnej zložky obyvateľstva.

Životné poistenie môžeme rozdeliť podľa odvetví, medzi ktoré patria:

- a) poistenie pre prípad smrti, dožitia, smrti alebo dožitia:
 - poistenie pre prípad smrti – základná forma životného poistenia, ktorá poisťuje iba poistné riziko bez sporiacej zložky,
 - poistenie pre prípad dožitia – spočíva vo vytváraní úspor a kapitálu pre budúce potreby, pre prípad dožitia sa veku určeného v poistnej zmluve,
 - zmiešané životné poistenie – zahŕňa poistenie pre prípad dožitia sa dohodnutého veku alebo pre prípad smrti poisteného.
- b) poistenie vena alebo prostriedkov na výživu detí – životné poistenie kumulujúce kapitál od narodenia, ktorého poistná suma sa vypláca v deň sobáša alebo pri dosiahnutí veku 25 rokov,
- c) kapitálové poistenie – poistenie súvisiace s kapitalizačnými zmluvami. Je to zmiešaný produkt, ktorý kryje životné riziká (úmrtie, dožitie, invaliditu a úraz), a okrem toho obsahuje aj sporiacu zložku, pričom peniaze sa zhodnocujú bezpečnejšie ako v prípade investičného poistenia,
- d) investičné životné poistenie – takisto obsahuje poistnú ochranu životných rizík, no zároveň ponúka aj investovanie do podielových fondov. Poistenec si zvolí stratégiu a mieru rizika, ktorú je ochotný podstúpiť,
- e) dôchodkové poistenie – ide o špeciálne životné poistenie pre prípad dožitia, kde sa pravidelne opakuje poistné plnenie formou výplat dôchodku. Cieľom je zabezpečenie finančných prostriedkov a udržanie životného štandardu aj v dôchodkovom veku, počas plnej invalidity či zabezpečenie pozostalých v prípade predčasnej smrti poisteného,
- f) poistenie pre prípad úrazu alebo choroby, ak je možné dojednať k životnému poisteniu pripoistenie.

1.2 Úmrtnosť a jej ukazovatele

Hoci môžu v životnom poistení nastať rôzne poistné udalosti, základnou z nich je úmrtie. Smrť je náhodný jav, pretože nevieme kedy presne nastane. Z tohto dôvodu je

predmetom poistenia, ktoré v prípade nastatia tejto poistnej udalosti poskytnú pozostalým finančné prostriedky.

Úmrtnosť sa vyznačuje vlastnosťami, ktoré môžeme opísať v nasledujúcich bodoch [3]:

- figurujú v nej práve dva stavy, ktoré označujeme pojmami stav živý a stav zomretý, pričom je možné jednoznačne rozhodnúť o stave každého z poistených,
- k prechodu medzi týmito dvoma stavmi môže dôjsť len jedným smerom, zo stavu živý do stavu zomretý a tento proces sa nazýva úmrtie,
- okamih nastatia úmrtia je jav náhodný a nezvratný, ktorý môžeme popísať len použitím nástrojov teórie pravdepodobnosti.

Faktorov ovplyvňujúcich úmrtnosť je mnoho. Medzi významné patrí životná úroveň, pohlavie, vierovyznanie, miera kriminality, nehodovosť, miera chorobnosti, vojny, držanie zbraní, samovraždy. Všeobecne ich môžeme rozdeliť do štyroch východiskových skupín:

a) *genetické faktory*, medzi ktoré patrí napríklad skutočnosť, že miera úmrtnosti mužov je vyššia ako miera úmrtnosti žien. Vzhľadom na fakt, že úmrtnosť žien je nižšia, a že ženy sa dožívajú vyššieho veku, je štúdium/analýza úmrtnosti oddelene za jednotlivé pohlavia logická. Avšak podľa nariadenia Európskej únie 2004/113/ES, ktoré v skratke označujeme ako Gender smernica, majú štáty Európskej únie zakázané rozlišovať faktor pohlavia pri analýze úmrtnosti a tiež aj v ostatných oblastiach. Gender smernici sa budeme ešte venovať v kapitole 1.3 tejto práce.

b) *ekologické faktory*, kde spadajú napríklad životné prostredie a klimatické podmienky,

c) *sociálno-ekonomické faktory* ďalej členiace sa na individuálne a tie, ktoré sú spôsobené vplyvom prostredia. Do individuálnych patrí životná úroveň, úroveň vzdelania, stravovacie návyky, fyzická aktivita, starostlivosť o zdravie. K vplyvom prostredia prislúcha úroveň a stav zdravotníctva, s ktorým súvisí dostupnosť a kvalita lekárskej starostlivosti, ďalej sociálne zabezpečenie, či ekonomická situácia krajiny.

d) *pokrok v medicíne*, ktorá za posledné desaťročie dosiahla neuveriteľné zlepšenie či už v dostupnosti liečiv alebo v nových liečebných postupoch pri špecifických civilizačných ochoreniach akými sú cukrovka, rakovina, srdcovo cievne ochorenia a mnohé ďalšie.

Pri analýze úmrtnosti môžeme využiť absolútne ukazovatele, ktoré ale nie sú vhodné na porovnávanie v čase a v priestore, preto sa častejšie sledujú relatívne ukazovatele úmrtnosti. Znalosť miery úmrtnosti poistených osôb je potrebná pri každom type poistných produktov životného poistenia. Okrem miery úmrtnosti sa pri posúdení poistného rizika berú do úvahy aj iné faktory, ako napríklad zdravotný stav, osobné návyky, zamestnanie a záľuby a iné, ktoré ovplyvňujú pravdepodobnosť nastatia poistnej udalosti. Individuálne miery

úmrtnosti poistencov nemusia byť rovnaké v porovnaní s mierami úmrtnosti celej populácie, ktoré sú zisťované centrálnymi štatistickými úradmi.

Najsledovanejšou charakteristikou je všeobecná miera úmrtnosti populácie, ktorá podáva informácie o procese vymierania populácie. Výsledná hodnota tohto ukazovateľa nám dáva informáciu o počte zomretých osôb pripadajúcich na 1 000 obyvateľov stredného stavu v danom kalendárnom roku. Keďže tento ukazovateľ nerešpektuje vekovú štruktúru obyvateľstva, nie je vhodný na porovnávanie. Aj z tohto dôvodu sú vhodnejšie a častejšie využívané špecifické miery úmrtnosti, ktoré nám podávajú informáciu o zmenách úmrtnosti podľa pohlavia a veku. Ukazovateľ môžeme definovať ako počet zomretých mužov alebo žien v dokončenom veku na 1 000 mužov, alebo žien v rovnakom veku. S rastúcim vekom úmrtnosť stúpa, výnimkou sú však najnižšie vekové kategórie. Vyššia úmrtnosť je v tomto prípade spojená s prvým rokom života. Ďalej vo veku 15 – 35 rokov sa úmrtnosť tiež mierne zvyšuje, a to predovšetkým v dôsledku vonkajších príčin smrti ako sú napríklad úrazy, dopravné nehody, otravy a podobne. Môžeme s určitosťou povedať, že vo všetkých vekových kategóriách miera úmrtnosti mužov prevyšuje mieru úmrtnosti žien a tento jav je označovaný ako tzv. mužská nadúmrtnosť.

Ďalším veľmi významným ukazovateľom je stredná dĺžka života novorodenca alebo stredná dĺžka života osoby v presnom veku x . Informuje nás o priemernom počte rokov, ktoré sa ešte novorodenec dožije. Jedná sa o hypotetický ukazovateľ, ktorý vychádza z predpokladu zachovania úmrtnostných pomerov a vyjadruje súčasnú úmrtnostnú situáciu. Keďže jeho hodnota poukazuje na súčasnú situáciu a nehovorí nič o budúcnosti sledovanej populácie, používa sa preto na porovnanie rôznych populácií. Okrem pomenovania stredná dĺžka života sa tiež často používajú pojmy očakávaná dĺžka života, či nádej dožitia.

Ukazovateľ strednej dĺžky života má niekoľko nevýhod, v dôsledku čoho sa častejšie používa medián dĺžky života. Tento medián sa nazýva pravdepodobná dĺžka života novorodenca a charakterizuje vek, ktorého sa dožije polovica obyvateľstva. Samozrejme tiež musí byť splnený predpoklad, že úmrtnostné pomery populácie ostanú nezmenené. Tento ukazovateľ nie je ovplyvňovaný extrémnymi hodnotami, no napriek tomu sa využíva len zriedka. Modus dĺžky života sa nazýva normálna dĺžka života a interpretujeme ho ako vek, v ktorom ľudia umierajú najčastejšie. V roku 2017 ľudia umierali najčastejšie vo veku 86 rokov, čo oproti 1993 predstavuje zvýšenie o 6 rokov.

Ako sme už spomenuli vyššie, úmrtnosť v prvom roku života jedinca je veľmi vysoká, a preto sa jej venuje osobitá pozornosť prostredníctvom ukazovateľa dojčenskej úmrtnosti. Tento ukazovateľ hovorí o počte detí, ktoré zomreli v priebehu prvého roku

života, teda 365 dní po narodení na 1 000 živo narodených detí. Mnohokrát sa ukazovateľ dojčenskej úmrtnosti využíva pre medzinárodné porovnávania. Aby sme zistili pravdepodobnosť úmrtia novorodenca v priebehu prvého roka života, použijeme kvocient dojčenskej úmrtnosti. Vypočítame ho ako podiel zomretých novorodencov k počtu narodených detí.

Ako dôsledok skvalitnenia zdravotníckej starostlivosti sa hranica úmrtnosti narodených detí posúva čím ďalej, tým viac k prvopočiatkom života, čo viedlo k vzniku ukazovateľov merajúcich úmrtnosť v skorších obdobiach života. Sem spadá ukazovateľ popôrodnej úmrtnosti, ktorý sleduje koľko detí zomrelo počas prvých troch dní ich života na 1 000 živo narodených detí. Včasná novorodenecká úmrtnosť zisťuje počet detí, ktoré zomreli počas prvých siedmich dní života. Tento ukazovateľ je súčasťou perinatálnej úmrtnosti, ktorá okrem počtu zomretých novorodencov v priebehu prvých siedmich dní života, berie do úvahy aj počet mŕtvo narodených detí na 1 000 živo narodených. Ďalším ukazovateľom je novorodenecká úmrtnosť, ktorú zaujíma úmrtnosť detí v priebehu prvých 28 dní života, taktiež na 1 000 živo narodených detí. Spomenieme si ešte ponovorodeneckú úmrtnosť, ktorá zastrešuje obdobie od 29. dňa života do jedného roka a sleduje počet zomretých detí v tomto období znovu na 1 000 živonarodených detí.

1.2.1 Úmrtnosť na Slovensku v roku 2017

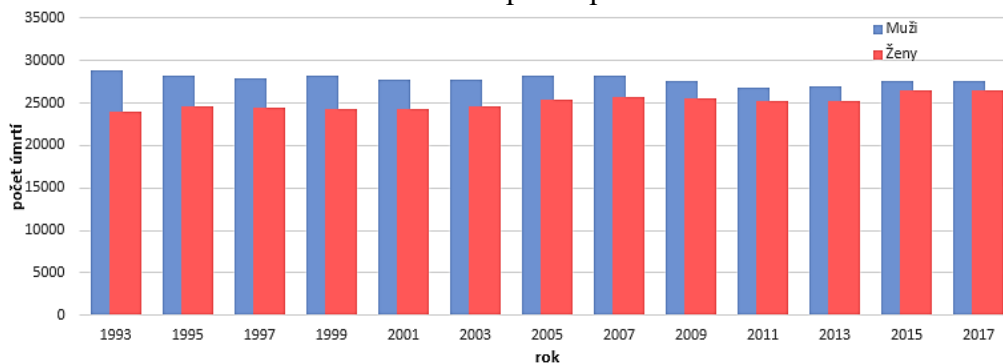
V roku 2017 zomrelo na Slovensku celkovo 53 914 obyvateľov, z čoho bolo 27 489 mužov a 26 425 žien. Nadúmrtnosť mužov bola viditeľná do veku 77 rokov vrátane. Najčastejšími príčinami úmrtia u oboch pohlaví boli choroby obehovej sústavy. Miera dojčenskej úmrtnosti v roku 2017 dosiahla úroveň 4,5 ‰, čo v absolútnom vyjadrení predstavuje úmrtie 263 detí vo veku do 1 roka.

V porovnaní s predchádzajúcim rokom počet zomretých vzrástol takmer o 1,6 tisíc obyvateľov, čo viedlo k zvýšeniu hrubej miery úmrtnosti o 0,3 bodu (z 9,6 ‰ na 9,9 ‰). Mierny medziročný pokles bol zaregistrovaný v prípade dojčenskej úmrtnosti, reálne o 48 úmrtí, čím sa miera dojčenskej úmrtnosti znížila o 0,9 bodu (z 5,4 ‰ na 4,5 ‰).

Počas pozorovaného obdobia medzi rokmi 1993 až 2017 sa počet úmrtí pohyboval v intervale 51 000 – 54 000, pričom najnižší počet úmrtí bol zaznamenaný v roku 1996, kedy zomrelo 51 236 ľudí. Naopak, najvyšší počet úmrtí pripadlo práve na rok 2017. Hrubá miera úmrtnosti sa udržiavala pod úrovňou 10 úmrtí na 1 000 obyvateľov, s výnimkou v roku

2008, kedy predstavovala presne 10,0 %. Najnižšie hodnoty hrubej miery úmrtnosti na úrovni 9,5 % boli zaznamenané v rokoch 1996 a 2014.

Graf č. 1 - Úmrtnosť na Slovensku podľa pohlavia v rokoch 1993 – 2017

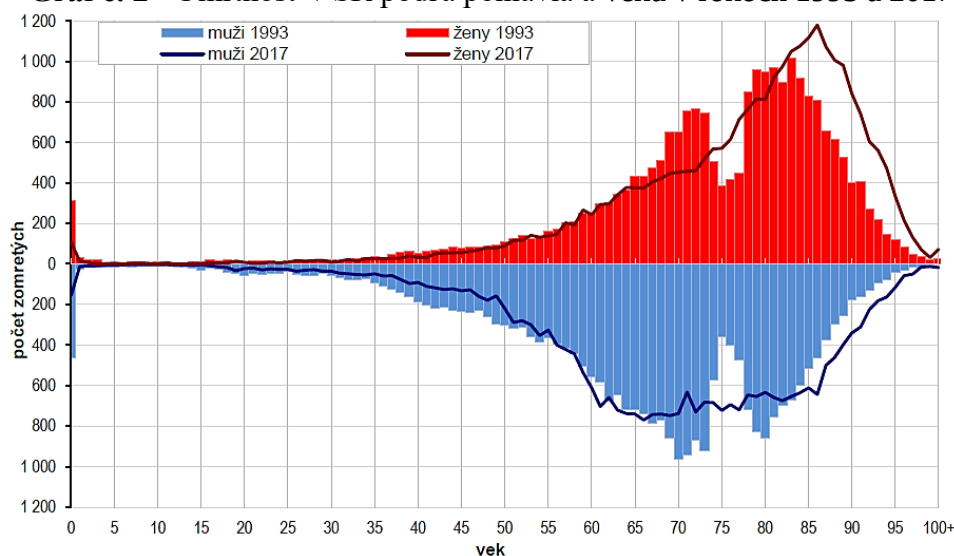


Zdroj: vlastné spracovanie

V sledovanom období pretrvávala nadúmrtnosť mužskej populácie. Rozdiely medzi úmrtnosťou mužskej a ženskej populácie sa však neustále znižujú, čo môžeme vidieť aj na grafe č. 1. Zatiaľ čo v roku 1993 pripadalo na 1 000 zomretých žien až 1 200 zomretých mužov, v súčasnosti je to len 1 040 zomretých mužov na 1 000 zomretých žien.

Zmena trendu úmrtnosti podľa pohlavia a veku v rokoch 1993 a 2017 je zobrazená na grafe č. 2. Z grafu je zrejmé, že úmrtnosť vo vyšších vekoch sa stále posúva, za čo môžeme vďačiť neustále sa rozvíjajúcemu technickému pokroku v medicíne, sociológii, farmácii a iných, so smrťou priamo či nepriamo súvisiacich vied. Kvalita života sa zlepšuje každým rokom, a to sa samozrejme odráža aj na vývoji úmrtnosti.

Graf č. 2 - Úmrtnosť v SR podľa pohlavia a veku v rokoch 1993 a 2017



Zdroj údajov: Štatistický Úrad SR

1.3 Legislatívna úprava poisťovníctva EÚ – Gender smernica

Členské štáty Európskej únie vstupom do aliancie prisľúbili akceptovať právo únie a spolu s tým súvisí aj rešpektovanie jej nariadení. Európska únia uverejnila viacero smerníc a nariadení týkajúcich sa oblasti poisťovníctva, ktoré slúžia na reguláciu a dohľad tohto odvetvia. Smernice vzťahujúce sa na oblasť poisťovníctva sú verejne dostupné pre obyvateľov Slovenskej republiky na internetovej stránke Ministerstva financií SR.

Za najpodstatnejšiu a najprelomovejšiu v tejto oblasti môžeme považovať smernicu Európskeho parlamentu a Rady 2009/138/ES z 25. novembra 2009 o začatí a vykonávaní poistenia a zaistenia (tzv. smernica Solventnosť II). Úmyslom smernice Solventnosť II bolo nastoliť nový režim regulácie poisťného sektora, a tým zaistiť vyššiu ochranu poistencom a ostatným príjemcom plnenia vyplývajúceho z poisťnej zmluvy. Ďalším cieľom smernice bol aj rozvoj integrácie poisťného trhu Európskej Únie prostredníctvom harmonizácie poisťného sektora, či nárastom medzinárodnej konkurencieschopnosti poisťovní. Mechanizmus Solventnosť II mal byť v Slovenskej republike účinný už v roku 2013, avšak do platnosti sa dostal až 1. januára 2016. Poisťovne nezaznamenali výraznejšie problémy s implementáciou tejto smernice.

Druhou veľmi dôležitou európskou smernicou formujúcou 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 poskytovaniu alebo jej kratšie pomenovanie tzv. Gender smernica. Ako už aj z názvu vyplýva tematikou smernice je antidiskriminácia podľa pohlavia, teda totožné zaobchádzanie s osobami bez ohľadu na ich pohlavie. Pre oblasť poisťovníctva sú direktívou vymedzené pravidlá, podľa ktorých, hoci je pohlavie jedným z kľúčových faktorov oceňovania produktov v poisťovní, je nevyhnutné sa usilovať o zachovanie jednotného zaobchádzania s oboma pohlaviami. Preto je nutné tento faktor z oceňovania vylúčiť. Vo všeobecnosti platilo a stále platí, že faktor pohlavia pomáha poisťovní čo najpresnejšie zachytiť a oceniť jej riziká a produkty, preto sa v odvetví poistenia určila výnimka, ktorej text znel: „Určité kategórie rizík sa môžu pre pohlavia líšiť. Pri hodnotení poisťných rizík je v niektorých prípadoch jedným, nie však nevyhnutne jediným, určujúcim faktorom. V prípade zmlúv, ktorými sa poisťujú tieto typy rizík, môžu členské štáty rozhodnúť o povolení výnimiek z pravidla o prémiech a zľavách jednotne pre obe pohlavia, pokiaľ môžu zabezpečiť, že základné poisťno-matematické a štatistické údaje, na ktorých sú výpočty založené, sú spoľahlivé,

pravidelne aktualizované a prístupné verejnosti.“ [26]. Táto výnimka dala voľnú ruku poisťovniam rozhodnúť sa, či chcú naďalej využívať faktor pohlavia pri svojich výpočtoch, za predpokladu podloženia odlišností medzi pohlaviami na základne aktuársko–štatistických údajov. Komisia tiež vyžadovala, aby členské štáty akceptujúce túto výnimku, zhromažďovali a zverejňovali aktuálne údaje týkajúce sa faktora pohlavia použitého na výpočty. Možnosť využili všetky členské krajiny, a to najčastejšie pri rizikovom životnom poistení a dôchodkovom poistení. Niektoré štáty túto výnimku používali pri poistení motorových vozidiel, z dôvodu štatistických zistení, podľa ktorých sú muži rizikovejšími vodičmi.

V marci 2011 nastal v tejto oblasti veľký zvrät. Na základne podnetu rozhodol Európsky súdny dvor o zrušení výnimky, čím nastolil zákaz používania faktora pohlavia pri aktuárskych výpočtoch týkajúcich sa poistného a poistných plnení ako kľúčového faktora. Rozsudok Súdneho dvora EÚ č. C-236/09 európskeho súdu o zrušení výnimky pre členské štáty Európskej Únie sa stal platný ku dňu 21. decembra 2012.

Zastaveniu používania faktora pohlavia v sektore poisťovníctva predchádzalo predloženie sťažnosti belgickej asociácie spotrebiteľov Test-Achats spolu s dvoma súkromnými osobami, konkrétne Charlesom Basselierom a Yann vanVugta na ústavný súd Belgicka v snahe zabrániť uplatňovaniu výnimky smernice. Argumentom Test-Achats bolo, že rozlišovanie pohlavia predstavuje diskrimináciu. Žaloba bola ďalej posunutá na Európsky súdny dvor, ktorý mal posúdiť kompatibilitu výnimky s právom Európskej Únie. Ako už z predchádzajúcej časti vyplýva, Európsky súd vyhovel žalobe a rozhodol sa výnimku týkajúcu sa poisťovníctva z Gender smernice odstrániť.

Verdikt súdu sa však nestretol s pozitívnymi reakciami, a to nielen u poisťovní, ale aj samotných ich klientov. Či už odborníci, zamestnanci alebo aj klienti v poistnom sektore nekompromisne nesúhlasia, aby sa prestalo pohlavie využívať ako kľúčový faktor oceňovania rizík a poistných produktov. Slovenská asociácia poisťovní (SLASPO), ktorá je združením komerčných poisťovní v Slovenskej republike, uverejnila takisto svoje stanovisko k antidiskriminačnej problematike na svojej internetovej stránke. SLASPO prejavila podporu zákazu diskriminácie ako takej, avšak nestotožňuje sa s tvrdením, že aj použitie tohto faktora na účely poistenia by malo byť chápané ako diskriminácia. Jedným z argumentov SLASPO bolo, že hlavnou úlohou poisťovní je dôsledná práca s rizikami, ktorá si vyžaduje ich čo najpresnejšie oceniť podľa všetkých prípustných faktorov, ešte skôr než ich poisťovňa preberie. A ako aj sama uvádza, kľúčovými faktormi pre oceňovanie v životnom poistení je vek, zdravotný stav, ale predovšetkým tiež pohlavie.

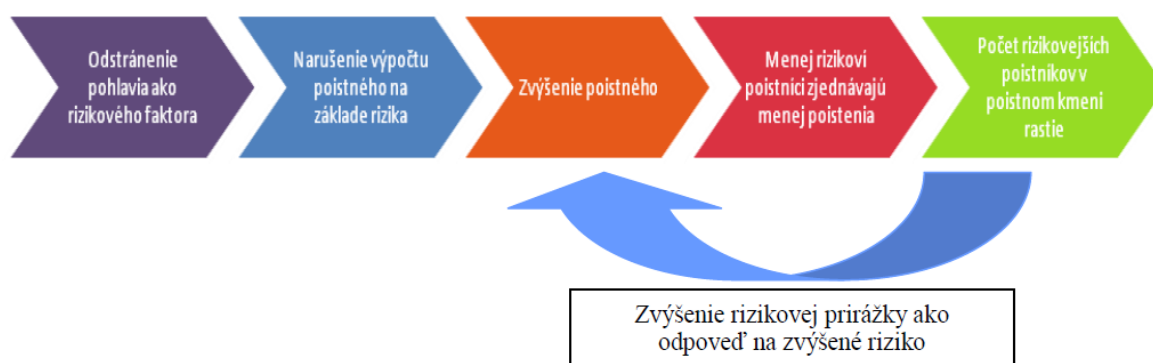
Očakávaným dosahom zrušenia výnimky smernice pre oblasť poisťovníctva sa už v roku 2011 zaoberala spoločnosť OXERA (The Oxford Economic Research Associates) – Oxfordská asociácia ekonomického výskumu, ktorá uverejnila štúdiu o predpokladanom efekte vynechania faktora pohlavia na samotných spotrebiteľov poisťovní. Uvedieme si niektoré postrehy spoločnosti OXERA:

- Riziko úmrtnosti mužov a žien sa vyznačuje významnými rozdielmi, ktoré značne vplyvajú na životné poistenie a dôchodky a odlišné je taktiež riziko zapríčinenia dopravnej nehody v každej z pozorovaných krajín. Predpokladá sa zotrvanie daných rozdielov aj v ďalších rokoch.
- S rizikom úmrtnosti súvisí fakt, že v každej z krajín platia ženy oproti mužom výrazne nižšie poistné pri rizikovom životnom poistení, a to v dôsledku nižšieho rizika ich úmrtia.
- Vzhľadom na skutočnosť, že ženy sa dožívajú v priemere vyššieho veku než muži, inkasujú nižšie mesačné platby pri dôchodkovom poistení oproti mužom, hoci výška zaplateného poistného je totožná.
- Rozdiely týkajúce sa pohlavia sa vyskytujú takisto aj pri poisteniach motorových vozidiel. Najviditeľnejšou skupinou sú mladí vodiči vo veku okolo 25-30 rokov. Ženy v tomto veku sú menej rizikové z dôvodu nižšej pravdepodobnosti účasti na dopravných nehodách.

Na základe týchto bodov štúdia konštatovala percentuálne zmeny v poistnom v prípade schválenia rovnoprávnosti pohlavia v poistení. Pre ženy okolo veku 40 rokov by to znamenalo nárast poistného pri životnom poistení v priemere o viac ako 30 %. Pre mužov okolo veku 65 rokov by sa znížilo poistné plnenie pri poistení na dožitie priemerne o viac ako 5 %. A v prípade poistenia motorového vozidla, by poistné na toto poistenie pre mladé šoférky vzrástlo v priemere cez 11 %.

Podľa OXERY môže nastať v dôsledku nezohľadnenia faktora pohlavia jav, ktorý sa nazýva fenomén nepriaznivého výberu. Môže zapríčiniť, že menej rizikovní poistenci budú platiť vyššie poistné na úkor rizikovejších poistencov. To bude mať za následok pokles záujmu menej rizikových poistencov a môže sa aj stať, že ukončia už existujúce poistné zmluvy, pretože poistenie sa stane pre ich potreby prídrahé. Vplyvom toho sa v poistnom kmeni zvýši podiel rizikovejších poistencov, čo spôsobí ďalšie zvyšovanie poistného a postupné odchádzanie ďalších a ďalších poistencov.

Obrázok č. 1 - Dôsledky odstránenia pohlavia ako faktora z výpočtu poistného



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

Ako sme už viackrát spomenuli, v súvislosti s rizikom existujú medzi pohlaviami výrazné rozdiely, ktoré spôsobujú odlišnosť cien niektorých poistných produktov. Príčiny, vďaka ktorým sú ženy považované za menej rizikové, okrem vyššej osobnej starostlivosti spojenej so zdravým životným štýlom, je taktiež všeobecne známe, že lepšie zvládajú stres a sú opatrnejšími šoférmi. Platnosť tohto tvrdenia je pravdivá väčšinou v ekonomicky vyspelých krajinách, v ktorých vďaka zlepšovaniu zdravotníctva klesla úmrtnosť žien pri pôrode, čo sa ukazuje ako jedna z najrozšírenejších príčin smrti žien v rozvojových krajinách. Dôkazom pravdivosti je práve skutočnosť, ktorú sme už spomenuli, a to, že ženy sa dožívajú v priemere vyššieho veku, čo potvrdzujú úmrtnostné tabuľky. Úmrtnostné tabuľky tvoria nástroj na zisťovanie úmrtnostných pomerov vybranej populácie. Obsahujú presné miery vystihujúce vymieranie populácie podľa pohlavia v jednotlivých vekových skupinách. Medzi najčastejšie používané informácie nachádzajúce sa v úmrtnostných tabuľkách sú pravdepodobnosť úmrtia a stredná dĺžka života pre všetky vekové kategórie. V oblasti poisťovníctva majú široké využitie. Sú nevyhnutné pre väčšinu výpočtov pri produktoch dôchodkového a životného poistenia ako takého. Zvrat v používaní faktora pohlavia v poisťovníctve mal najväčší efekt práve na používanie tabuliek, podľa ktorého sa po novom musia na výpočty používať len a len unisex úmrtnostné tabuľky.

Gender smernica so sebou priniesla mnoho výhod, a tiež nevýhod či už pre poisťovne alebo pre spotrebiteľov. Podľa viacerých odborníkov, ktorí sa venovali tematike zavedenia Gender smernice a jej možným dopadom ešte pred jej platnosťou, je možné zosumarizovať dopady smernice do nasledujúcich bodov:

1. Niektorí poistenci sa budú snažiť profitovať z jednotnej sadzby, preto zrušia svoju aktuálnu poistku a založia si novú, ktorá im zabezpečí nižšie poistné.

2. Spotrebitelia budú pred zavedením smernice špekulovať, či je výhodnejšie sa poistiť pred alebo po zavedení jednotnej sadzby.
3. Poisťovne budú nútené úplne prehodnotiť a zmeniť svoje marketingové a cenové stratégie.
4. Zmena sadzieb si vyžiada prevádzku viacerých systémov, ktoré budú fungovať pre platné portfólio starých poistných zmlúv a ďalšie pre nové poistné zmluvy.

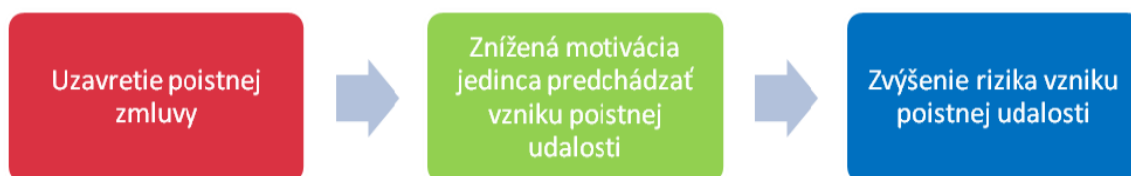
Okrem predpokladaných negatívnych dopadov nastolenia rovnoprávnosti medzi mužmi a ženami v poisťovníctve sa spotrebitelia obávali ovplyvnenia cien poistných produktov a poisťovne o fungovanie tohto nového systému ako celku. Dnes už s určitosťou vieme povedať, ktoré predpovede sa naplnili, a ktoré vôbec nie.

Zo strany poisťovní táto zmena priniesla množstvo dodatočných nákladov spojených s tvorbou nových produktov a ich oceňovaním, či s preceňovaním už existujúcich produktov. Ďalšou veľkou výzvou pre poisťovne bolo upriamiť pozornosť pri oceňovaní produktov na iné významné faktory. Aby sa nenaplnili obavy spotrebiteľov o raste cien, rozhodli sa poisťovne zvýhodniť ceny pripoistení na produktoch, ktorých sa týkali zmeny. Okrem modernizácie a vylepšovania starých produktov dostali poisťovne možnosť orientovať sa na produkty charakteristické pre jednotlivé pohlavia.

Za najväčšiu hrozbu všetkých zmien a obmedzení v legislatíve Európskej Únie sa považuje oslabenie ekonomickej efektívnosti poisťovníctva. Strategickým krokom na znemožnenie nastatia tejto situácie je prevencia antiselekcie a morálneho hazardu.

Morálny hazard spočíva v klesajúcej motivácii konať opatrne, čím rastie riziko, na ktoré sa vzťahuje ochrana. O morálnom hazarde môžeme hovoriť teda len v prípade, že existuje poistná ochrana, ktorá má za následok nárast rizika. Morálny hazard je zjednodušene znázornený na obrázku č. 2.

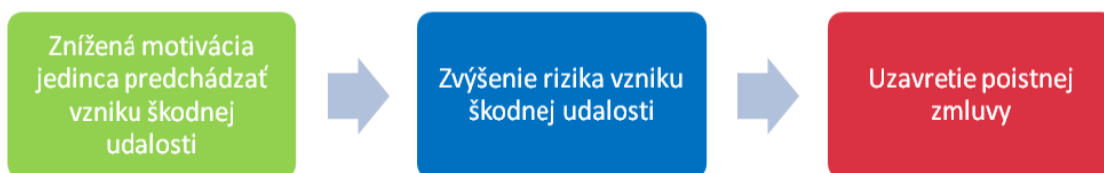
Obrázok č. 2 - Schéma morálneho hazardu



Zdroj: vlastné spracovanie

V prípade nepriaznivého výberu, ktorý sme už spomínali vyššie, je postup opačný. K uzavretiu poisťnej zmluvy dôjde až potom, ako dôjde k poklesu motivácie konať opatrne. Neopatrné správanie je kvázi príčinou toho, prečo sa človek rozhodne uzavrieť poisťnú zmluvu. Schematické znázornenie nepriaznivého výberu je zachytené na obrázku č. 3.

Obrázok č. 3 - Schéma nepriaznivého výberu



Zdroj: vlastné spracovanie

2 Cieľ práce

Cieľom diplomovej práce je predovšetkým modelovanie úmrtnostného správania populácie Slovenskej republiky pomocou známych zákonov úmrtnosti. Existuje veľké množstvo zákonov úmrtnosti, z ktorých vzorku sme zvolili tak, aby obsahovala výber od najjednoduchších zákonov úmrtnosti po najzložitejšie zákony úmrtnosti, od najstarších až po tie, ktoré boli publikované len nedávno. Okrem dôslednej selekcie zákonov úmrtnosti, sme modelovali úmrtnosť v rôznych rokoch s časovými odstupmi 12 rokov – 1993, 2005 a 2017. Týmto sme pripravili vhodné podmienky pre vzájomné porovnanie výsledných modelov.

Začneme od podstaty a budeme sa snažiť dať čitateľom základný pohľad na problematiku týkajúcu sa úmrtnosti v oblasti poisťovníctva. Tu sa čitateľ oboznámi so základnými pojmami životného poistenia, pre ktoré je úmrtnosť veľmi dôležitou súčasťou. Cieľom aplikačnej časti práce bude nájdenie funkcií opisujúcich úmrtnosť populácie na Slovensku spoločne s odhadom ich parametrov. Pokiaľ takýto model nájdeme, dostávame efektívny a úsporný nástroj na najrozmanitejšie operácie s praktickými dátami o úmrtnosti populácie. Ďalej si v kapitole bližšie opíšeme úmrtnosť a jej ukazovatele a pozrieme sa aj na vývoj úmrtnosti medzi rokmi 1993 až 2017 s úmyslom informovať čitateľa o aktuálnej situácii v oblasti úmrtnosti na Slovensku. Jedným z faktorov, ktorý ovplyvňuje úmrtnosť je aj pohlavie, ktoré bolo prostredníctvom Gender smernice zakázané používať pri aktuárskych výpočtoch. Preto je našim zámerom poukázať na rozdiely v úmrtnosti mužov a žien, a taktiež popísať vplyv Gender smernice na oblasť poisťovníctva.

V 3. kapitole sa budeme detailne venovať vybraným zákonom úmrtnosti, ich prednostiam aj nedostatkom, ale aj vzťahom dôležitým pre ich modelovanie. Tiež sa budeme sústrediť na to, aby sme čo najzrozumiteľnejšie opísali metódy odhadu parametrov zákonov úmrtnosti, ktoré spadajú pod nelineárnu regresnú analýzu a sú nevyhnutné pre aplikáciu modelov na reálne dáta o úmrtnosti populácie Slovenska. Aby sme boli schopní porovnať výsledné modely, budeme sledovať miery kvality modelov a presnosti ich prognóz, z ktorých si zvolíme priemerný súčet štvorcov rezíduí *MSE* a štandardnú odchýlku rezíduí *RMSE*. Na dosiahnutie cieľov sme použili analytický softvér SAS Enterprise Guide, ktorého základným rozhraním a ukážkou procedúry nelineárnej regresie spolu s výstupom sme sa zaoberali tiež v tejto časti práce.

V praktickej časti, ktorá tvorí záverečnú kapitolu diplomovej práce, využijeme teoretické poznatky z predchádzajúcich častí práce na samotnú aplikáciu zákonov úmrtnosti s cieľom porovnať projekčné vlastnosti jednotlivých modelov pre dáta Slovenskej republiky. Predmetom modelovania je vekový interval od 0 do 100 rokov. Najskôr graficky porovnáваме vekovo špecifické miery úmrtnosti v závislosti od pohlavia s úmyslom poukázať na rozdiely v úmrtnosti medzi mužmi a ženami. Na záver, pri rešpektovaní Gender smernice, zanedbáme faktor pohlavia pri modelovaní úmrtnosti a pokúsime sa nájsť model, ktorý by čo najlepšie zodpovedal reálnym dátam úmrtnosti slovenskej populácie v rokoch 1993, 2005 a 2017. Okrem softvéru SAS Enterprise Guide sme v praktickej časti použili aj Microsoft Office Excel, pomocou ktorého sú na základe výsledkov zhotovené rôzne grafy a tabuľky.

V konečnom dôsledku je našim úmyslom vytvoriť materiál, ktorý by čitateľovi rozšíril vedomosti z oblasti zákonov úmrtnosti a ich modelovania. Snažíme sa tiež poukázať na význam aktuárskej demografie a štatistiky v poisťovníctve.

3 Metodika práce a metody skúmania

Kapitola obsahuje stručný prehľad zákonov úmrtnosti, ktoré sa používajú na modelovanie úmrtnosti nielen ľudskej populácie. Pri predstavovaní modelov budeme vo všeobecnosti vychádzať z publikácií [3], [7] a [17].

Naším cieľom je odhadnúť parametre vybraných zákonov úmrtnosti pomocou SAS Enterprise Guide. Uvedieme si tri základné metódy odhadu parametrov nelineárnej regresie (Gaussova-Newtonova metóda, Gradientná metóda, Levenbergova-Marguardtova metóda), ktoré sú v tomto analytickom softvéri dostupné. Bližšie informácie o metódach môžeme nájsť v [9].

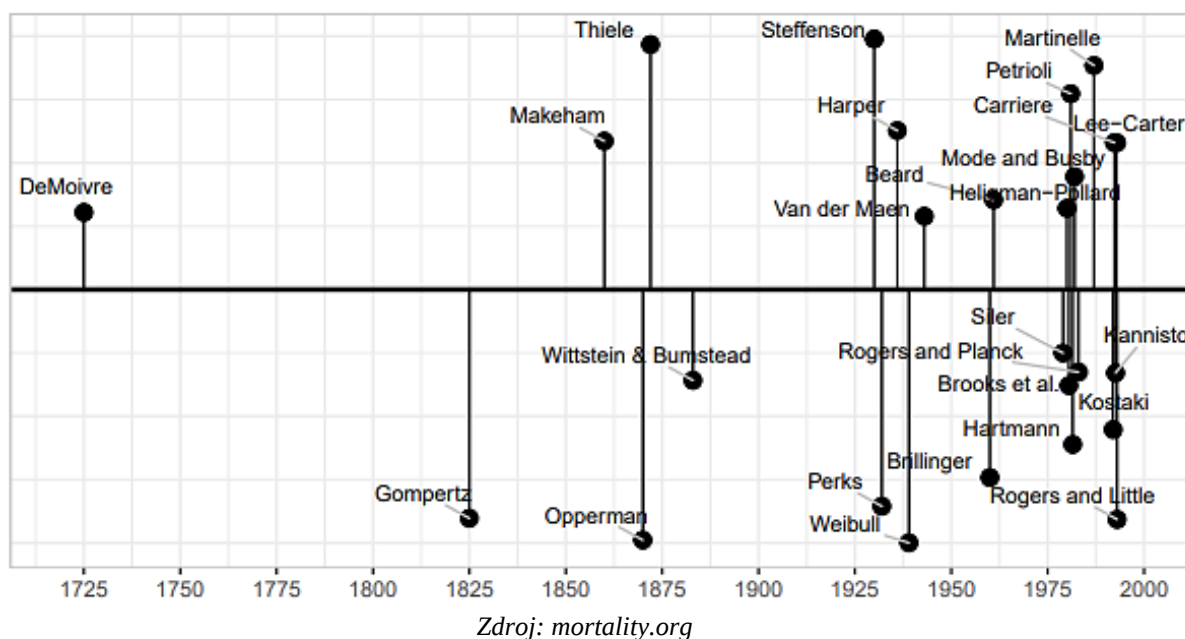
V kapitole sú zahrnuté aj základné miery kvality modelov a presnosti ich prognóz (priemerný súčet štvorcov rezíduí MSE , štandardná odchýlka rezíduí $RMSE$), ktoré sme v praktickej časti použili na vzájomné porovnanie modelov úmrtnosti. Viac informácií o týchto, ale aj ďalších mierach kvality modelov nájdeme v [10].

3.1 Zákony úmrtnosti

Zákony úmrtnosti popisujú úmrtnosť pomocou matematických spojitých funkcií veku. Najčastejšie sú zákonmi úmrtnosti vyjadrené základné úmrtnostné funkcie, a to miera úmrtnosti q_x , intenzita úmrtnosti μ_x a centrálna miera úmrtnosti m_x . Počet parametrov týchto funkcií sa obvykle pohybuje od dvoch do deviatich. Niektoré zákony sú platné pre celé vekové rozpätie, iné sa vzťahujú iba k vybraným vekovým intervalom. Uplatnenie nachádzajú pri vyrovnávaní dát a redukcii štatistických chýb, pri konštrukcii úmrtnostných tabuliek, taktiež sú vhodné na dopĺňovanie neúplných dát a veľmi často sú používané na prognózy vývoja úmrtnosti populácie.

Zákony úmrtnosti sa postupom času vyvíjali a upravovali. Na obrázku č. 4 môžeme vidieť časovú os úmrtnostných modelov do roku 2000. Vedci, ktorí sa pričínili o rozvoj danej problematiky, častokrát vychádzali z pôvodných matematických funkcií. Pridaním, respektíve úpravou parametrov boli aj vďaka lepšej informovanosti o úmrtnosti populácie schopní odstrániť niektoré nedostatky. Okrem týchto vznikali však aj nové zákony úmrtnosti, ktoré boli založené na iných predpokladoch.

Obrázok č. 4 - Časová os zákonov úmrtnosti



Najväčším problémom pretrvávajúcim až dodnes je kvalita a spoľahlivosť dát. Nezrovnalosti v dátach sa týkajú okrem údajov osôb v mladšom veku predovšetkým najstarších skupín obyvateľstva. V minulosti sa vyšších vekov dožíval len veľmi malý počet osôb, a preto existovalo nedostatočné množstvo informácií o úrovni úmrtnosti v tomto veku.

Pokiaľ sa nám teda podarí skonštruovať funkcie opisujúce úmrtnosť populácie v danom čase pre populáciu, poprípade z pohľadu poisťovne pre poisťný kmeň, a to spoločne s odhadom ich parametrov, dostávame efektívny a úsporný nástroj na najrozmanitejšie operácie s praktickými dátami o úmrtnosti populácie, ktoré sú väčšinou vyjadrované vo forme úmrtnostných tabuliek.

Zákony úmrtnosti uvedené v tabuľke č. 1 opisujú úmrtnosť v pevnom časovom bode. Skutočná úmrtnosť je však stochastická a neustále sa vyvíjajúca. Teda, zatiaľ čo modely úmrtnosti opísané v tabuľke sú statické, parametre sa musia nastavovať periodicky, aby vyhovovali zmenám v úmrtnostnom správaní populácie.

Tabuľka č. 1 – Základné parametrické funkcie ľudskej 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	$q(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + \frac{GH^x}{K}}$
<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>	1998	$\mu_x = \frac{B \cdot e^{\alpha x}}{1 + B \cdot e^{\alpha x}}$
<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

3.1.1 Konštantná intenzita úmrtnosti

Najjednoduchším matematickým modelom a zároveň najpoužívanejším modelom, ktorý zjednodušuje aktuárske výpočty, je konštantná intenzita úmrtnosti. Vzhľadom na fakt, že počas štúdia aktuárstva sme sa s konštantnou intenzitou stretli najčastejšie, popíšeme ju, aj keď nie je vhodná pre ľudskú populáciu a vysvetlíme si aj z akého dôvodu.

Pre intenzitu úmrtnosti tohto modelu platí, že je nezávislá od veku x , a teda môžeme ju zapísať ako

$$\mu_x = \lambda, \lambda = \text{konštanta}$$

Pravdepodobnosť, že x – ročná osoba sa dožije veku $(x + t)$ je

$${}_t p_x = e^{-\int_0^t \mu_{x+s} ds} = e^{-\int_0^t \lambda ds} = e^{-\lambda t}$$

Funkciu ${}_t p_x = S_x(t)$ nazývame aj miera prežitia, respektíve distribúcia prežitia. Pre distribučnú funkciu x – ročnej osoby potom platí

$$F_x(t) = 1 - {}_t p_x = 1 - e^{-\lambda t}$$

Ako sme už spomenuli, nedostatkom tohto predpokladu je nezávislosť intenzity úmrtnosti od veku, a teda aj funkcie prežitia od veku osoby. Keďže očakávame rast intenzity úmrtnosti s narastajúcim vekom, môžeme skonštatovať, že takýto zákon je očividne nevyhovujúci pre ľudskú populáciu.

3.1.2 Gompertzov zákon úmrtnosti

Jeden z prvých a zároveň najúspešnejších pokusov o matematické vyjadrenie závislosti úrovne úmrtnosti na veku učinil *Benjamin Gompertz* v roku 1825. Bol to anglický matematik a aktuár, ktorý štúdiom ľudskej populácie došiel k záveru, že pravdepodobnosť úmrtia v danom roku sa zdvojnásobuje každých 8 rokov. Zákon spočíva na všeobecnom predpoklade, že odolnosť človeka voči smrti klesá s narastajúcim vekom. Inak povedané,

úmrtnosť narastá exponenciálne v závislosti od veku a intenzita úmrtnosti je vyjadrená vzťahom

$$\mu_x = B \cdot c^x \quad (3.1)$$

kde je intenzita úmrtnosti určená dvoma reálnymi parametrami B a c , pre ktoré platí $B > 0$ a $c > 1$.

Funkciu prežitia vyjadríme vo všeobecnosti vzťahom

$$l_x = l_\alpha \cdot e^{-\int_\alpha^x \mu_s ds} \quad (3.2)$$

Ak do (3.2) dosadíme $\alpha = 0$, potom

$$l_x = l_0 \cdot e^{-\int_0^x B \cdot c^t dt} \quad (3.3)$$

kde l_0 bude určovať tzv. koreň úmrtnostnej tabuľky (najčastejšie sa stanoví 100 000 osôb). Výpočtom integrálu nachádzajúceho sa v exponente vzťahu (3.3) dostaneme

$$\int_0^x B \cdot c^t dt = B \cdot \left[\frac{c^t}{\ln c} \right]_0^x = \frac{B}{\ln c} (c^x - 1) \quad (3.4)$$

a dosadením do funkcie prežitia v (3.3) vyjadríme vzťah

$$l_x = l_0 \cdot e^{-\frac{B}{\ln c}(c^x - 1)} = l_0 \cdot \left(e^{-\frac{B}{\ln c}} \right)^{(c^x - 1)} \quad (3.5)$$

v ktorom označíme $e^{-\frac{B}{\ln c}} = g$ a $\frac{l_0}{g} = k$. Ďalšou úpravou vzťahu (3.5) dostaneme

$$l_x = k \cdot g^{c^x} \quad (3.6)$$

Aby sme vyjadrili pravdepodobnosť prežitia ${}_t p_x$, dosadíme predošlý vzťah (3.6) do definície pravdepodobnosti ${}_t p_x$ a upravíme na požadovaný tvar

$${}_t p_x = \frac{l_{x+t}}{l_x} = \frac{k \cdot g^{c^{x+t}}}{k \cdot g^{c^x}} = g^{c^x \cdot (c^t - 1)} \quad (3.7)$$

Parametre Gompertzovho zákona dostatočne spoľahlivo opisujú úmrtnostné správanie populácie. Je prekvapujúce, že zákon platí vo veľkom počte krajín, vo veľkom počte časových období a dokonca aj pre rôzne druhy organizmov, nielen pre človeka. Zatiaľ čo skutočná priemerná dĺžka života osôb sa líši z krajiny na krajinu a od zvierat'a po zviera, platí to isté všeobecné pravidlo, že pravdepodobnosť úmrtia v danom veku sa zdvojnásobi každých x rokov. Je to úžasný fakt a nikto nerozumie tomu, prečo je to pravda.

3.1.3 Makehamov zákon úmrtnosti

Už Gompertz vo svojej práci poznamenal, že okrem exponenciálne rastúcej úmrtnosti v závislosti od veku, môže existovať aj zložka úmrtnosti, ktorá je nezávislá od veku. Avšak pre analýzu životných tabuliek, ktoré boli vtedy k dispozícii, sa obmedzil len na exponenciálnu zložku úmrtnosti. O pár rokov neskôr, v roku 1860 pridal ďalší poistný matematik *William Makeham* ku Gompertzovmu vzťahu člen nezávislý od veku. Táto zložka, obvykle označovaná písmenom A , je nazývaná ako Makehamov parameter a platí $A > 0$. Vznikol vzťah, ktorý je dnes označovaný ako Makehamov zákon úmrtnosti

$$\mu_x = A + B \cdot c^x \quad (3.8)$$

Tento parameter posúva graf funkcie μ_x o konštantu A . Keďže ide o zovšeobecnenie už existujúceho Gompertzovho vzťahu, často je nazývaný ako Gompertzov-Makehamov zákon úmrtnosti. Zákon je používaný k opisu rozdelenia dĺžky života pre široké spektrum organizmov (komáre, myši, kone, vysokohorské ovce) vrátane ľudí.

Výhodou tohto zákona úmrtnosti je relatívne jednoduchá interpretácia, kompatibilita modelu spojená s pomerne malým množstvom parametrov, a tiež skutočnosť, že k odhadu parametrov nie je nevyhnutný špeciálny softvér. Na druhej strane slabou stránkou je ale skeptickosť úmrtnosti vo vysokých vekoch, ktorú mimoriadne nadhodnocuje.

Analogicky ako v predchádzajúcej časti odvodíme funkciu prežitia

$$l_x = l_\alpha \cdot e^{-\int_\alpha^x (A+B \cdot c^t) dt} \quad (3.9)$$

Pre $\alpha = 0$ dostaneme

$$l_x = l_0 \cdot e^{-\int_0^x (A+B \cdot c^t) dt} \quad (3.10)$$

Úpravou integrálu a dosadením do funkcie prežitia dospejeme k výslednému vzťahu

$$\int_0^x (A + B \cdot c^t) dt = \left[A \cdot t + B \cdot \frac{c^t}{\ln c} \right]_0^x = A \cdot x + \frac{B}{\ln c} (c^x - 1)$$

$$l_x = l_0 \cdot e^{-\left[Ax + \frac{B}{\ln c} (c^x - 1) \right]} \quad (3.11)$$

Pre $s = e^{-A}$, $g = e^{-\frac{B}{\ln c}}$ a $k = \frac{l_0}{g}$ dostaneme

$$l_x = k \cdot s^x \cdot g^{c^x} \quad (3.12)$$

Vzťah pre pravdepodobnosť, že x – ročná osoba sa dožije veku $(x + t)$ bude vyzeráť takto

$$\begin{aligned} {}_t p_x &= e^{-\int_0^t \mu_{x+s} ds} = e^{-\int_0^t (A+Bc^{x+s}) ds} = \\ &= e^{-\int_0^t A ds + Bc^x \int_0^t c^s ds} = e^{-At} g^{c^x(c^t-1)} = s^t \cdot g^{c^x \cdot (c^t-1)} \end{aligned} \quad (3.13)$$

Zaujímavé je, že Gompertz aj Makeham si uvedomovali, že ich vzťah pre intenzitu úmrtnosti (3.8) rozhodne nezodpovedá úmrtnostnému správaniu sa populácie vo všetkých vekoch. Už oni ho považovali za vzťah, ktorý najvierohodnejšie opisuje úmrtnostné správanie ľudskej populácie v rozmedzí od veku 20 rokov až po 60 rokov. V súčasnej dobe sa však používa skôr pre vyššie veku, čo môže byť vysvetlené vysokým progresom životnej úrovne, čím sa mohla vtedy určená hranica posunúť vyššie.

3.1.4 Coaleov - Kiskerov zákon úmrtnosti

Alternatívou pre Gompertzov-Makehamov zákon úmrtnosti je model Coale-Kiskera, ktorý prikladá dôraz na zmenu miery úmrtnosti v dvoch po sebe idúcich rokoch. Hlavným rozdielom týchto alternatívnych zákonov je, že Coaleov-Kiskerov zákon úmrtnosti predpokladá, že rast miery úmrtnosti klesá vo vysokých vekoch lineárne, zatiaľ čo

pri Gompertzovom-Makehamovom zákone sa domnievame, že úmrtnosť vo vysokých vekoch má konštantný priebeh rastu.

Coaleov-Kiskerov model môžeme zapísať pomocou špecifickej miery úmrtnosti m_x , ktorá je bližšie definovaná v kapitole 1.2, vzťahom

$$m_x = e^{Ax^2+Bx+C} \quad (3.14)$$

v ktorom A, B a C predstavujú parametre modelu. V rámci Slovenskej republiky sa tento zákon nikdy formálne nepoužil, no vo zvyšných európskych štátoch bol už v minulosti významnou alternatívou už niekoľkokrát spomínaného Gompertzovho-Makehamovho zákona. Jeho prednosťou je predovšetkým vyššia flexibilita, a tiež numerická ustálenosť pri odhade parametrov.

Model požaduje dodržanie dvoch predpokladov. Prvým je nevyhnutnosť toho, aby miery úmrtnosti okolo veku 85 boli vierohodné a my sme tak mohli parameter pre vek 85 počítať priamo z dát. V prípade Slovenskej republiky budeme tento predpoklad považovať za zachovaný. Ďalším predpokladom je určenie tzv. maximálne dosiahnuteľného veku, ktorý pri modelovaní úmrtnosti populácie Švédska stanovili Coale a Kisker fixne na vek 110 rokov. V prípade, že zmeníme maximálny dosiahnuteľný vek ω , špecifické miery úmrtnosti sa postupne vyrovnávajú v súlade s nádejou dožitia pri narodení jednotlivca.

3.1.5 Kannistov zákon úmrtnosti

Kannistov zákon úmrtnosti patrí do skupiny logistických modelov, ktoré sú v poslednej dobe stále viac populárnejšími pri projekcii úmrtnosti.

Vo všeobecnosti možno intenzitu úmrtnosti zapísať v tvare

$$\mu_x = \frac{Ae^{Bx}}{1 + Ae^{Bx}} \quad (3.15)$$

kde A, B reprezentujú parametre modelu, pričom parametre A, B sú nezáporné a intenzita úmrtnosti smeruje asymptoticky k číslu 1. Makeham pridal k vzťahu (3.15) nový parameter C , ktorý ako v prípade Gompertzovho-Makehamovho zákona spôsobuje posun krivky intenzity úmrtnosti o konštantu.

$$\mu_x = \frac{Ae^{Bx}}{1 + Ae^{Bx}} + C \quad (3.16)$$

Kannistov model môžeme jednoduchou úpravou prepísať do logistickej podoby nasledovne

$$\text{logit}(q_x) = \ln(a) + bx \quad (3.17)$$

a pre funkciu $\text{logit}(q_x)$ platí vzťah

$$\text{logit}(q_x) = \ln\left(\frac{q_x}{p_x}\right) = \ln(q_x) - \ln(p_x) \quad (3.18)$$

Obrovskou výhodou tohto modelu je flexibilita, keďže veľmi dobre opisuje súčasný trend úmrtnostného správania populácie, a tiež vykazuje dlhodobú spoľahlivosť až do veku $\omega = 120$. Pre vysoké veku 100 rokov a viac sú to pre krajiny vo vyspelej Európe momentálne najvhodnejšie modely.

3.1.6 Helligmanov - Pollardov zákon úmrtnosti

Helligman a Pollard navrhli v roku 1980 niekoľko parametrických modelov úmrtnosti, ktoré sú schopné modelovať trend úmrtnostného správania ľudskej populácie pre všetky veku s výnimkou nultého roka a majú biologickú interpretáciu. Práve z tohto dôvodu patria k najrozšírenejším a najviac používaným zákonom úmrtnosti v demografii. Model, ktorý predstavili ako prvý, mal 8 parametrov a môžeme ho napísať takto

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x \quad (3.19)$$

kde A, B, C, D, E, F, G, H sú odhadované parametre.

Helligmanov-Pollardov model sa delí na 3 členy. Prvý člen $A^{(x+B)^C}$ predstavuje úmrtnosť detskej populácie, v ktorej parameter A predstavuje dojčenskú úmrtnosť a parameter B odzrkadľuje mieru úmrtnosti detí vo veku 1 roka. Parameter C odráža mieru úmrtnosti až po začiatok dospelosti. Všetky tri parametre A, B, C nadobúdajú hodnoty z intervalu $(0,1)$.

Druhý člen $De^{-E(\ln x - \ln F)^2}$ reprezentuje úmrtnosť hlavne z dôvodu nehôd, v dôsledku čoho dostal aj názov havarijný hrb. Parameter D meria výšku havarijného hrbu, pričom nadobúda hodnoty z intervalu $\langle 0,1 \rangle$. E udáva jeho šírku a jeho hodnoty sú kladné. Parameter F predstavuje umiestnenie havarijného hrbu na vekovej osi.

Posledný člen tvorený Gompertzovým zákonom úmrtnosti reprezentuje úmrtnostné správanie dospelých ľudí. Hodnoty parametrov sa nachádzajú v intervale $\langle 0,1 \rangle$, respektíve $\langle 0, \infty \rangle$. Parameter G pritom meria základnú úroveň úmrtnosti dospelých a H predstavuje mieru nárastu úmrtnosti alebo inak sklon Gompertzovej krivky.

Helligman a Pollard svoj prvý model ďalej upravovali, až kým neprišli na jeho posledný tvar, ktorý vyzerá takto

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^{xK}} \quad (3.20)$$

Vhodnosť ich zákonov úmrtnosti cez všetky vekové kategórie sa osvedčila na empirických pozorovaniach úmrtnosti populácie Spojených štátov amerických, a tiež v Austrálii a sú znovu používané na modelovanie úmrtnosti ako alternatívny spôsob ku Gompertzovmu-Makehamovmu modelu.

3.1.7 CoDe zákon úmrtnosti

CoDe model patrí medzi najmladšie zákony úmrtnosti. Umožňuje modelovať úmrtnostné správanie populácie v celom vekovom rozpätí. Názov dostal na základe anglických výrazov *Compression* (zhustovanie) a *Delay* (oneskorenie), pretože pomáha rozlišovať účinky zhustovania a oneskorenia úmrtnosti ľudskej populácie.

CoDe model obsahuje 10 parametrov, z ktorých päť sa v čase mení. Týchto päť interpretovateľných časovo závislých parametrov odráža oneskorenie úmrtnosti a zmeny v úmrtnosti v detstve, v dospievaní, v strednom veku a starobe.

Formulácia tohto modelu, kde $I(*)$ je takzvaná indikátorová funkcia, b_1 určuje sklon krivky pravdepodobnosti úmrtia pred vekom $(M - h)$, b_2 medzi vekom $(M - h)$ a M a b_3 nad hodnotou modálneho veku úmrtia M , je nasledovná

$$q_x = \frac{A}{x+B} + \frac{ae^{(x-m)}}{1+e^{(x-m)}} + I(x \leq M-h) \left[\frac{b_1 e^{b_1(x-M)}}{1 + \frac{b_1}{g} e^{b_1(x-M)}} \right] + \quad (3.21)$$

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

Parametre c_1 a c_2 sú konštanty, ktoré sú vyjadrené pomocou nasledujúcich vzťahov

$$c_1 = \frac{b_1 e^{b_1(-h)}}{1 + \frac{b_1}{g} e^{b_1(-h)}} - \frac{b_2 e^{b_2(-h)}}{1 + \frac{b_2}{g} e^{b_2(-h)}} \quad (3.22)$$

$$c_2 = \frac{b_2}{1 + \frac{b_2}{g}} + c_1 - \frac{b_3}{1 + \frac{b_3}{g}} \quad (3.23)$$

Výsledný model CoDe sa teda skladá z piatich členov, ktoré predstavujú úmrtnosť v po sebe nasledujúcich etapách života. Prvé dve sa používajú na modelovanie úmrtnosti v mladom veku a v dospievaní. Zvyšné tri zahŕňajú výrazy, ktoré rozlišujú úmrtnosť v skorej dospelosti, neskorej dospelosti a v starobe.

Model úmrtnosti pre nízky vek je opísaný pomocou jednoduchej funkcie, ktorá predpokladá inverzný vzťah medzi úmrtnosťou a vekom $\frac{A}{x+B}$. Parameter A reprezentuje úroveň detskej úmrtnosti a B je ovplyvnený mierou poklesu úmrtnosti v detstve. Vzhľadom na fakt, že efekt zmien v hodnote B v čase je relatívne malý, zmeny v úmrtnosti v mladšom veku možno pripísať najmä zmenám hodnoty parametra A . Interpretácia tohto parametra je potom triviálna, nižšia úmrtnosť detí zodpovedá nižšej hodnote parametra A .

V rokoch dospievania narastá pravdepodobnosť úmrtnosti pomerne rýchlo a následne sa ustáli pred začiatkom dospelosti. Toto sa zvyčajne nazýva, ako už spomínaný, havarijný hrb úmrtnosti. V CoDe modeli je tento stav opísaný pomocou logistickej funkcie $\frac{ae^{(x-m)}}{1+e^{(x-m)}}$, kde m určuje vek, v ktorom je nárast v rokoch dospievania viditeľne najprudší a parameter a predstavuje zvyšnú úmrtnosť spôsobenú ďalšími faktormi. Pre naše dáta je charakteristický prudký nárast úmrtnosti vo veku 16 rokov, čo sa zhoduje s krajinami ako Japonsko, Francúzsko a Dánsko, na ktorých bol tento model testovaný.

Keďže máme dospelosť rozdelenú do troch častí (skorá dospelosť, neskorá dospelosť a staroba), hodnotu h zvolíme 30 rokov. Priemerná hodnota modálneho veku v našich dátach je okolo 80 rokov, čo znamená, že parameter b_1 zaznamenáva úmrtnosť približne 50 rokov (skorá dospelosť), b_2 vplýva na úmrtnosť medzi 50 až 80 rokov (neskorá dospelosť) a parameter b_3 vyjadruje úmrtnosť nad 80 rokov (staroba). Parameter g predstavuje približnú hodnotu pravdepodobnosti úmrtia nad 110 rokov, ktorá podľa vedcov z [18] je rovnaká a rovná sa zhruba 0,7. My túto hodnotu prevezmeme, a teda výsledný model po doplnení do vzťahu (3.21) bude mať tvar

$$q_x = \frac{A}{x+B} + \frac{ae^{(x-16)}}{1+e^{(x-16)}} + I(x \leq 50) \left[\frac{b_1 e^{b_1(x-80)}}{1 + \frac{b_1}{0,7} e^{b_1(x-80)}} \right] + \quad (3.24)$$

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

3.2 Odhad parametrov nelineárnych regresných modelov

Keďže zákony úmrtnosti predstavujú pokusy ako opísať empirické pozorovania, sú úzko spojené so štatistickými technikami analýzy údajov o úmrtnosti a v dnešnej dobe by boli označené ako parametrické modely nelineárnej regresnej analýzy.

Základná myšlienka nelineárnej regresie je totožná s myšlienkou regresnej analýzy. Cieľom regresnej analýzy je nájsť vzťah medzi závislou a jednou alebo viacerými nezávislými premennými regresného modelu. Ak závislú premennú vyjadríme pomocou lineárneho vzťahu, hovoríme o lineárnej regresii. Ak závislosť nie je lineárna, jej priebeh sa vyjadrí vhodnou nelineárnou (polynomickou, logaritmickou, exponenciálnou, atď.) regresnou funkciou.

V našej práci sa zameriame na nelineárnu regresiu a modelovanie úmrtnosti slovenskej populácie v závislosti od veku. V tejto časti si opíšeme najpoužívanejšie iteračné metódy odhadu parametrov nelineárneho regresného modelu, ktoré sú dostupné v softvéri SAS Enterprise Guide, a tiež spôsoby výberu modelu, ktorý najlepšie vystihuje danú závislosť pomocou mier presnosti odhadu parametrov regresného modelu a ich prognóz.

3.2.1 Gaussova - Newtonova metóda

Gaussova-Newtonova metóda patrí medzi iteračné metódy využívajúce deriváciu funkcie. Jej hlavnou nevýhodou je, že občas konverguje veľmi pomaly, čo vedie k veľkému počtu iterácií. V určitých prípadoch môže dokonca oscilovať, čo sa prejavuje poklesom v jednej iterácii a následným rastom súčtu štvorcov v nasledujúcej iterácii. Rast a pokles súčtu štvorcov sa opakuje až pokým sa hodnota nestabilizuje.

Gaussova-Newtonova metóda vychádza z lineárnej aproximácie funkcie $f(\mathbf{x}_i, \boldsymbol{\theta})$ Taylorovým rozvojom v bode $\boldsymbol{\theta} = (\theta_1^0, \theta_2^0, \theta_3^0, \dots, \theta_k^0)$ so zachovaním len lineárnych podmienok. To nám dáva

$$f(\mathbf{x}_i, \boldsymbol{\theta}) = f(\mathbf{x}_i, \boldsymbol{\theta}^0) + \sum_{j=1}^k \left\{ \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \cdot (\theta_j - \theta_j^0) \right\} + \varepsilon_i \quad (3.25)$$

Ak označíme

$$\begin{aligned} f_i^0 &= f(\mathbf{x}_i, \boldsymbol{\theta}^0) \\ \beta_j^0 &= \theta_j - \theta_j^0 \\ J_{ij}^0 &= \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_j} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \end{aligned}$$

potom sme schopní nelineárny regresný model prepísať do tvaru

$$y_i - f_i^0 = \sum_{j=1}^k \beta_j^0 J_{ij}^0 + \varepsilon_i \quad (3.26)$$

kde $i = 1, 2, \dots, n$. Touto úpravou sme získali lineárny regresný model. Vektor $\boldsymbol{\theta}^0$ zvyčajne nazývame počiatočná hodnota parametrov. Rovnicu vieme tiež zapísať ako

$$\mathbf{y}^0 = \mathbf{J}^0 \boldsymbol{\beta}^0 + \boldsymbol{\varepsilon}^0 \quad (3.27)$$

a teda odhadom vektora $\boldsymbol{\beta}^0$ je vektor $\hat{\boldsymbol{\beta}}^0$, pre ktorý platí

$$\hat{\boldsymbol{\beta}}^0 = (\mathbf{J}^{0T} \mathbf{J}^0)^{-1} \cdot \mathbf{J}^{0T} \cdot \mathbf{y}^0 = (\mathbf{J}^{0T} \mathbf{J}^0)^{-1} \cdot \mathbf{J}^{0T} \cdot (\mathbf{y} - \mathbf{f}^0) \quad (3.28)$$

a $\mathbf{J}^{0T}$ je transponovaná matica k matici $\mathbf{J}^0$.

Vzhľadom na to, že $\boldsymbol{\beta}^0 = \boldsymbol{\theta} - \boldsymbol{\theta}^0$, ďalšiu iteráciu, čiže odhadnutý vektor parametrov $\hat{\boldsymbol{\theta}}^1$ môžeme definovať takto

$$\hat{\boldsymbol{\theta}}^1 = \hat{\boldsymbol{\beta}}^0 + \hat{\boldsymbol{\theta}}^0 \quad (3.29)$$

Postup vyjadrenia vzťahu pre ďalšie vektory parametrov je obdobný ako v (3.29), preto všeobecný predpis iterácií vieme zapísať pomocou (3.28) týmto spôsobom

$$\hat{\boldsymbol{\theta}}^{k+1} = \hat{\boldsymbol{\beta}}^k + \hat{\boldsymbol{\theta}}^k = (\mathbf{J}^{kT} \mathbf{J}^k)^{-1} \cdot \mathbf{J}^{kT} \cdot (\mathbf{y} - \mathbf{f}^k) + \hat{\boldsymbol{\theta}}^k \quad (3.30)$$

v ktorom

$$\hat{\boldsymbol{\theta}}^k = [\theta_1^k, \theta_2^k, \theta_3^k, \dots, \theta_n^k]$$

$$\mathbf{J}^k = [J_{ij}^k]$$

$$\mathbf{f}^k = [f_1^k, f_2^k, f_3^k, \dots, f_n^k]^T$$

Iteračný postup pokračuje, až pokiaľ nebude splnená podmienka konverencie, teda pokiaľ neplatí

$$\left| \frac{\hat{\theta}_j^{k+1} - \hat{\theta}_j^k}{\hat{\theta}_j^k} \right| < \delta \quad (3.31)$$

kde $j = 1, 2, \dots, k$ a δ je vopred dané malé kladné číslo, zvyčajne 10^{-6} .

Gaussova-Newtonova metóda relatívne dobre konverguje, pokiaľ platí aspoň jedna z nasledujúcich podmienok:

- Rezíduá sú malé, tzn. počiatočný odhad parametrov bol veľmi dobrý.
- Regresná funkcia $f(\mathbf{x}, \boldsymbol{\beta})$ je takmer lineárna.
- Rezíduá majú striedavé znamienka.

V prípade, že metóda nekonverguje alebo konverguje veľmi pomaly, môžeme použiť rôzne modifikácie Gaussovej-Newtonovej metódy, poprípade tzv. kvázi-newtonovské metódy alebo tiež Gillovu-Murrayho metódu. Viac o modifikáciách Gaussovej-Newtonovej metódy je možné zistiť napríklad v [14].

3.2.2 Gradientná metóda

Gradientná metóda (*Metóda najstrmšieho zostupu*) určuje globálne minimum funkcie súčtu štvorcov reziduálnych odchýlok (Sum of Square Errors) $S(\boldsymbol{\theta})$ priamou minimalizáciou. Jej cieľom je dostať sa z počiatočného východiskového bodu $\boldsymbol{\theta}^0$ v smere vektora s prvkami danými parciálnymi deriváciami funkcie súčtu štvorcov reziduálov vzhľadom na prvky vektora parametrov $\boldsymbol{\theta}$. Je to typ iteračnej metódy využívajúcej gradient funkcie.

Funkcia S rapídne klesá v smere záporného gradientu. V prípade, že teda máme hodnotu $S(\boldsymbol{\theta})$ funkcie $f(\mathbf{x}_i, \boldsymbol{\theta})$, potom smer funkcie kolíše v smere

$$\frac{1}{2} \frac{\partial S(\boldsymbol{\theta})}{\partial \theta^i} = -\mathbf{J}^T [\mathbf{y} - f(\boldsymbol{\theta}^i)] \quad (3.32)$$

kde matica $\mathbf{J}$ je Jakobián funkcie $f(\mathbf{x}_i, \boldsymbol{\theta})$.

Ďalšiu iteráciu vektora parametrov $\boldsymbol{\theta}$ môžeme vypočítať dosadením do všeobecného predpisu takto

$$\boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^k + \lambda \mathbf{J}^T \cdot [\mathbf{y} - f(\boldsymbol{\theta}^k)] \quad (3.33)$$

kde λ reprezentuje skalár zvolený takým spôsobom, aby platilo

$$S(\boldsymbol{\theta}^k + \lambda \mathbf{J}^T [\mathbf{y} - f(\boldsymbol{\theta}^k)]) < S(\boldsymbol{\theta}^k) \quad (3.34)$$

Metóda najstrmšieho zostupu sa častokrát používa na stanovenie oblasti s najväčšou pravdepodobnosťou dosiahnutia optima pri postupe z východiskového počiatočného odhadu optimálnych podmienok. Medzi najväčšie nevýhody gradientnej metódy pri riešení nelineárnej regresnej úlohy patrí, že niekedy môže veľmi pomaly konvergovať. Najstrmší zostup spravidla funguje najlepšie vtedy, ak je počiatočný východiskový bod ďalej od optimálneho bodu. Aj keď sa iteráciami dostávame do blízkosti optima, procedúra bude vykazovať kratšie kroky spolu s náhodným správaním. To je už spomínaný problém konvergenzie gradientnej metódy.

V porovnaní s Gaussovou-Newtonovou metódou, ktorá je opísaná v predchádzajúcej časti 3.2.1, metóda najstrmšieho zostupu funguje efektívnejšie v prvých iteráciách, v ktorých aj v prípade zlých vstupných odhadov dokáže rýchlejšie nájsť zodpovedajúci smer k minimu funkcie súčtu štvorcov. Naopak v konečných iteráciách pracuje horšie, pretože v oblasti minima funkcie štvorcov častokrát konverguje pomaly, alebo tiež osciluje.

3.2.3 Levenbergova - Marquardtova metóda

Ďalšou populárnou metódou odhadu parametrov nelineárneho regresného modelu je tiež modifikácia jednoduchšej linearizácie. Ako prvý publikoval tento algoritmus odhadu parametrov K. Lavenberg, a to v roku 1944 a následne bol znovuobjavený v roku 1963 D. Marquardtom. Z času na čas sa táto metóda označuje pojmom Marquardtov kompromis, pretože sa usiluje o sklbenie predností už vyššie spomenutých metód – Gaussovej-Newtonovej metódy a Gradientnej metódy.

Začíname rovnako ako pri predchádzajúcich metódach nultou iteráciou, resp. počiatočným východiskovým vektorom parametrov θ^0 . Pri nasledujúcej iterácii θ^1 vychádzame zo vzťahu

$$\theta^1 = \theta^0 + \delta \quad (3.35)$$

kde δ je parameter. Na určenie parametra δ sú funkcie $f(\mathbf{x}_i, \theta + \delta)$ aproximované linearizačným vzťahom

$$f(\mathbf{x}_i, \theta + \delta) \approx f(\mathbf{x}_i, \theta) + J\delta \quad (3.36)$$

v ktorom $J = \frac{\partial f(\mathbf{x}_i, \theta)}{\partial \theta_j}$ predstavuje gradient funkcie $f(\mathbf{x}_i, \theta)$ vzhľadom na parameter θ .

V minimálnej sume štvorcov $S(\theta)$ bude gradient vzhľadom na vektor parametrov θ nulový. Prostredníctvom aproximácie (3.36) vyjadríme

$$S(\theta + \delta) \approx \|\mathbf{y} - f(\theta) - J\delta\|^2 \quad (3.37)$$

Ak funkciu zderivujeme podľa δ a položíme ju rovnú nule, dostaneme

$$(\mathbf{J}^T \mathbf{J}) \delta = \mathbf{J}^T [\mathbf{y} - f(\boldsymbol{\theta})] \quad (3.38)$$

Kde $\mathbf{J}$ je Jakobián, ktorého i -ty riadok označujeme J_i a f spolu s $\mathbf{y}$ reprezentujú vektory i -tej zložky funkcie $f(\mathbf{x}_i, \boldsymbol{\theta})$. Takto dostaneme sústavu lineárnych rovníc, ktorej riešením je parameter δ (vid'[15]).

Levenberg sa pričínal k rozvoju tejto metódy tak, že nahradil rovnicu (3.38), tzv. tlmeným variantom

$$(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I}) \delta = \mathbf{J}^T [\mathbf{y} - f(\boldsymbol{\theta})] \quad (3.39)$$

v ktorom $\mathbf{I}$ je jednotková matica, ktorá udáva prírastok δ k odhadovanému vektoru parametrov $\boldsymbol{\theta}$.

Tlmený faktor $\lambda \geq 0$ je prispôsobený vo všetkých iteráciách. V prípade, že dochádza k rapidnému poklesu súčtu štvorcov $S(\boldsymbol{\theta})$, využívajú sa menšie hodnoty λ a algoritmus sa približuje ku Gaussovej-Newtonovej metóde výpočtu opísanej v časti 3.2.1. V opačnom prípade, keď iteráciou neklesá súčet štvorcov dostatočne rýchlo, λ sa zväčšuje a algoritmus prepne na výpočet pomocou metódy najstrmšieho zostupu opísanej v podkapitole 3.2.2.

Nevýhodou Levenbergovho postupu je, že pokiaľ je δ priveľmi veľké, inverziu matice $(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})$ nemá zmysel použiť. Preto Marquardt rovnicu (3.39) pozmenil tak, že namiesto jednotkovej matice $\mathbf{I}$ dosadil do vzťahu diagonálnu maticu súčinu $\mathbf{J}^T \mathbf{J}$. Napokon výsledný Levenbergov-Marquardtov algoritmus je vyjadrený nasledujúcim vzťahom

$$(\mathbf{J}^T \mathbf{J} + \lambda \cdot \text{diag}(\mathbf{J}^T \mathbf{J})) \delta = \mathbf{J}^T [\mathbf{y} - f(\boldsymbol{\theta})] \quad (3.40)$$

kde, ako sme už spomenuli, $\text{diag}(\mathbf{J}^T \mathbf{J})$ reprezentuje diagonálnu maticu obsahujúcu prvky diagonály súčinu matíc $\mathbf{J}^T \mathbf{J}$ (vid'[16]).

Rozličné počítačové softvéry stanovujú hodnotu λ rôzne, napríklad PROC NLIN v SAS Enterprise Guide sa jej hodnota začína na $\lambda = 10^{-8}$. Séria krokov použitím metódy pokus-omyl je sprevádzaná v každej iterácii s λ , ktorá je opakovane násobená desiatkou, až pokiaľ nie je splnená podmienka

$$S(\hat{\theta}^{k+1}) < S(\hat{\theta}^k) \quad (3.41)$$

Stratégiou je zachovať čo najmenšie dosiahnuteľné λ , medzitým čo zabezpečujeme, aby suma štvorcov reziduálnej odchýlky sa znižovala každou iteráciou. Výsledný vektor prírastkov vypočítaný pomocou tohto algoritmu sa zvyčajne nachádza medzi Gaussovým-Newtonovým vektorom hodnôt a vektorom gradientnej metódy.

V súčasnosti existujú nové modifikované postupy založené na Levenbergovom-Marquardtovom algoritme spočívajúce na rozličných ohraničeníach individuálnych krokov, tzv. trust-region metódy, ktoré napomáhajú k vyriešeniu nedostatkov týchto metód. Pre naše potreby sú však vyššie opísané metódy postačujúce, podrobnosti k trust-region metódam môžeme nájsť napríklad v [9].

3.3 Meranie kvality modelov a presnosti ich prognóz

V štatistike je pojem *priemerný súčet štvorcov rezíduí* dôležitým kritériom, ktoré sa používa na meranie kvality odhadov regresných modelov a ich prognóz. Priemerný súčet štvorcov rezíduí, skrátene *MSE*, je veľmi dôležitá hodnota potrebná na posúdenie presnosti štatistického odhadu.

Pod pojmom reziduá v regresnej analýze rozumieme rozdiely medzi skutočnou hodnotou a odhadnutou hodnotou. Priemerný súčet štvorcov rezíduí môžeme nazývať rizikovou funkciou, ktorá zodpovedá očakávanej hodnote straty štvorcovej chyby. Tento rozdiel alebo strata vzniká v dôsledku náhodnosti alebo kvôli tomu, že odhady neuvádzajú informácie, ktoré by mohli poskytnúť presnejší odhad.

Predpokladajme, že $\hat{\mathbf{y}}_i$ je vektor označujúci očakávané hodnoty závislej premennej a $\mathbf{y}_i$ je vektor reprezentujúci skutočné hodnoty závislej premennej. Potom vzťah pre priemerný súčet štvorcov rezíduí bude mať tvar

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{\mathbf{y}}_i - \mathbf{y}_i)^2 \quad (3.42)$$

Vo všeobecnosti, ak θ je nejaký neznámy parameter a $\hat{\theta}$ predstavuje zodpovedajúci odhad tohto parametra, potom vzorec pre priemerný súčet štvorcov rezíduí daného odhadu je

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] \quad (3.43)$$

Základné charakteristiky priemerného súčtu štvorcov rezíduí sú znázornené nižšie:

- Ak $MSE = 0$, znamená to, že odhadovaná hodnota sa rovná parametru θ . Táto situácia je ideálna, ale v praxi prakticky nemožná.
- Čím menšia je hodnota MSE , tým je model presnejší, pretože menšie hodnoty znamenajú menšie chyby.
- Hodnoty MSE sú vhodné na porovnávanie dvoch alebo viacerých štatistických modelov. Sú nástrojom na meranie presnosti odhadu na daných údajoch.
- V mnohých metódach krokovej regresie sa MSE využíva na stanovenie počtu premenných, ktoré majú byť zaradené do modelu z daného súboru pozorovaní.
- Analýza rozptylu, t. j. ANOVA, odhaduje priemerný súčet štvorcov rezíduí ako súčasť štatistickej analýzy.

Štandardná odchýlka rezíduí ($RMSE$) je ďalší spôsob kvantifikácie rezíduí. Ide o veľmi často používané meranie rozdielov medzi odhadovanými hodnotami a skutočnými pozorovanými hodnotami. $RMSE$ je definovaná ako druhá odmocnina rozdielov medzi vyrovnanými hodnotami a pozorovanými hodnotami. Jednotlivé rozdiely v tomto výpočte sú známe ako rezíduá. $RMSE$ odhaduje veľkosť chýb. Je to dobrá miera presnosti, ktorá sa často používa na porovnanie kvality regresných modelov.

Predpokladajme, že $\hat{\theta}$ predstavuje odhad daného parametra θ , potom $RMSE$ je v skutočnosti druhá odmocnina priemerného súčtu štvorcov rezíduí MSE . Platí

$$RMSE(\hat{\theta}) = \sqrt{MSE(\hat{\theta})} \quad (3.44)$$

$$RMSE(\hat{\theta}) = \sqrt{E[(\hat{\theta} - \theta)^2]} \quad (3.45)$$

Existuje ešte mnoho mier, ktorými vieme posúdiť kvalitu modelu a presnosť jeho prognózy. Spomenieme si napríklad priemernú absolútnu chybu rezíduí (MAE), priemernú absolútnu percentuálnu chybu rezíduí ($MAPE$), relatívnu chybu nazývanú tiež Theilov koeficient a iné.

3.4 Softvér SAS Enterprise Guide

SAS (Statistical Analysis Software) je moderný systém spracovania informácií vyvinutý v roku 1974. V minulosti sa využíval hlavne na analýzu štatistických údajov. Jeho súčasná podoba je súborom mnohých modulov s veľmi širokým rozsahom použitia a jednoduchým používaním. Viac informácií o systéme SAS a jeho moduloch nájdete v [20]. Na internete stojí za to navštíviť webovú stránku www.sas.com, ktorá obsahuje informácie o produktoch, konferenciách organizovaných užívateľskými skupinami, ako aj mnoho programov poskytovaných ich tvorcami.

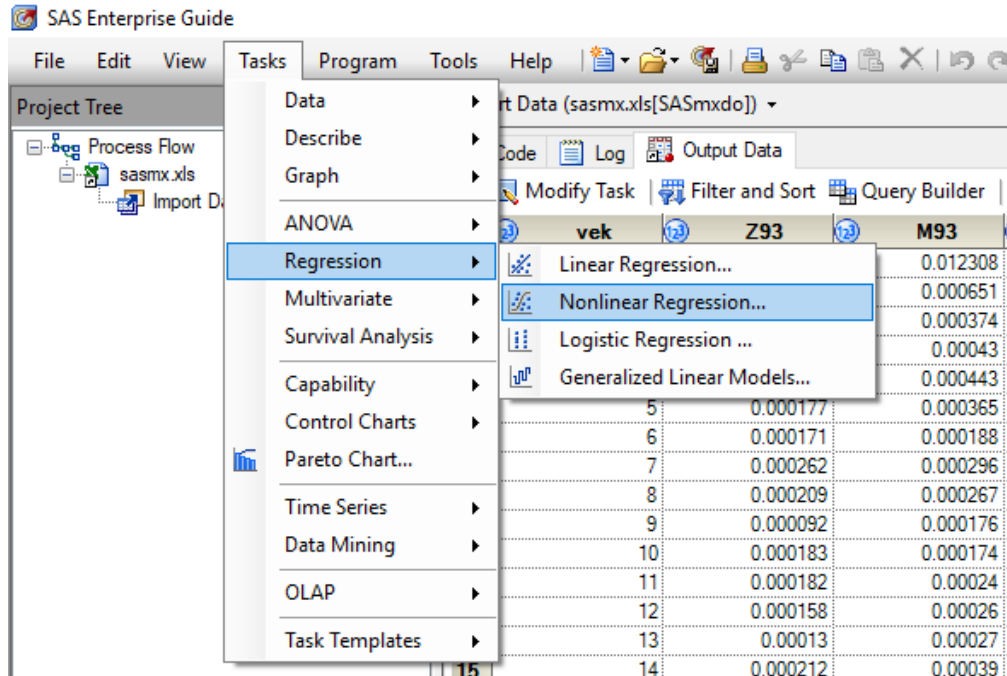
SAS Enterprise Guide je jedným zo samostatných nástrojov SAS Foundation, ktoré môžu analytici využiť pri svojej každodennej práci s dátami. Je to grafické používateľské rozhranie, ktoré používa mnoho spoločností v oblasti telekomunikácií, bankovníctva, energetiky, farmácie a mimo nej. K práci s dátami nepotrebuje používateľ nutne znalosť programovacieho jazyku, stačí mu poznať základné nástroje tohto prostredia. Cez hlavný panel nástrojov sa dokáže dostať k základným príkazom.

SAS Enterprise Guide 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 cez OLE DB pre OLAP,
- vizualizácia OLAP dát vo forme diagramov,
- prepojenie s údajmi z iných zdrojov prostredníctvom mechanizmu "pass through" – možnosť pripojenia k údajom v databáze Oracle, MSSQL alebo Teradata, zobrazenie tabuliek, dotaz na tabuľky z inej databázy (v závislosti od licencie),

- automatizácia správ a výpočtov na serveri alebo prostredníctvom programu Microsoft Windows Scheduler.

Obrázok č. 5 - Nelineárna regresná analýza v SAS Enterprise Guide



Zdroj: vlastný výstrižok zo softvéru SAS Enterprise Guide

Nelineárna regresná analýza je taktiež súčasťou rozhrania SAS Enterprise Guide. Na obrázku č. 5 vidíme jej umiestnenie v softvérovom prostredí. Dostupné metódy odhadu parametrov regresnej analýzy sú Gaussova-Newtonova metóda, Gradientná metóda a Levenbergova-Marquardtova metóda, ktoré sú priblížené v kapitole 3.2. Ponúka široké spektrum možností nastavenia výstupu. Hlavnou nevýhodou nelineárnej regresnej analýzy v SAS Enterprise Guide je, že obsahuje len niekoľko predvolených tvarov nelineárnych modelov, ktoré nekorešpondujú so vzťahmi zákonov úmrtnosti opísanými v 3.1, ktoré sú predmetom tejto diplomovej práce. Súčasťou softvéru SAS je však okrem interaktívnych modulov aj vlastný programovací jazyk. Aj v tomto prostredí sa dajú upravovať kódy a písať vlastné programy pomocou programovacieho jazyka SAS v časti Preview Code, ktorý tu generuje softvér automaticky.

Ukážka programu nelineárnej regresie v SAS Enterprise Guide:

```
/* -----
Code generated by Jana Kútiková

Input Data: Local:WORK.SASMX
```

```

Server: Local
----- */
%_eg_conditional_dropds(WORK.SORTTempTableSorted,
                        WORK.TMP0TempTableForPlots);
/* -----
Sort data set Local:WORK.SASMX
----- */

PROC SQL;
    CREATE VIEW WORK.SORTTempTableSorted AS
        SELECT T.S2017, T.vek
        FROM WORK.SASMX as T
;
QUIT;
TITLE;
TITLE1 "Nonlinear Regression";
TITLE2 "Results";
FOOTNOTE;
FOOTNOTE1 "Generated by the SAS System (&SASSERVERNAME, &SYSSCPL)
on %TRIM(%QSYSFUNC(DATE()), NLDATE20.) at
%TRIM(%SYSFUNC(TIME()), TIMEAMPM12.)";
PROC NLIN DATA=WORK.SORTTempTableSorted
    METHOD=MARQUARDT
    MAXITER=100
    CONVERGE=1E-05
    SINGULAR=1E-08
    MAXSUBIT=30
;
MODEL S2017 = a + b * c ** vek
;
PARMS
a=0.0004
b=0.0000043
c=0.136
;

OUTPUT OUT=WORK.TMP0TempTableForPlots
    H=_h1
    L95M=_l95m
    L95=_l95
    PARMS=a b c
    PREDICTED=_predicted1
    RESIDUAL=_residual1
    SSE=_sse1
    STDI=_stdi1
    STDP=_stdp1
    STDR=_stdr1
    STUDENT=_rstudent1
    U95M=_u95m
    U95=_u95
    WEIGHT=_weight1
;
RUN;

TITLE;
TITLE1 "Regression Analysis Plots";

```

```

PROC SORT DATA=WORK.TMP0TempTableForPlots
  OUT=WORK.TMP0TempTableForPlots;
  BY vek;
RUN;
AXIS1 MAJOR=(NUMBER=5) WIDTH=1;
AXIS2 OFFSET=(10 PCT) WIDTH=1;
AXIS3 MAJOR=(NUMBER=5) OFFSET=(5 PCT) WIDTH=1;
PROC GPLOT DATA=WORK.TMP0TempTableForPlots
  ;
  WHERE S2017 IS NOT MISSING AND
        vek IS NOT MISSING;

/* ***** PREDICTED plots ***** */

  TITLE4 "Observed S2017 by Predicted S2017";
  SYMBOL1 C=BLUE V=DOT HEIGHT=2PCT INTERPOL=NONE L=1 W=1;
  LABEL _predicted1 = "Predicted S2017";
  WHERE S2017 IS NOT MISSING AND _predicted1 IS NOT MISSING;
  PLOT S2017 * _predicted1 /
        VAXIS=AXIS1 VMINOR=0 HAXIS=AXIS3 HMINOR=0
        DESCRIPTION = "Observed S2017 by Predicted S2017"
  ;
RUN;

  TITLE4 "Observed S2017 by vek";
  SYMBOL1 C=BLUE V=DOT HEIGHT=2PCT INTERPOL=NONE L=1 W=1;
  SYMBOL2 C=BLUE V=NONE INTERPOL=JOIN L=1 W=1;
  SYMBOL3 C=BLUE V=NONE INTERPOL=JOIN L=1 W=1;
  SYMBOL4 C=BLUE V=NONE INTERPOL=JOIN L=1 W=1;
  WHERE S2017 IS NOT MISSING AND vek IS NOT MISSING;
  PLOT (S2017 _predicted1 _l95 _u95) * vek /
        VAXIS=AXIS1 VMINOR=0 HAXIS=AXIS3 HMINOR=0
        DESCRIPTION = "Observed S2017 by vek"
        OVERLAY
  ;
RUN;

/* ***** RESIDUAL plots ***** */

  TITLE4 "Residuals of S2017 by Predicted S2017";
  SYMBOL1 C=BLUE V=DOT HEIGHT=2PCT INTERPOL=NONE L=1 W=1;
  LABEL _residual1 = "Residual of S2017";
  LABEL _predicted1 = "Predicted S2017";
  WHERE _residual1 IS NOT MISSING AND _predicted1 IS NOT
        MISSING;
  PLOT _residual1 * _predicted1 /
        VAXIS=AXIS1 VMINOR=0 HAXIS=AXIS3 HMINOR=0 VREF=0
        DESCRIPTION = "Residuals of S2017 by Predicted S2017"
  ;
RUN;

  TITLE4 "Studentized Residuals of S2017 by Predicted S2017";
  SYMBOL1 C=BLUE V=DOT HEIGHT=2PCT INTERPOL=NONE L=1 W=1;
  LABEL _rstudent1 = "Studentized Residual of S2017";
  LABEL _predicted1 = "Predicted S2017";
  WHERE _rstudent1 IS NOT MISSING AND _predicted1 IS NOT

```

```

MISSING;
PLOT _rstudent1 * _predicted1 /
    VAXIS=AXIS1 VMINOR=0 HAXIS=AXIS3 HMINOR=0 VREF=0
    DESCRIPTION = "Studentized Residuals of S2017 by
                    Predicted S2017"
;
RUN;

TITLE4 "Residuals of S2017 by vek";
SYMBOL1 C=BLUE V=DOT HEIGHT=2PCT INTERPOL=NONE L=1 W=1;
LABEL _residual1 = "Residual of S2017";
WHERE _residual1 IS NOT MISSING AND vek IS NOT MISSING;
PLOT _residual1 * vek /
    VAXIS=AXIS1 VMINOR=0 HAXIS=AXIS3 HMINOR=0 VREF=0
    DESCRIPTION = "Residuals of S2017 by vek"
;
RUN;

TITLE4 "Studentized Residuals of S2017 by vek";
SYMBOL1 C=BLUE V=DOT HEIGHT=2PCT INTERPOL=NONE L=1 W=1;
LABEL _rstudent1 = "Studentized Residual of S2017";
WHERE _rstudent1 IS NOT MISSING AND vek IS NOT MISSING;
PLOT _rstudent1 * vek /
    VAXIS=AXIS1 VMINOR=0 HAXIS=AXIS3 HMINOR=0 VREF=0
    DESCRIPTION = "Studentized Residuals of S2017 by vek"
;
RUN;

SYMBOL;
QUIT;

/* -----
End of task code.
----- */

RUN; QUIT;
%_eg_conditional_dropds(WORK.SORTTempTableSorted,
                        WORK.TMP0TempTableForPlots);
TITLE; FOOTNOTE;

```

Hore vidíme ukážku programu nelineárnej regresnej analýzy pre rok 2017 pri zanedbaní pohlavia pomocou Gompertzovho-Makehamovho zákona úmrtnosti, ktorého intenzita úmrtnosti má tvar (3.8). Pre naprogramovanie daného kódu sa predpokladá znalosť práce so štyrmi typmi procedúr (PROC SQL, PROC NLIN, PROC SQRT, PROC GPLOT). PROC GPLOT slúži na programovanie grafických výstupov nelineárnej regresie, konkrétne grafy predikovaných hodnôt, rezíduí, atď.. PROC SORT sa využíva na usporiadanie pozorovaných hodnôt podľa hodnôt jednej alebo viacerých kvantitatívnych alebo kvalitatívnych premenných. Procedúra vytvára len výstupný dátový súbor. Tieto dve procedúry nie sú pre zbehnutie programu nelineárnej regresie nevyhnutné. PROC SQL patrí k základným procedúram, ktoré sú súčasťou každého programu a slúži na vytváranie,

udržiavanie, získavanie a aktualizáciu údajov v tabuľkách. Najdôležitejšou procedúrou pre nelineárnu regresnú analýzu je PROC NLIN. Je to procedúra priamo vytvorená pre tento účel. Obsahuje veľké množstvo príkazov, z ktorých pre nás najrelevantnejšie sú [21]:

- **METHOD** (=MARQUARDT, NEWTON, GRADIENT) – špecifikuje iteračnú metódu, ktorú používa procedúra NLIN pri riešení nelineárneho problému najmenších štvorcov.
- **MAXITER** (=n) – udáva maximálny počet iterácií optimalizačného procesu. Prednastavená hodnota je n=100.
- **CONVERGE** (=c) – určuje konvergenčné kritérium pre procedúru. Prednastavená na hodnotu 10^{-5} .
- **MODEL** – slúži na zadefinovanie tvaru regresného modelu.
- **PARMS** – priestor pre špecifikáciu parametrov a nastavenie ich počiatočných hodnôt. Hodnoty sú prednastavené na 0.001.

Výstup procedúry obsahuje fázy iteračnej metódy (Obrázok č. 6), zhrnutie procesu odhadovania (Obrázok č. 7), analýzu rozptylu (Obrázok č. 8), ktorej súčasťou je aj priemerný súčet štvorcov rezíduí MSE opísaný v kapitole 3.3, odhadovaná hodnota parametrov spolu so štandardnými odchýlkami a s intervalmi spoľahlivosti odhadu parametrov (Obrázok č. 9), korelačná matica parametrov (Obrázok č. 10) a nakoniec ukážka grafu pozorovaných hodnôt v závislosti od veku (Obrázok č. 11).

Obrázok č. 6 - Výstup softvéru SAS Enterprise Guide 1

Nonlinear Regression				
Results				
The NLIN Procedure				
Dependent Variable S2017				
Method: Marquardt				
Iterative Phase				
Iter	a	b	c	Sum of Squares
0	0.000400	4.3E-6	0.1360	0.000085
1	0.000406	0.000041	0.0837	0.000084
2	0.000457	0.000376	-0.1000	0.000079
3	0.000710	0.00200	-0.1086	0.000059
4	0.000934	0.00339	-0.1135	0.000053
5	0.000955	0.00352	-0.1112	0.000053
6	0.000955	0.00352	-0.1119	0.000053
7	0.000955	0.00352	-0.1117	0.000053
8	0.000955	0.00352	-0.1118	0.000053

NOTE: Convergence criterion met.

Zdroj: vlastné spracovanie

Obrázok č. 7 - Výstup softvéru SAS Enterprise Guide 2

Estimation Summary	
Method	Marquardt
Iterations	8
Subiterations	11
Average Subiterations	1.375
R	6.346E-6
PPC(c)	0.000117
RPC(c)	0.000436
Object	4.03E-10
Objective	0.000053
Observations Read	51
Observations Used	51
Observations Missing	0

Zdroj: vlastné spracovanie

Obrázok č. 8 - Výstup softvéru SAS Enterprise Guide 3

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	2	0.000012	6.163E-6	5.54	0.0068
Error	48	0.000053	1.112E-6		
Corrected Total	50	0.000066			

Zdroj: vlastné spracovanie

Obrázok č. 9 - Výstup softvéru SAS Enterprise Guide 4

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
a	0.000955	0.000150	0.000654	0.00126
b	0.00352	0.00106	0.00138	0.00566
c	-0.1118	0.2967	-0.7083	0.4847

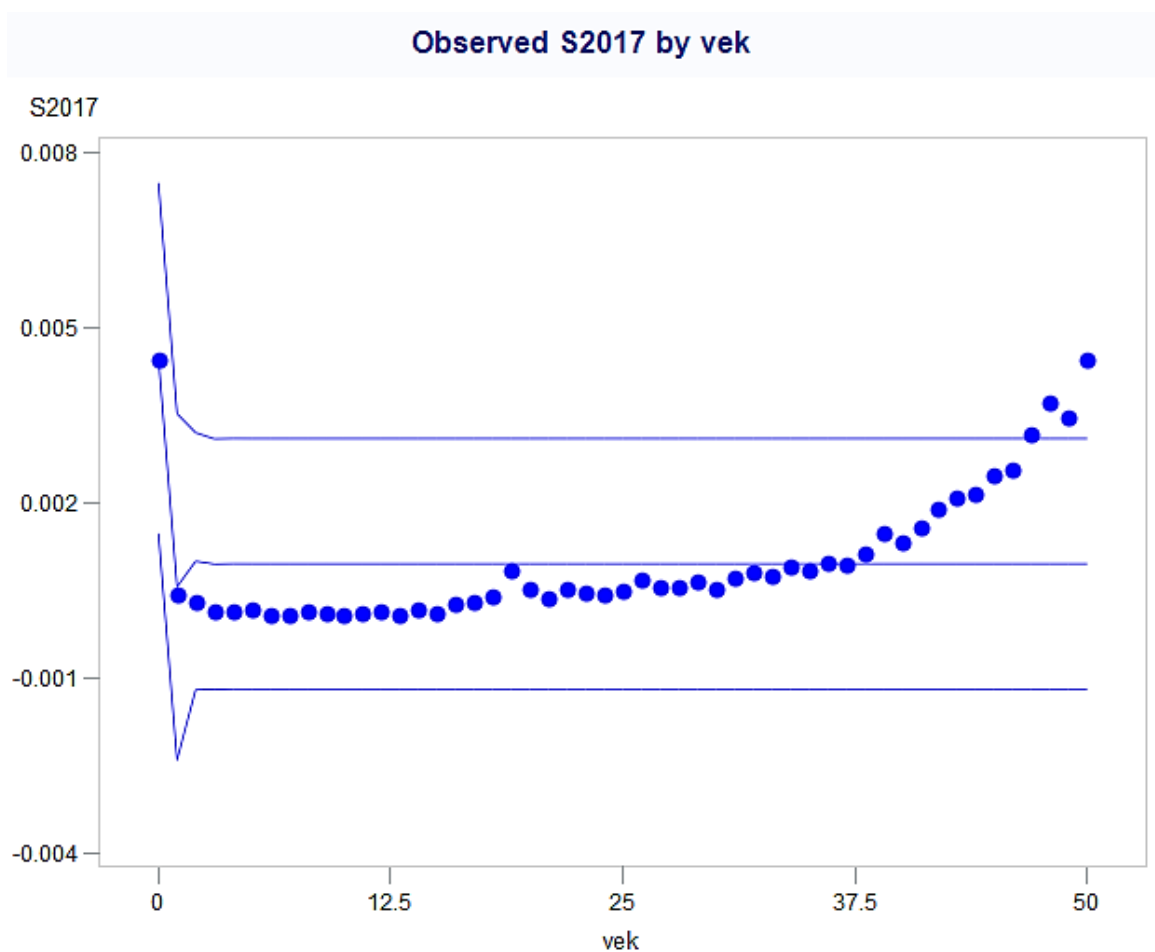
Zdroj: vlastné spracovanie

Obrázok č. 10 - Výstup softvéru SAS Enterprise Guide 5

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: vlastné spracovanie

Obrázok č. 11 - Výstup softvéru SAS Enterprise Guide 6



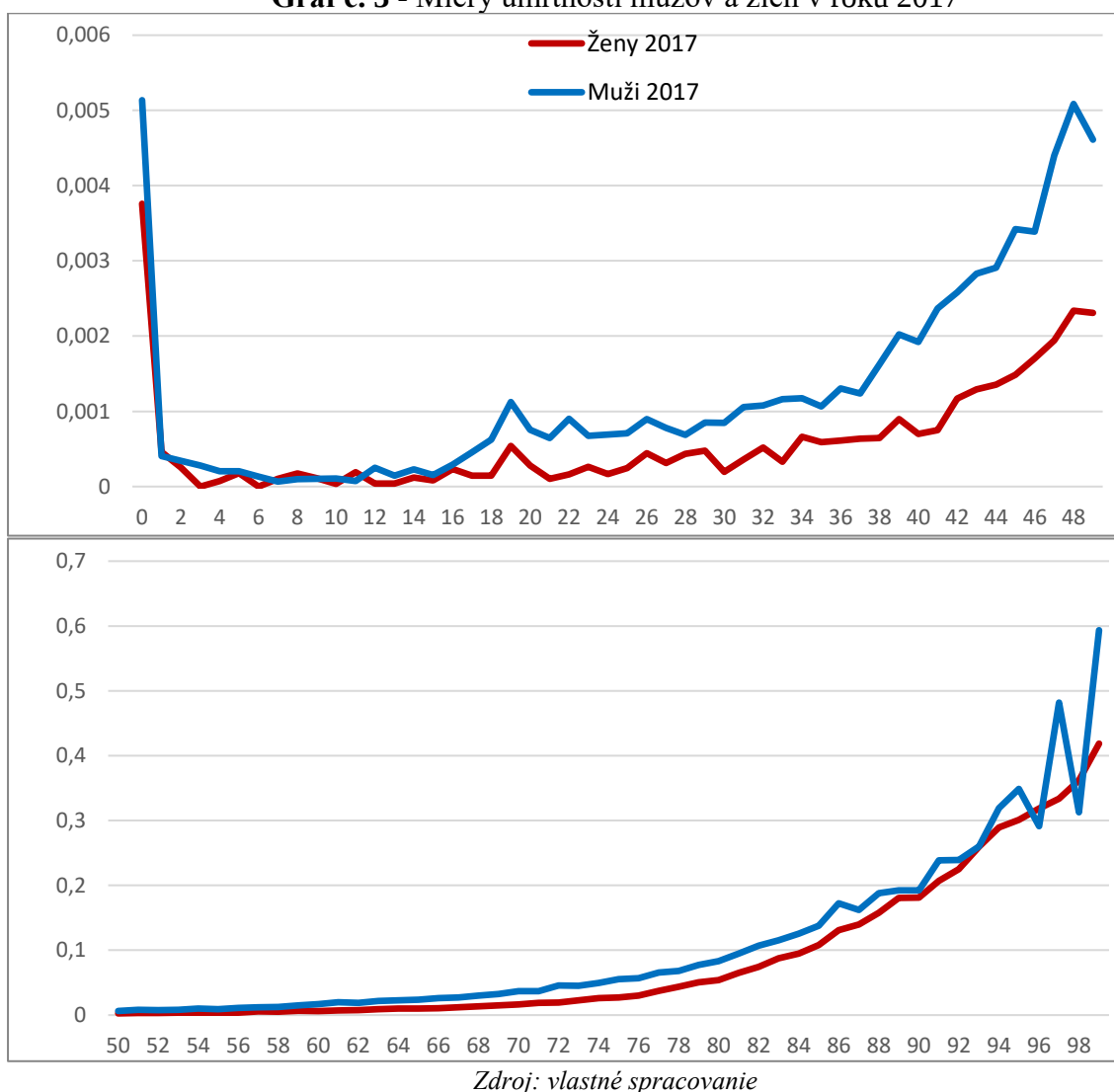
Zdroj: vlastné spracovanie

SAS Enterprise Guide je bezpečné pracovné prostredie analytika, ktoré umožňuje prístup k rovnakým údajom viacerým používateľom. Jednoduché používateľské rozhranie a kompatibilita s inými databázami umožňuje intuitívnu manipuláciu s údajmi a komplexnú analýzu pomocou len jedného nástroja.

4 Výsledky práce a diskusia

Základným problémom modelovania pravdepodobnosti úmrtia alebo miery či intenzity úmrtnosti je, že spojité teoretické pravdepodobnosti úmrtia sa odhadujú štatistickými výpočtami z reálnych dát, ktoré sú triedené do relatívne veľkých vekových intervalov. Inak povedané, okamžité pravdepodobnosti úmrtia sa nahrádzajú pravdepodobnosťami vypočítanými pre ročný vekový interval. Toto spôsobuje nesystematické nepravidelnosti.

Graf č. 3 - Miery úmrtnosti mužov a žien v roku 2017



Na grafe č. 3 sú zobrazené miery úmrtnosti populácie mužov a žien na Slovensku v roku 2017. Podstatou tohto grafu je ukázať dve skutočnosti. Prvou z nich sú

nepravidielnosti najviac viditeľné vo vekoch nad 90 rokov, ktoré sa dajú vysvetliť malým počtom úmrtí a veľkými výkyvmi v ich počtoch. Podobné výkyvy sú aj v mladších vekoch, ale vzhľadom na malé hodnoty výkyvov nie sú také výrazné. Takže prakticky na celom vekovom intervale sú nežiadúce výkyvy, ktoré je potrebné odstrániť, aby sme sa priblížili k skutočným spojeným hodnotám miery úmrtnosti. Takéto metódy nazývame **vyrovnávanie**, resp. **graduácia**. Dôvodom graduácie je tiež to, že klesajúca tendencia miery úmrtnosti je nereálna, hlavne vo vysokých vekoch. Významným nástrojom graduácie sú aj zákony úmrtnosti, ktoré sú predmetom tejto diplomovej práce.

Druhou významnou skutočnosťou je rozdiel v miere úmrtnosti v závislosti od pohlavia. Na celom vekovom intervale je miera úmrtnosti žien nižšia ako miera úmrtnosti mužov, čím sa opäť ako v kapitole 1.3 dostávame k otázke, či je správne (spravodlivé) zjednotenie úmrtnosti mužov a žien. Rodové rozdiely v miere úmrtnosti vedú k odlišnostiam vo výške poistného obzvlášť pre poistenie na dožitie a pri dôchodkovom poistení.

V nasledujúcej časti sa budeme zaoberať podrobnejšie jednotlivými modelmi zákonov úmrtnosti spomínanými v predchádzajúcej kapitole a budeme charakterizovať ich použiteľnosť na konkrétne dáta o úmrtnosti mužov a žien na Slovensku v rokoch 1993, 2005 a v neposlednom rade 2017. Posúdime vhodnosť týchto modelov na základe vyrovnania miery úmrtnosti na celé vekové rozpätie od 0 do 100 rokov. Vekový interval sme z dôvodu lepšej prehľadnosti, a tiež presnejšieho odhadu parametrov modelov rozdelili na dva úseky, oblasť do 50 rokov vrátane a oblasť nad 50 rokov. V závere tejto kapitoly sa zamierame už na modely so zjednoteným pohlavím a vyberieme ten, ktorý najlepšie opisuje úmrtnostné správanie v Slovenskej republike na celom jeho vekovom intervale.

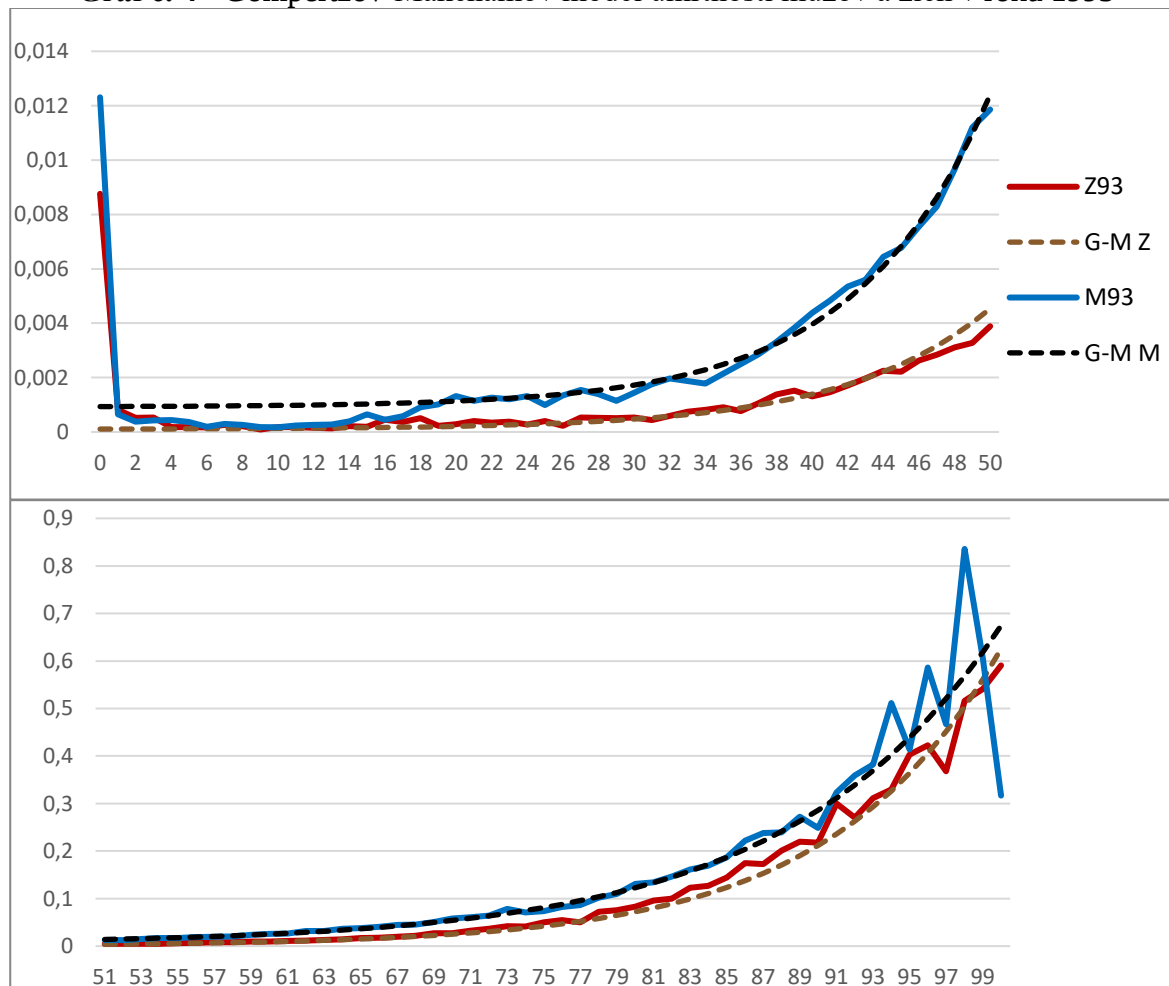
4.1 Modely úmrtnosti pri zohľadnení pohlavia

Pred prijatím Gender smernice (1.3) bolo pohlavie dôležitým faktorom pri modelovaní úmrtnosti. V tejto kapitole aplikujeme vybrané zákony úmrtnosti na populáciu Slovenska v rokoch 1993, 2005 a 2017 s použitím metód uvedených v kapitole 3.2 prostredníctvom softvéru SAS Enterprise Guide.

4.1.1 Gompertzov - Makehamov model úmrtnosti

Prvý zákon, ktorým sme skúsili modelovať úmrtnosť populácie Slovenska je Gompertzov-Makehamov model. Má tvar rastúcej exponenciálnej krivky a je vyjadrený vzťahom (3.8). Jeho parametre sme odhadli pomocou jednej z metód nelineárnej regresie opísanej v kapitole 3.2. Vzhľadom na to, že softvér SAS Enterprise Guide má v ponuke len pár prednastavených typov nelineárnych modelov nezhodujúcich sa s modelmi zákonov úmrtnosti, ktorými sa v našej diplomovej práci venujeme, bolo potrebné vytvoriť vlastný program na odhad parametrov. Ukážku programu je možné nájsť v kapitole 3.4.

Graf č. 4 - Gompertzov-Makehamov model úmrtnosti mužov a žien v roku 1993



Zdroj: vlastné spracovanie

Graf č. 4 znázorňuje miery úmrtnosti mužov M93 a žien Z93 pre celé vekové rozpätie v roku 1993 a následne ich hodnoty vyrovnané pomocou Gompertzovho-Makehamovho zákona (G-M M, G-M Z). Odhady parametrov spolu s mierami presnosti odhadu sú

zobrazené v tabuľke č. 2. Ako vidíme v prvej časti grafu č.4, tento model úmrtnosti nie je vhodný na modelovanie dojčenskej úmrtnosti do 1 roku života, pretože skutočné miery úmrtnosti sa veľmi odlišujú od mier vyrovnaných modelom.

Čo sa týka mužskej časti populácie v roku 1993, model mierne nadhodnocuje ich miery úmrtnosti až do 18 roku života. Od 18 roku života až po 90 rokov môžeme povedať, že Gompertzov-Makehamov model skoro presne opisuje úmrtnostné správanie populácie. Vo vyšších vekoch nad 90 rokov sú na grafe č. 4 mužskej populácie výrazné nepravidelnosti, ktoré pri vyrovnaní mierne podhodnocujú vypočítané miery úmrtnosti.

V prípade opačného pohlavia, je volatilita dát oveľa miernejšia ako u mužov, čím kvalita Gompertzovho-Makehamovho modelu vzrástla. Hodnoty reziduálnej odchýlky tiež naznačujú, že v roku 1993 je tento model vhodnejší pre ženskú populáciu.

Tabuľka č. 2 - Odhad parametrov G-M pre rok 1993

Gompertz-Makeham model 1993				
	Parametre 50-		Parametre 50+	
	Ženy	Muži	Ženy	Muži
c	9,88E-05	0,000919	0,000868	0,00441
A	9,19E-06	0,000015	0,000012	0,000115
B	0,1236	0,1328	0,1086	0,0867
MSE	1,52E-06	2,71E-06	0,000422	0,004606
RMSE	0,001234	0,001648	0,020539	0,067867

Zdroj: vlastné spracovanie

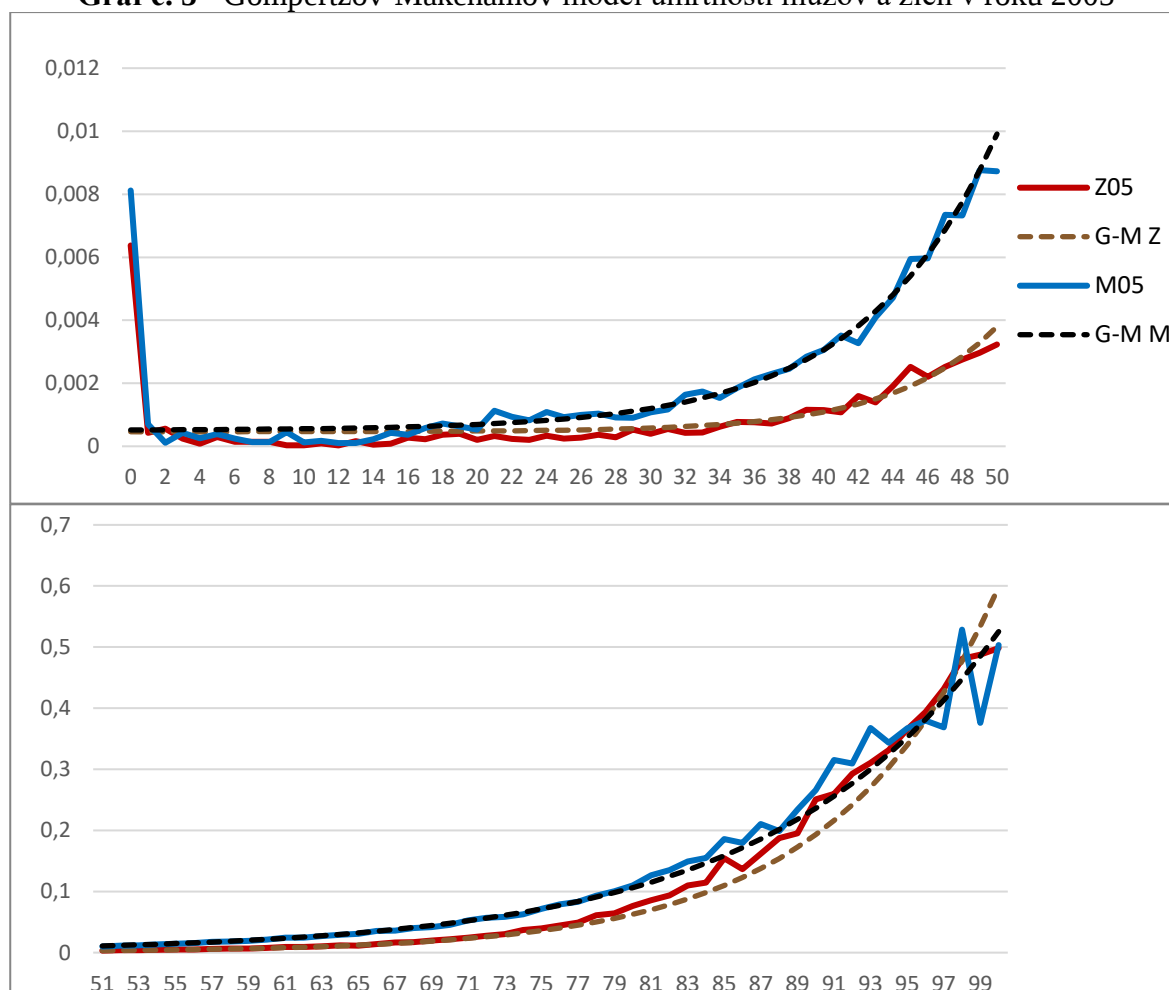
V roku 2005 je kvalita vyrovnaní mier úmrtnosti modelom v prvej polovici vekového intervalu pri oboch pohlaviach relatívne vhodná (Graf č. 5). Na rozdiel od roku 1993 sú miery úmrtnosti do 18 roku života výraznejšie nadhodnotené pri ženách. Čo sa týka žien vo vyšších vekoch, miera úmrtnosti odhadnutého modelu je nižšia ako skutočná miera na vekovom intervale od 80 do 95 roku života. Pri mužskej populácii z dôvodu vysokej volatility je ťažké hodnotiť presnosť vyrovnaní modelu. Hodnoty v mladom veku do 17 rokov sú mierne nadhodnotené, naopak na vekovom intervale od 80 do 96 rokov hodnoty modelu podhodnocujú skutočnú mieru úmrtnosti mužov. Tabuľka č. 3 obsahuje odhadované parametre Gompertzovho-Makehamovho modelu pre rok 2005.

Tabuľka č. 3 - Odhad parametrov G-M pre rok 2005

Gompertz-Makeham model 2005					
	Parametre 50- Ženy Muži		Parametre 50+ Ženy Muži		
c	0,000461	0,000505	0,000941	0,000474	
A	8,13E-07	0,000014	7,1E-06	0,000176	
B	0,1664	0,1302	0,1134	0,080005	
MSE	7,52E-07	1,23E-06	0,000585	0,000674	
RMSE	0,000867	0,001107	0,024181	0,025962	

Zdroj: vlastné spracovanie

Graf č. 5 - Gompertzov-Makehamov model úmrtnosti mužov a žien v roku 2005



Zdroj: vlastné spracovanie

Na grafe č. 6 sa viditeľne prejavuje podobné úmrtnostné správanie ako v roku 1993. Hoci výkyvy miery úmrtnosti u mužov sú pri vysokých vekoch menej volatilné. Odhadnuté parametre sú vyobrazené v tabuľke č. 4. Vo všeobecnosti, pozorovaním Gompertzovho-Makehamovho modelu pre roky 1993, 2005 a 2017, je možné zhrnúť, že model je vhodný najmä pre vekové rozpätie od 18 do 90 rokov. V mladšom veku jemne nadhodnocuje

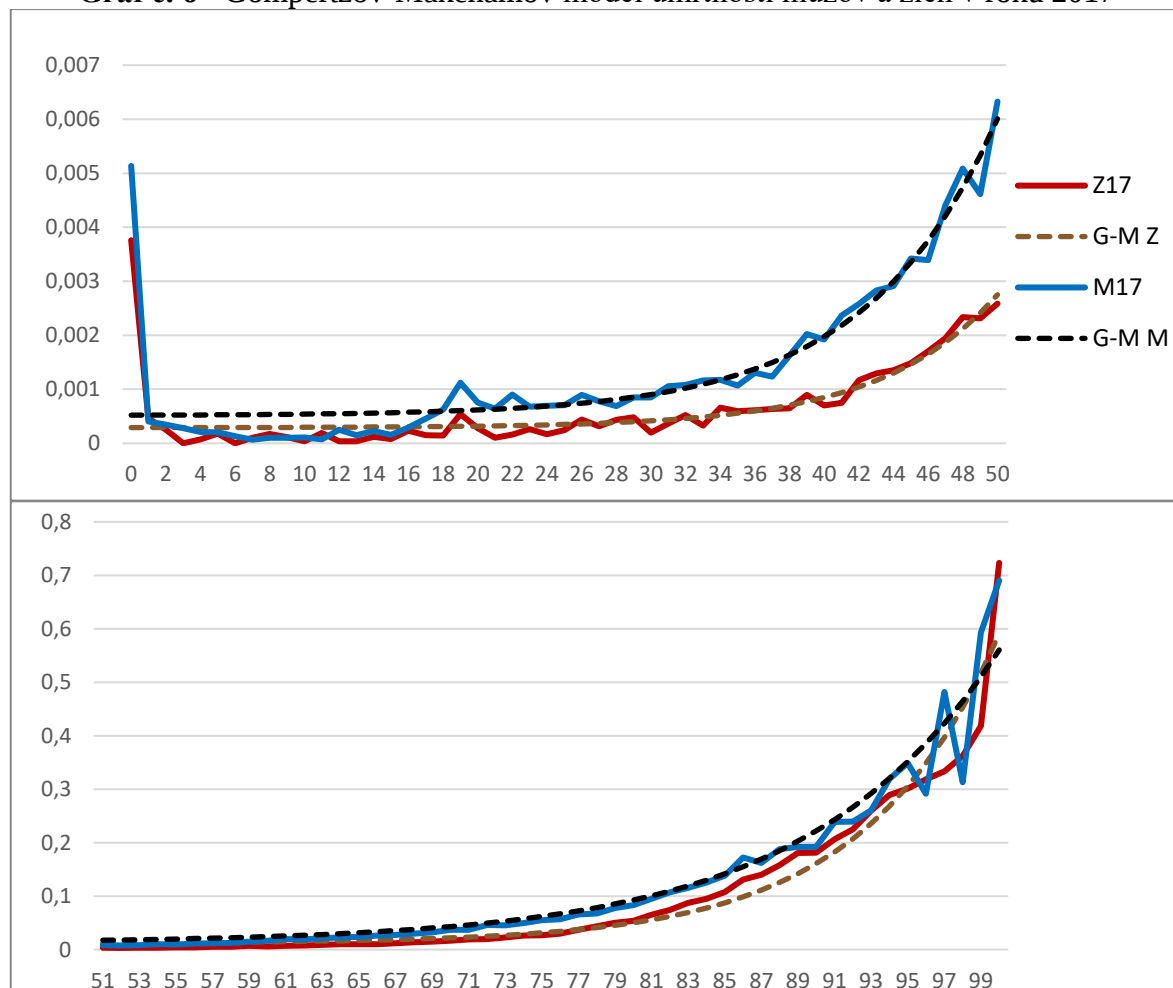
úmrtnosť, a naopak vo vysokých vekoch ju podhodnocuje. V nasledujúcich častiach budeme skúmať či ďalšie modely lepšie odzrkadľujú trend úmrtnosti populácie Slovenska.

Tabuľka č. 4 - Odhad parametrov G-M pre rok 2017

Gompertz-Makeham model 2017				
	Parametre 50-		Parametre 50+	
	Ženy	Muži	Ženy	Muži
c	0,000289	0,000513	0,012	0,01275
A	1,5E-06	7,33E-06	7,32E-07	0,000036
B	0,1481	0,1324	0,1358	0,0963
MSE	2,59E-07	4,9E-07	0,000998	0,00129
RMSE	0,000509	0,0007	0,031589	0,035917

Zdroj: vlastné spracovanie

Graf č. 6 - Gompertzov-Makehamov model úmrtnosti mužov a žien v roku 2017



Zdroj: vlastné spracovanie

4.1.2 Coaleov - Kiskerov model úmrtnosti

V predchádzajúcej časti 4.1.1 sme si mohli všimnúť, že aplikovaný model má v rôznych rokoch, medzi ktorými sú odstupy 12 rokov, veľmi podobné charakteristiky a správa sa podobne v rovnakých vekových intervaloch. Z tohto dôvodu sa v nasledujúcich častiach zameriame len na jeden z rokov 1993, 2005 a 2017.

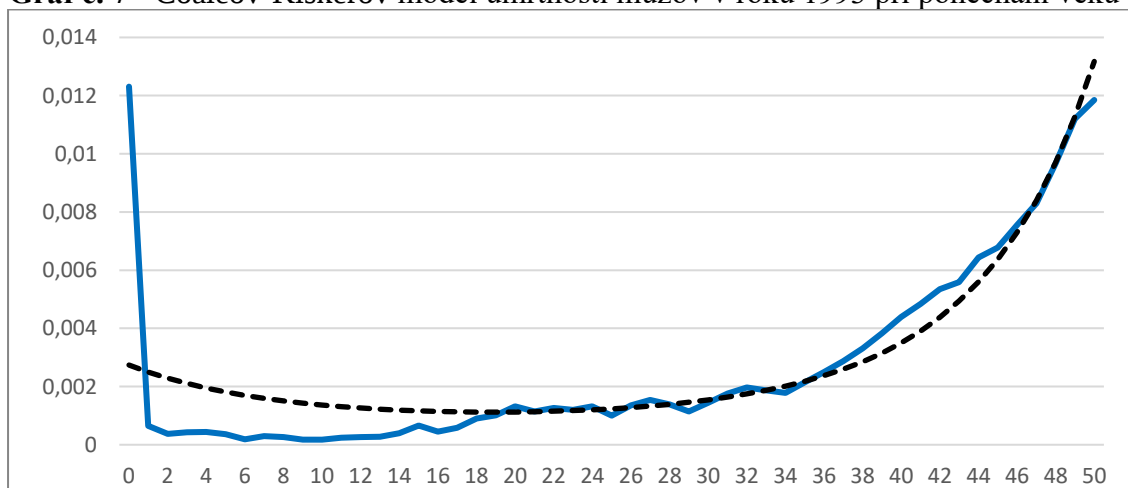
V prípade Coaleovho-Kiskerovho zákona úmrtnosti budeme modelovať rok 1993. Výsledky odhadov parametrov spolu s mierami presnosti vyrovňania sú prezentované v tabuľke č. 5. Rozdiely v odhadovaných parametroch pre vekový interval $\langle 0,50 \rangle$ sa výrazne líšia od intervalu nad 50 rokov.

Tabuľka č. 5 - Odhad parametrov C-K pre rok 1993

Coale-Kisker model 1993				
	Parametre 50-		Parametre 50+	
	Ženy	Muži	Ženy	Muži
A	0,00134	0,000854	-0,00048	-0,00157
B	-0,0142	0,0317	0,181	0,359
C	-8,1811	-8,1361	-13,8863	-20,7498
MSE	1,43E-06	2,89E-06	0,000297	0,004129
RMSE	0,001197	0,0017	0,017238	0,064259

Zdroj: vlastné spracovanie

Graf č. 7 - Coaleov-Kiskerov model úmrtnosti mužov v roku 1993 pri ponechaní veku 0



Zdroj: vlastné spracovanie

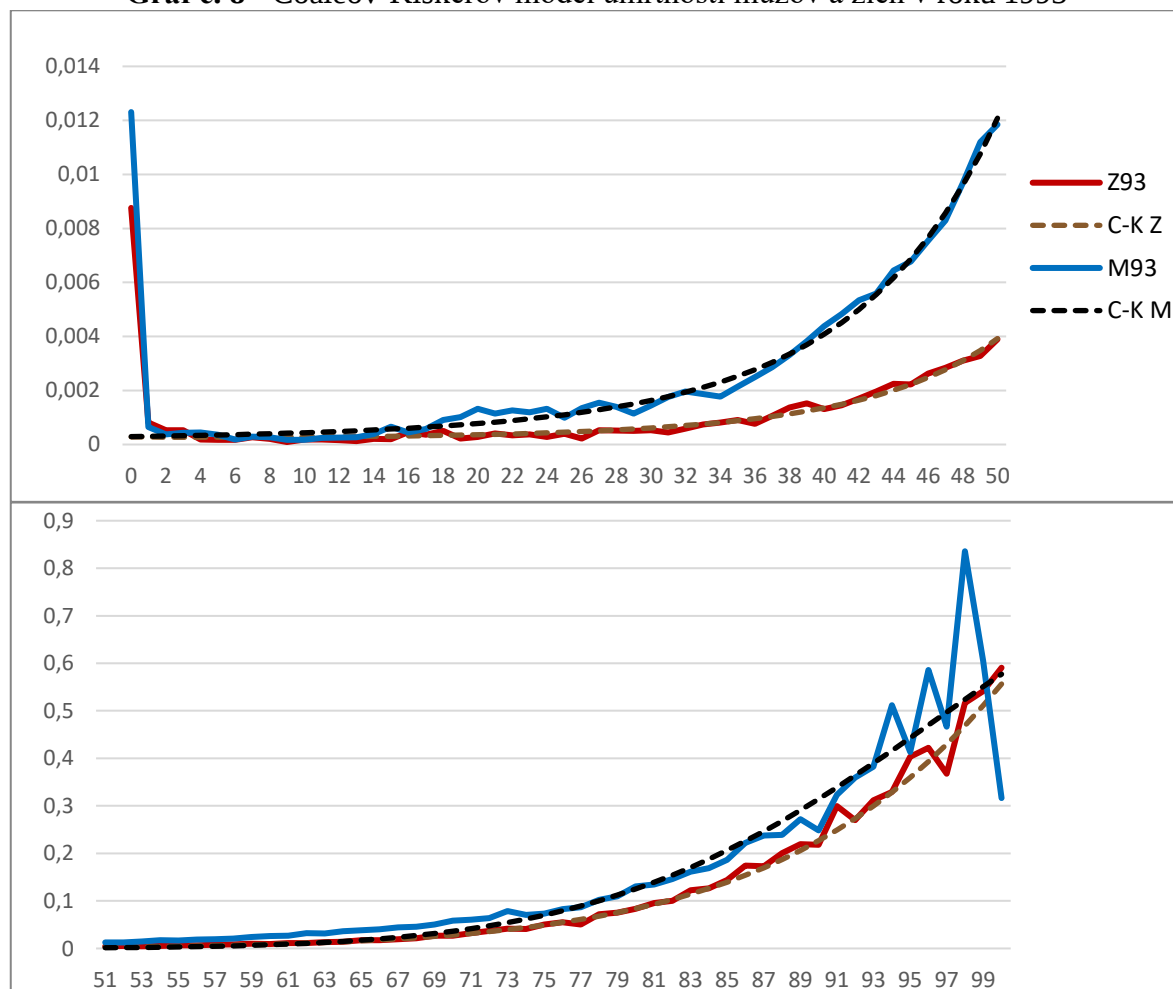
Na grafe č. 8 môžeme pozorovať správanie tohto modelu počas celého života sledovanej populácie. Ešte predtým ako sa detailnejšie pozrieme na celú populáciu, zameriame sa na dojčenskú úmrtnosť. Coaleov-Kiskerov model nie je vhodný

pre modelovanie úmrtnosti do 1 roku života, v závislosti od čoho bolo potrebné pri odhadovaní parametrov odstrániť pozorovanie vo veku 0. Pri ponechaní, model v prvej polovici sledovaného intervalu vykazoval neadekvátny charakter krivky, ktorý okrem iného spôsobil nadhodnotenie úmrtnosti do veku 18 rokov. Jav je zachytený na grafe č. 7.

Keď sa pozrieme na graf č. 8, ktorý vznikol vynechaním veku 0 pri odhade parametrov modelu, môžeme konštatovať, že pre ženskú populáciu sa model zdá byť viac vyhovujúci než pre opačné pohlavie. Výrazné nadhodnocovanie alebo podhodnocovanie nie je na grafe viditeľné a výkyvy vysokých vekov sú pomocou neho primerane vyrovnané.

To isté sa o mužskej populácii nedá povedať, o čom vypovedajú aj výsledné hodnoty mier vyrovnania, ktoré sú na vekovom intervale 50 až 100 rokov až 13 násobne vyššie. Treba však brať do úvahy aj fakt, že výkyvy mužskej populácie vo vysokých vekoch sú omnoho volatilnejšie. Napriek tomu reálne hodnoty mier úmrtnosti mužov sú zreteľne nadhodnotené oproti krivke Coaleovho-Kiskerovho modelu aj na úseku od 51 do 75 rokov.

Graf č. 8 - Coaleov-Kiskerov model úmrtnosti mužov a žien v roku 1993



Zdroj: vlastné spracovanie

4.1.3 Kannistov model úmrtnosti

Kannistov model úmrtnosti má podobné charakteristické črty ako Gompertzov-
Makehamov model. V poslednej kapitole 4.2, v ktorej budeme porovnávať presnosť
jednotlivých modelov medzi sebou navzájom, zistíme, že v prípade neuvažovania pohlavia
ako faktora vplývajúceho na úmrtnosť, sú tieto dva modely takmer totožné.

Predsa sa však bližšie pozrieme na presnosť modelu, a to konkrétne na rok 2005.
V tabuľke č. 6 sú odhadnuté parametre a miery presnosti modelu *MSE* a *RMSE* pre dva
vekové intervaly, menej ako 50 rokov a od veku 50 do 100 rokov.

Tabuľka č. 6 - Odhad parametrov K pre rok 2005

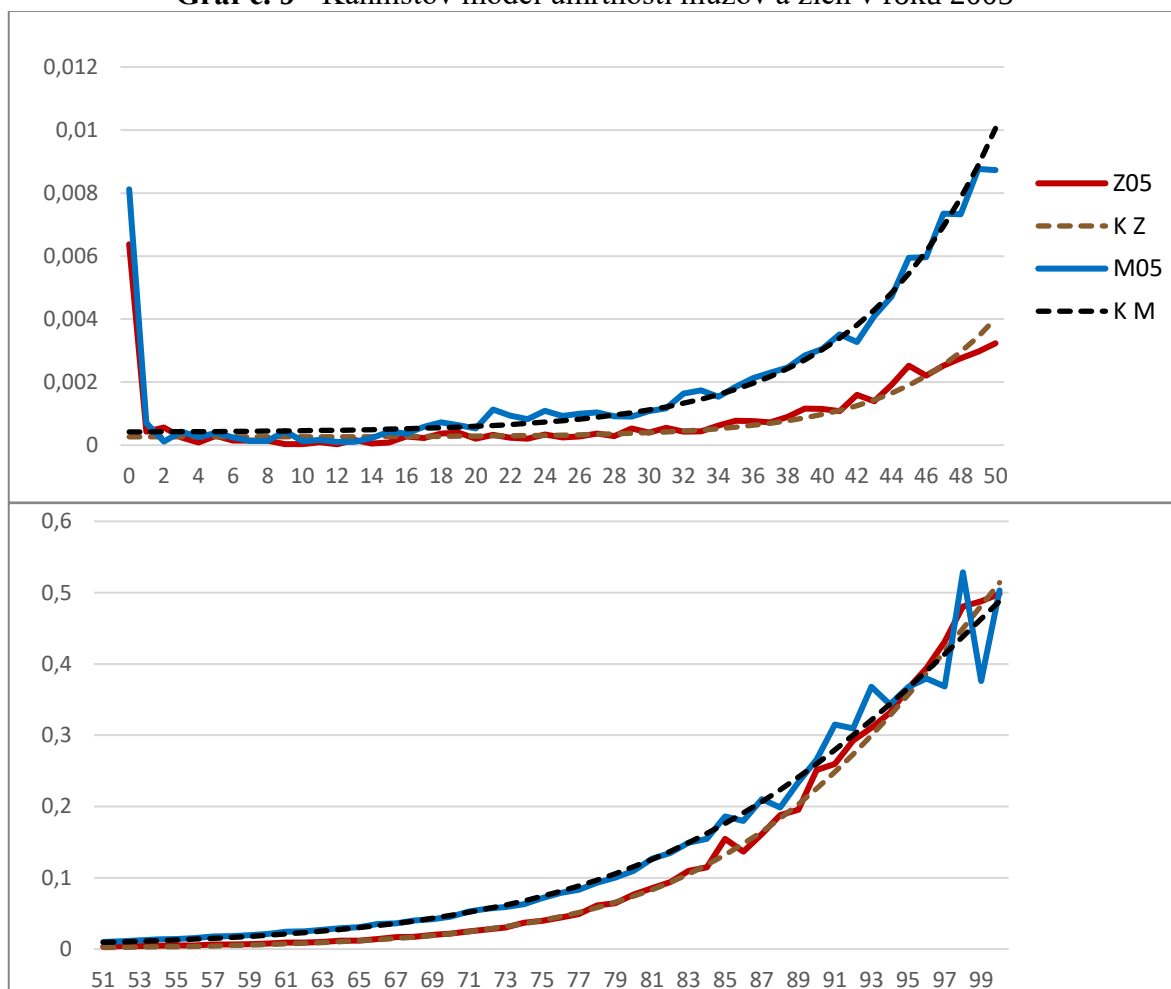
Kannisto model 2005				
	Parametre 50-		Parametre 50+	
	Ženy	Muži	Ženy	Muži
C	0,000261	0,000406	0,00036	0,00252
A	9,05E-07	0,000014	2,56E-06	0,000042
B	0,1667	0,1309	0,1293	0,1002
MSE	7,77E-07	1,25E-06	7,36E-05	0,000457
RMSE	0,000881	0,00112	0,008578	0,021378

Zdroj: vlastné spracovanie

Rovnako ako pri predchádzajúcich modeloch, hlavnou nevýhodou Kannistovho
modelu úmrtnosti je, že pri modelovaní zanedbáva hodnoty dojčenskej úmrtnosti. Z grafu
č. 9 je zrejmé, že hodnoty mier úmrtnosti mužov a žien sa do veku 20 rokov veľmi približujú.
Tento istý jav vidíme aj vo veku nad 90 rokov, kedy dokonca vo veku 95 rokov miera
úmrtnosti žien prevýšila hodnotu mužskej miery úmrtnosti.

Pre ženskú populáciu tento model pozoruhodne presne zachytáva vývoj úmrtnosti
v závislosti od veku s miernym nadhodnotením odhadovaných hodnôt v mladom veku, a tiež
jemným podhodnotením vo vysokých vekoch. Rovnako to platí aj pre opačné pohlavie,
avšak vo vysokých vekoch v dôsledku zvýšenej volatility hodnôt, nevieme jednoznačne
hovoríť o podhodnotení úmrtnosti.

Graf č. 9 - Kannistov model úmrtnosti mužov a žien v roku 2005



Zdroj: vlastné spracovanie

4.1.4 Helligmanov - Pollardov model úmrtnosti

Helligmanov-Pollardov zákon úmrtnosti patrí k zložitejším zákonom úmrtnosti, ktorý pozostáva z väčšieho počtu parametrov. Napriek tomu nie je schopný prekonať hendikep predošlých modelov spočívajúci v neschopnosti modelovať umieranie dojčenskej časti populácie. Vlastne sa o to ani nepokúša, keď vezmeme do úvahy fakt, že definičný obor logaritmickej funkcie nezahŕňa 0, a teda hodnotu v tomto bode na základe vzťahu (3.20) nemôžeme vypočítať.

Tabuľka č. 7 obsahuje prehľad odhadov parametrov modelu spolu s mierami presnosti odhadu. Čo je potrebné spomenúť, že výhodou tohto modelu je pomerne veľká flexibilita. U predchádzajúcich modeloch boli parametre často dosť odlišné už pri vekových intervaloch $\langle 0,50 \rangle$ a $\langle 51,100 \rangle$, zatiaľ čo parametre Helligmanovho-Pollardovho modelu sa líšia len minimálne medzi všetkými pozorovanými rokmi. Ďalšou výhodou oproti ostatným

modelom je, že nemá iba jednoduchý tvar exponenciálnej krivky, ale je viac tvarovateľný, čo sa prejavuje pri modelovaní tzv. havarijného hrbu.

Tabuľka č. 7 - Odhad parametrov H-P pre rok 2017

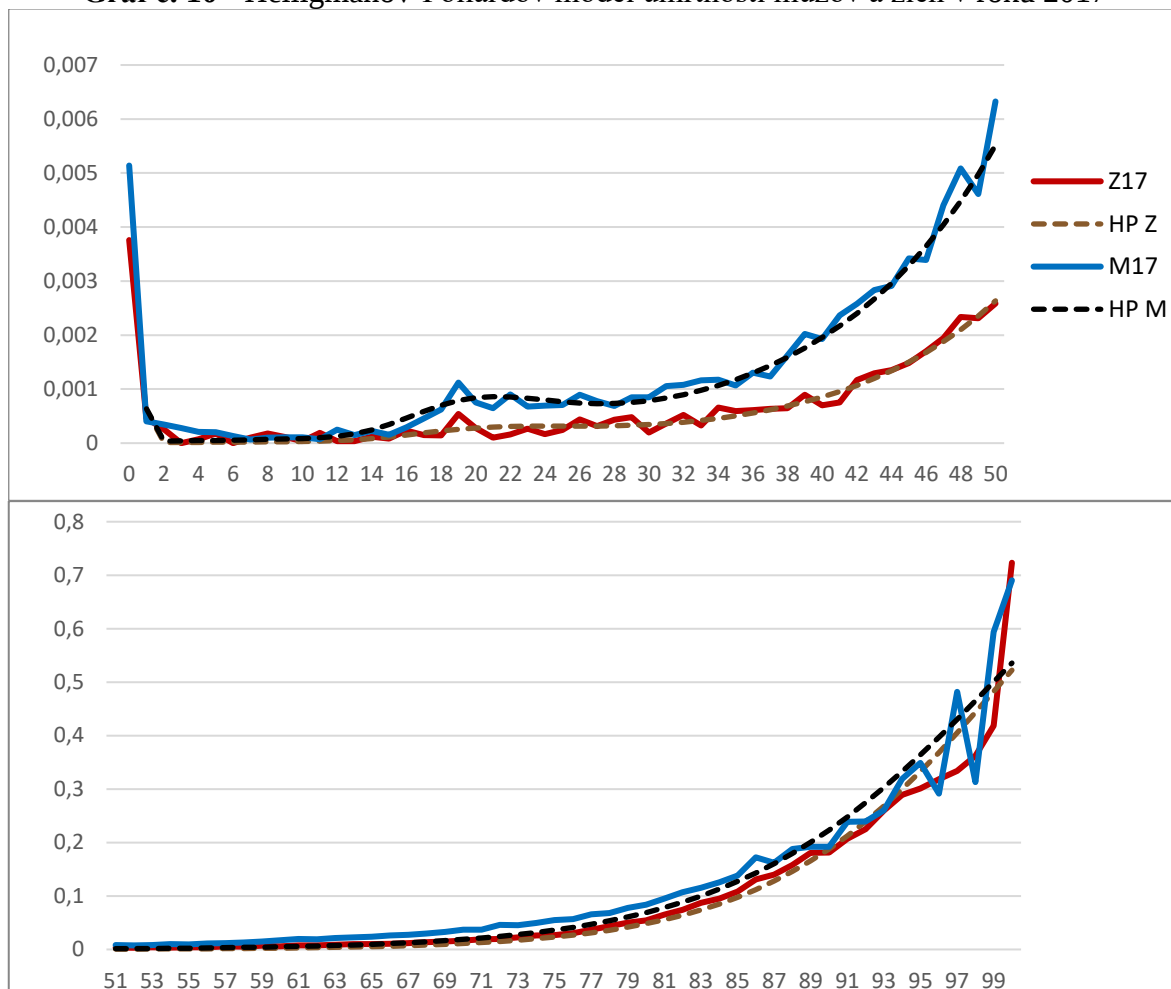
Helligman-Pollard model 2017				
	Parametre 50-		Parametre 50+	
	Ženy	Muži	Ženy	Muži
A	0,0005934	0,0005934	0,0005934	0,0005934
B	0,010536	0,010536	0,010536	0,010536
C	0,9875	0,9875	0,9875	0,9875
D	0,000200492	0,000600492	0,000200492	0,000400492
E	8,9761	11,9761	8,9761	8,9761
F	21,328	20,328	21,328	21,328
G	0,0000091	0,000029991	2,844E-07	0,00000154
K	1,04243	1,4243	0,95	0,95
H	1,1201	1,1101	1,1631	1,1443
MSE	1,35891E-08	4,65981E-08	0,001246799	0,001641765
RMSE	0,000116572	0,000215866	0,035310037	0,040518697

Zdroj: vlastné spracovanie

Na grafe č. 10 môžeme zreteľne vidieť vyššie spomínané výhody a nevýhody modelu. Pri mužskej populácii do veku 5 rokov pozorujeme mierne podhodnotenie reálnych hodnôt. Viditeľný je tiež mierne zakrivený tvar s výstupom na vekovom intervale od 18 do 25 veku života, čo predstavuje spomínaný havarijný hrb, ktorý je v populácii mužov oveľa viac výrazný. Na vekovom intervale (51,100) sa v dôsledku zložitosti modelu nepodarilo nájsť optimálne hodnoty parametrov. Pomocou ani jednej z metód nelineárnej regresie dostupných v softvéri SAS Enterprise Guide nedošlo k splneniu konvergenčného kritéria pri danom počte iterácií. Vďaka tomu vidíme, že krivka Helligmanovho-Pollardovho modelu spoľahlivo neopisuje krivku mier úmrtnosti mužov.

Naopak, pri ženskej populácii odhadovaná krivka Helligmanovho-Pollardovho modelu výstižne opisuje úmrtnostné správanie na celom vekovom intervale. Vo vysokých vekoch sú badateľné menšie nezrovnalosti, ktoré sú však pravdepodobne spôsobené neprimeranosťou reálnych údajov a ich volatilitou vo vysokých vekoch.

Graf č. 10 - Helligmanov-Pollardov model úmrtnosti mužov a žien v roku 2017



Zdroj: vlastné spracovanie

4.1.5 CoDe model úmrtnosti

CoDe model úmrtnosti je najmladší a zároveň najzložitejší zo zákonov úmrtnosti, ktorými sa v diplomovej práci zaoberáme. V tejto časti sa pokúsime modelovať úmrtnostné správanie populácie v roku 2017. Je to z dôvodu, že predpokladáme, že model publikovaný v roku 2016 by mal najlepšie opísať úmrtnosť v tak blízkom roku, hoci bol testovaný na populácii v období od 1950 do 2010.

Zložitosť modelu spočíva v tom, že je zložený z 12 neznámych ukazovateľov. Až 5 z týchto parametrov sa nevyvíja v čase a sú totožné v rôznych krajinách, preto sme ich na základe grafického výstupu a názorov expertov z výskumu [18] vopred stanovili, čím sa model zjednodušil. Okrem iného model obsahuje dve konštanty, ktorých hodnota závisí od odhadnutých parametrov, ktoré prácu v softvéri značne sťažujú. Zásadných je pre nás

zvyšných 5 parametrov, ktoré sme sa snažili odhadnúť. Odhadnuté parametre spolu s fixnými parametrami a konštantami sa nachádzajú v tabuľke č. 8.

Tabuľka č. 8 - Odhad parametrov CoDe pre rok 2017

CoDe model 2017					
	Odhadované parameter			Fixné parametre	
	Ženy	Muži		Ženy	Muži
A	0,000275	0,000138	m	16	16
B	0,0735	0,0272	h	30	30
a	0,000209	0,000398	M	80	80
b₁	0,1321	0,1	g	0,7	0,7
b₂	0,0675	0,1143	c₁	-0,0063	0,001258
b₃	0,121	0,1159	c₂	-0,0479	7,75E-05
MSE	0,001002057	0,001257653			
RMSE	0,031655278	0,035463407			

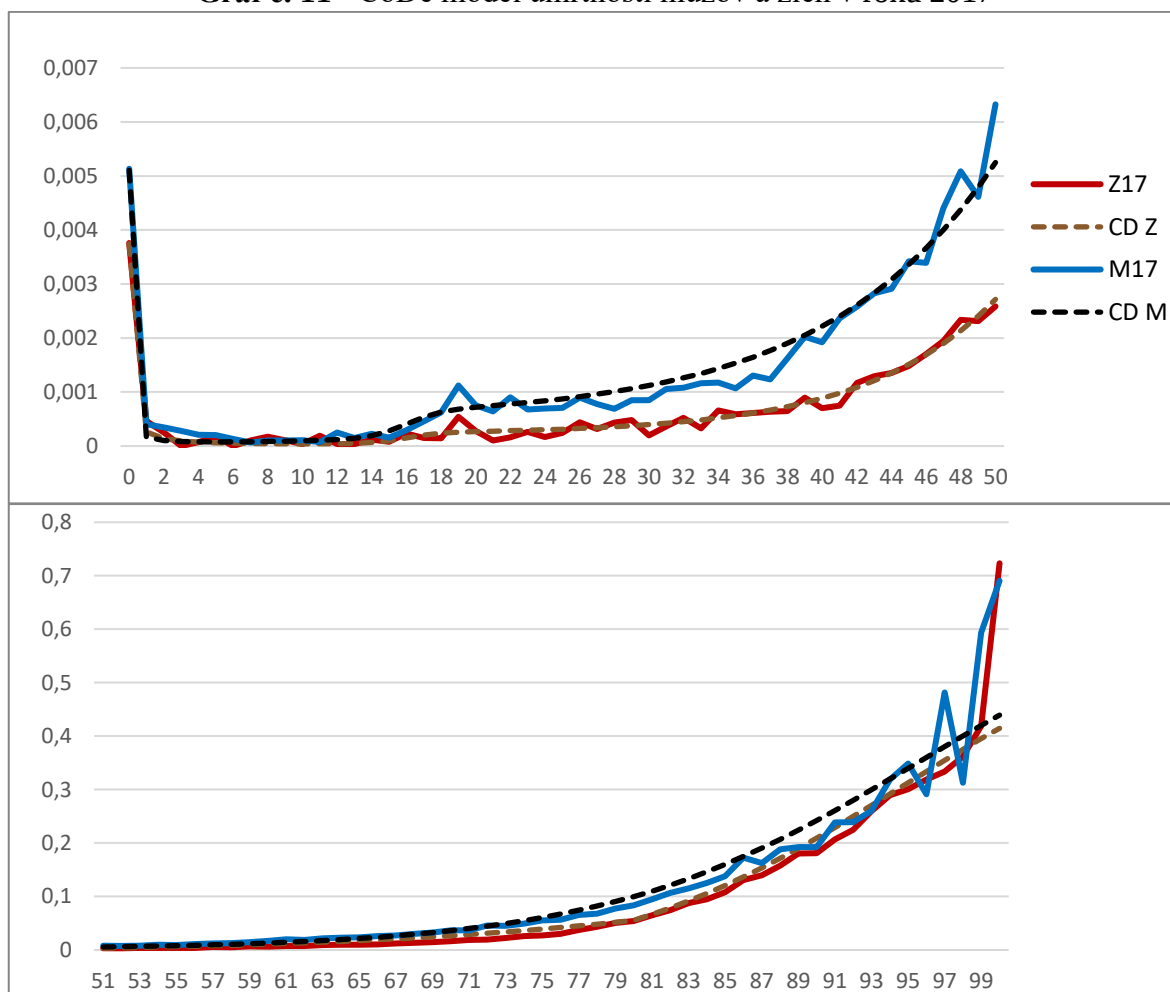
Zdroj: vlastné spracovanie

V porovnaní s ostatnými modelmi je v tabuľke č. 8 dominantná jedna výhoda a to tá, že odhadované parametre modelu vyhovujú celému vekovému intervalu. Spôsobené je to tým, že samotný model obsahuje podmienku, ktorá rozdeľuje vekový interval na 3 časti približené v časti 3.1.7, a teda napríklad na vek 52 rokov nemajú vplyv parametre, ktoré sa týkajú úmrtnosti do veku 50 rokov.

Ďalšou významnou výhodou je jeho schopnosť zakomponovania dojčenskej úmrtnosti do modelu, na rozdiel od predchádzajúcich modelov, ktoré vekový interval $\langle 0,1 \rangle$ úplne ignorovali, alebo pri ktorých spôsobil výraznú zmenu krivky odhadovaného modelu.

U žien CoDe model pekne vyrovnáva výkyvy na celom vekovom intervale s miernym nadhodnocovaním v druhej polovici intervalu, čo je čiastočne spôsobené aj veľmi vysokými mierami úmrtnosti v posledných rokoch (Graf č. 11). U mužskej populácii dochádza k väčším rozdielom. Od havarijného hrbu nebol model schopný zachytiť tvar krivky reálnych dát mier úmrtnosti, čo môže byť zapríčinené problémom splnenia konvergenčného kritéria v SAS Enterprise Guide z dôvodu náročnosti modelu a vysokej volatility mužských mier úmrtnosti. To spôsobilo nadhodnotenie hodnôt medzi vekmi 20 až 40 rokov a takisto vo vysokých vekoch od 75 rokov. Hodnoty sa zdajú byť tiež mierne podhodnotené od 45 do 50 roku života.

Graf č. 11 - CoDe model úmrtnosti mužov a žien v roku 2017



Zdroj: vlastné spracovanie

4.2 Porovnanie modelov úmrtnosti populácie Slovenska

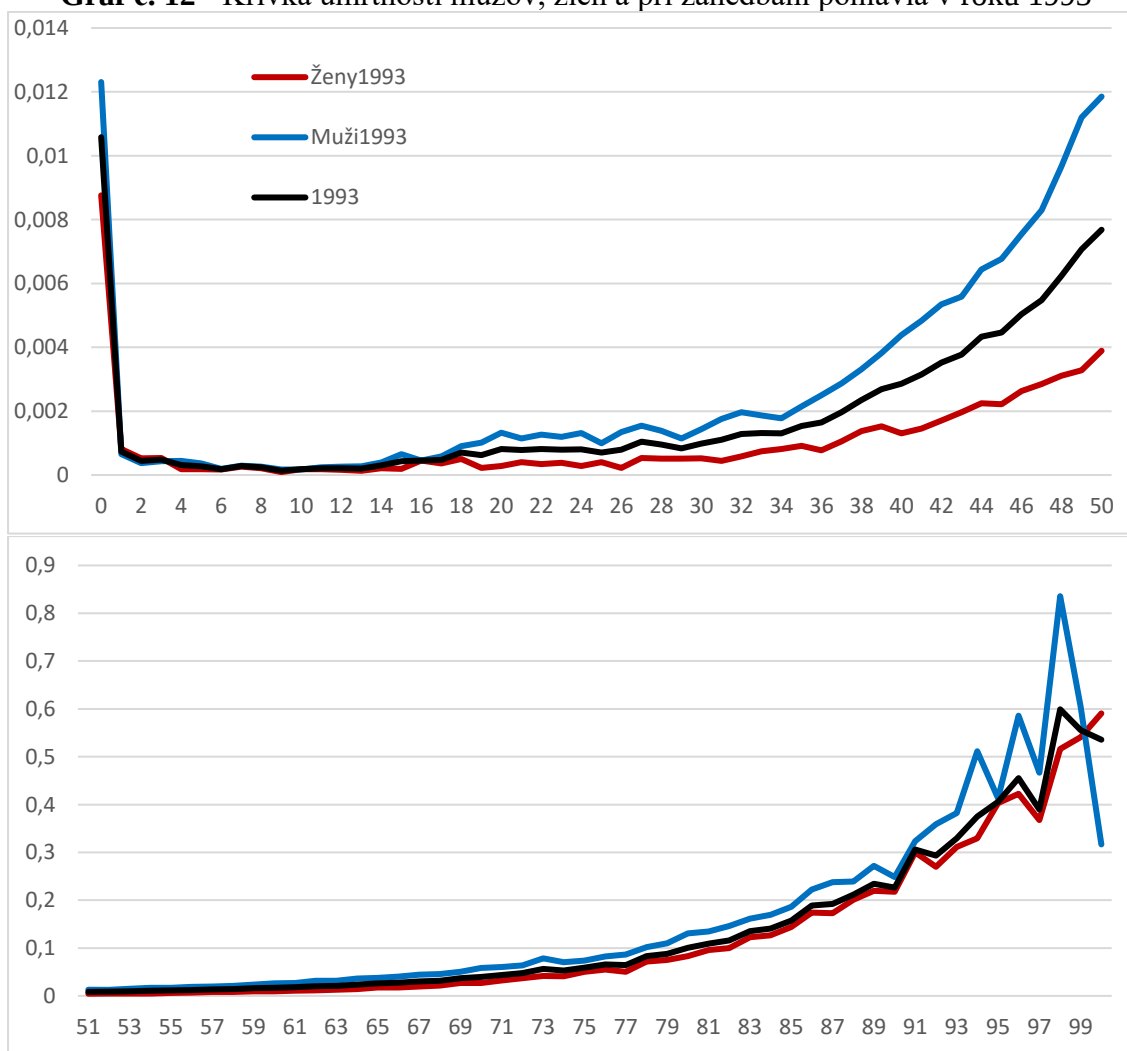
Výber modelu, ktorý najlepšie opisuje úmrtnostné správanie populácie, tvorí jednu z najdôležitejších súčastí modelovania, keďže pre poisťovňu je dôležité vedieť odhadnúť úmrtnosť čo najpresnejšie. Existuje niekoľko ukazovateľov, ktorými je možné sledovať kvalitu vyrovnania modelu. Jedným z nich je aj ukazovateľ priemerný súčet štvorcov rezíduí *MSE* opísaný v kapitole 3.3.

V tejto časti práce budeme porovnávať 5 modelov úmrtnosti opísaných v kapitole 3.1 pre vybrané roky 1993, 2005 a 2017 pri zanedbaní pohlavia ako faktora vplyvajúceho na úmrtnosť populácie Slovenskej republiky.

4.2.1 Analýza v roku 1993

Zanedbanie faktora pohlavia pri modelovaní úmrtnosti spôsobí nadhodnotenie úmrtnosti ženskej populácie a podhodnotenie mužskej populácie, výsledkom čoho je krivka, ktorá je akýmsi priemerom hodnôt úmrtí mužov a žien v danom veku. Graf č. 12 zobrazuje rozdielnosť kriviek úmrtnosti mužov, žien a následne krivku pri zanedbaní pohlavia, ktorá je predmetom modelovania v tejto kapitole.

Graf č. 12 - Krivka úmrtnosti mužov, žien a pri zanedbaní pohlavia v roku 1993



Zdroj: vlastné spracovanie

Rozdiely medzi úmrtnosťou mužov a žien do 18 roku života sú porovnateľné, preto po tento vek je zanedbanie pohlavia prípustné. Môžeme si všimnúť na vekovom intervale $\langle 51,100 \rangle$ rokov, že čierna krivka s označením 1993, ktorá nerozlišuje pohlavie pri úmrtnosti, sa približuje viac k úmrtnosti ženskej populácie, čo ale spôsobuje značné podhodnotenie

úmrtnosti mužov na tomto intervale. Výhodou je, že výkyvy viditeľné vo vysokých vekoch pri mužoch sú pozorovateľnejšie menej volatilné.

Tabuľka č. 9 obsahuje výsledné hodnoty ukazovateľa priemerného súčtu štvorcov rezíduí na 25 ročných vekových intervaloch a následne na celom modelovanom vekovom rozpätí. Výsledky sú taktiež zobrazené v grafoch č. 13 až 16.

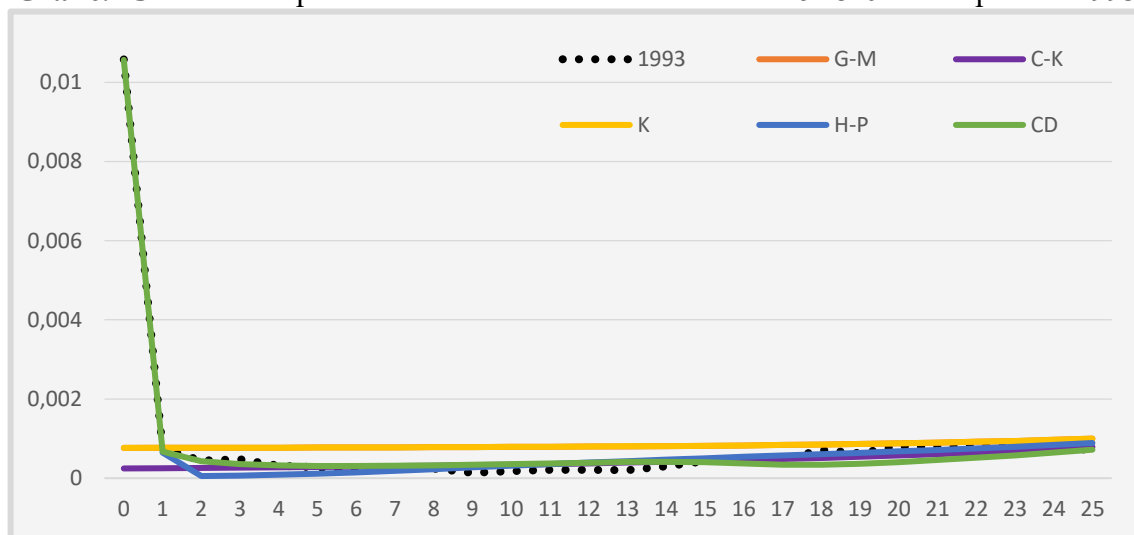
Tabuľka č. 9 - Hodnoty MSE pre rôzne vekové intervaly pre rok 1993

MSE	G-M	C-K	K	H-P	CD
0-25	3,871E-06	4,13334E-06	3,87068E-06	4,33171E-06	3,42474E-08
26-50	4,56125E-08	2,26167E-08	4,52148E-08	5,58282E-08	8,90391E-08
51-75	2,36779E-06	1,90551E-05	1,70619E-05	7,41522E-05	5,7423E-06
76-100	0,001564269	0,000840738	0,000807622	0,000846197	0,00138666
0-100	0,000388789	0,00021389	0,000205137	0,000228938	0,000344685

Zdroj: vlastné spracovanie

Najlepšie výsledky pre vekový interval do 25 rokov dosiahol, ako sme očakávali, CoDe model. Dôvodom je, že ako jediný je schopný modelovať dojčenskú úmrtnosť, čím hodnota ukazovateľa MSE bola niekoľko násobne nižšia ako pri zvyšných modeloch, konkrétne $3,42474 \cdot 10^{-8}$. Na grafoch je CoDe model znázornený zelenou farbou.

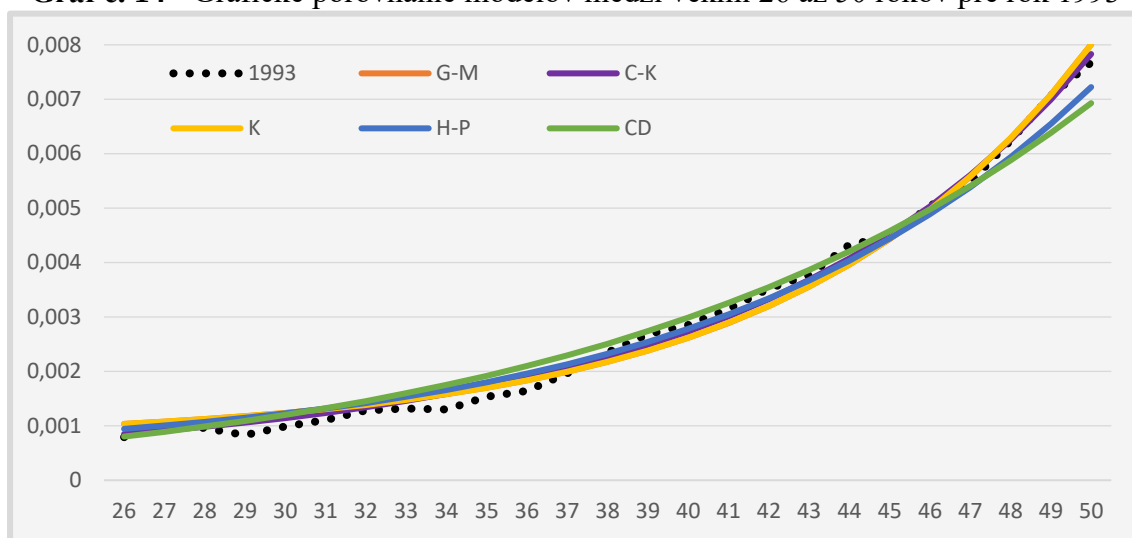
Graf č. 13 - Grafické porovnanie modelov od narodenia do 25 roku života pre rok 1993



Zdroj: vlastné spracovanie

Na nasledujúcom vekovom intervale $\langle 26,50 \rangle$ sa najlepším modelom stal Coaleov-Kiskerov model s hodnotou $MSE = 2,26167 \cdot 10^{-8}$. Na tomto intervale dosahovali všetky z vybraných modelov veľmi dobré výsledky, čo vidíme aj na grafe č. 14.

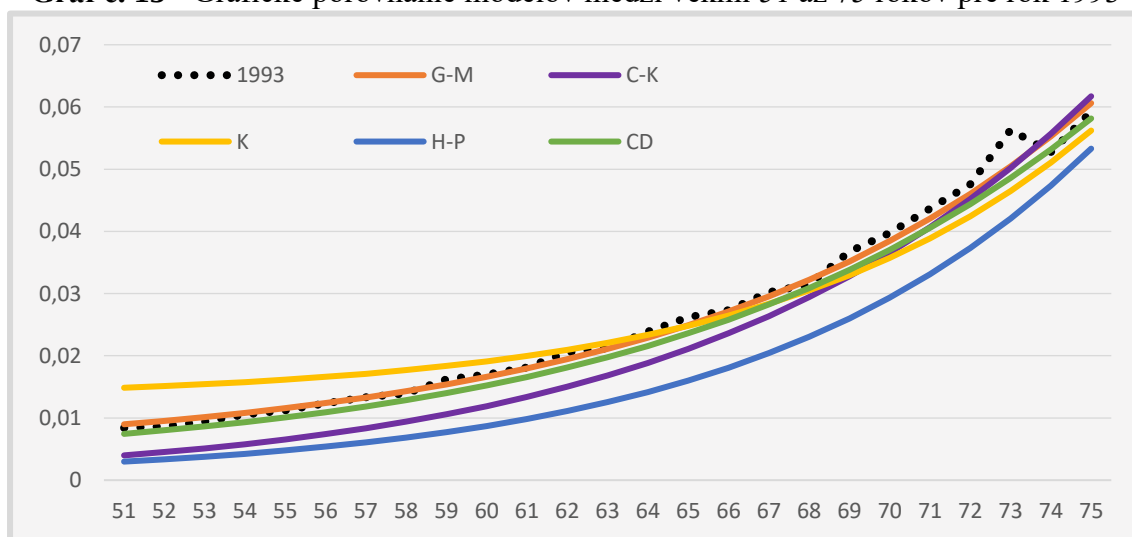
Graf č. 14 - Grafické porovnanie modelov medzi vekmi 26 až 50 rokov pre rok 1993



Zdroj: vlastné spracovanie

Na ďalšom vekovom intervale medzi vekmi 51 až 75 rokov môžeme jednoznačne za najlepší model považovať najznámejší a zároveň najjednoduchší Gompertzov- Makehamov model s $MSE = 2,36779 \cdot 10^{-6}$. Tento model sa často používa pri modelovaní úmrtnosti pre veky od 30 do 70 rokov, čo je zrejmé aj z grafu č. 15.

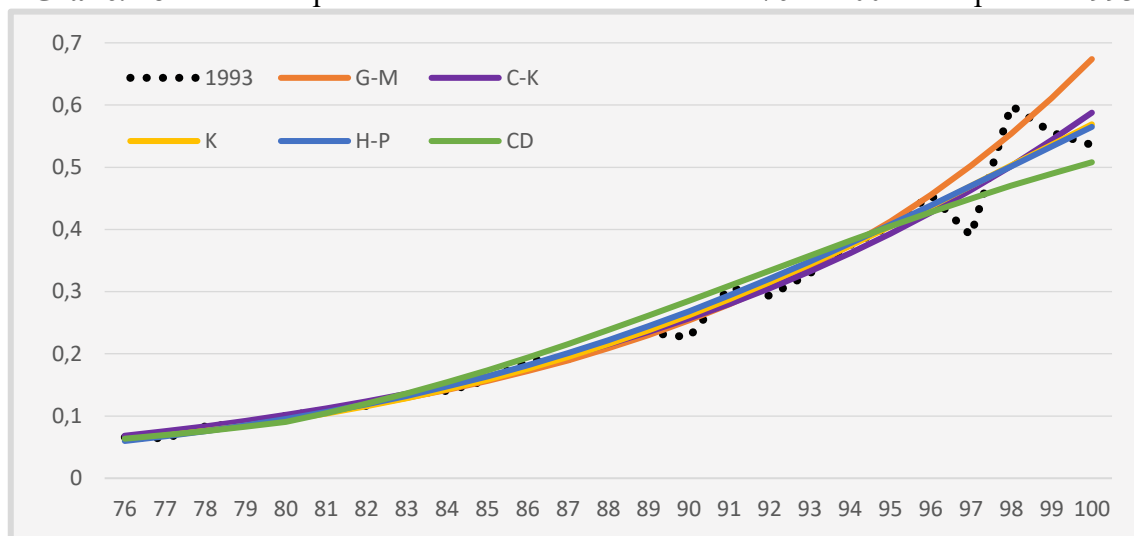
Graf č. 15 - Grafické porovnanie modelov medzi vekmi 51 až 75 rokov pre rok 1993



Zdroj: vlastné spracovanie

Na poslednom sledovanom vekovom intervale od 76 do 100 rokov sa za najlepší model považuje Kannistov model. To bolo takisto predvídateľné, lebo ako sme už spomenuli v kapitole 3.1.5, tento model vykazuje spoľahlivé výsledky hlavne vo vysokých vekoch, a to až do 120 roku života. Jeho hodnota MSE na tomto intervale je 0,000807622.

Graf č. 16 - Grafické porovnanie modelov medzi vekmi 76 až 100 rokov pre rok 1993



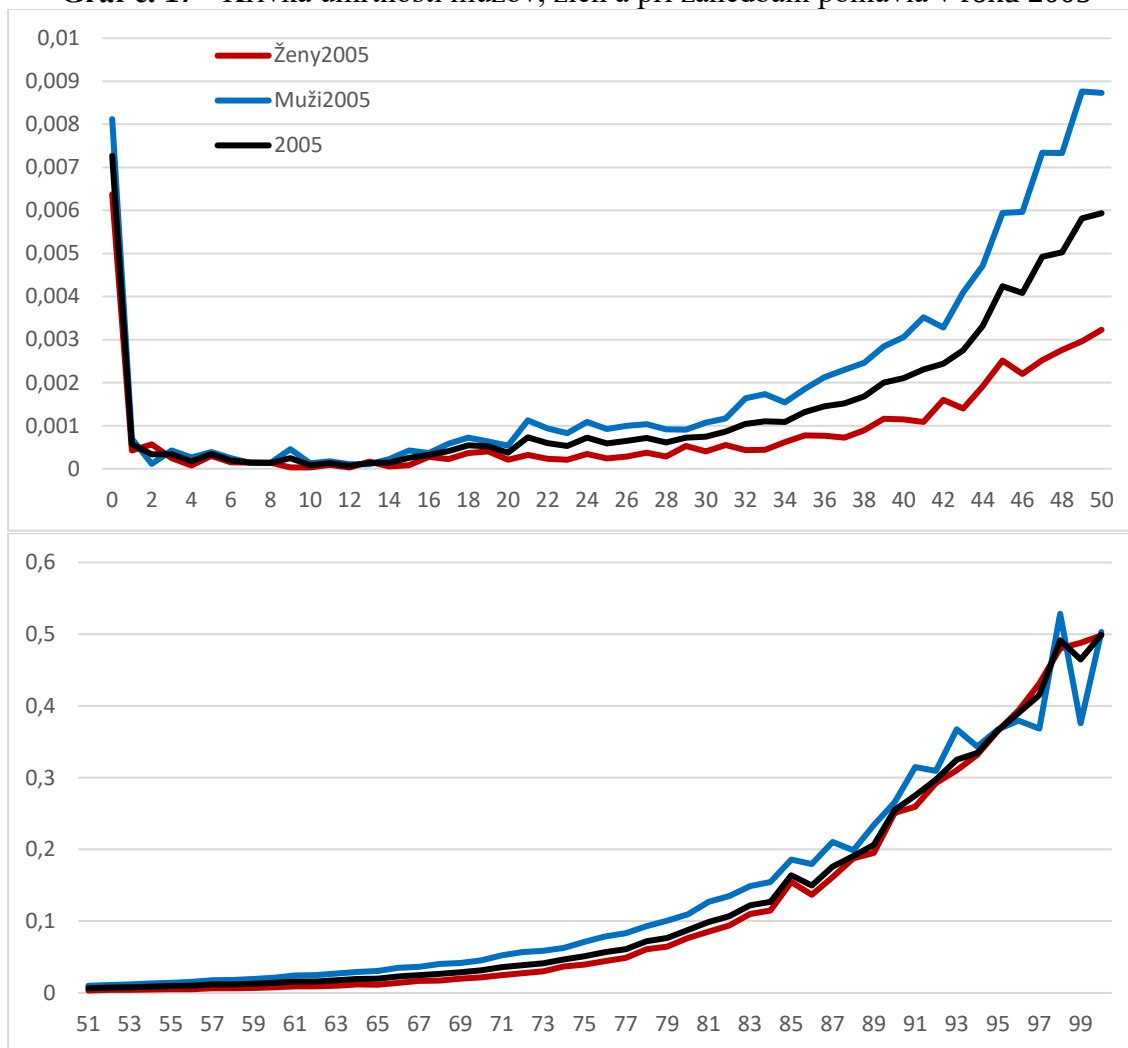
Zdroj: vlastné spracovanie

Model, ktorý najpresnejšie opisuje úmrtnostné správanie populácie Slovenskej republiky v roku 1993 je Kannistov model. Najviac k tomu prispel fakt, že vo vysokých vekoch mal najlepšie výsledky aproximácie, čo bolo najproblémovjšou oblasťou pre ostatné modely. Druhým najlepším modelom pre tento rok je Coaleov-Kiskerov model a hneď za ním je Helligmanov-Pollardov model. Rozdiely v *MSE* sú pre tento rok veľmi malé, preto všetky tri môžeme považovať za vhodné.

4.2.2 Analýza v roku 2005

Vplyv faktora pohlavia sme spomenuli už v predchádzajúcej časti. Z údajov a aj z grafu č. 17 môžeme pozorovať, že úmrtnosť mužov a žien sa oproti roku 1993 navzájom k sebe mierne priblížila. Na grafe č. 17 je zobrazená čierna krivka, ktorá predstavuje úroveň miery úmrtnosti pri zanedbaní pohlavia. V mladom veku do 14 rokov predstavuje spoľahlivý opis úmrtnosti oboch pohlaví a so stúpajúcim vekom sa krivka približuje viac ku krivke úmrtnosti ženskej populácie. Táto krivka rovnako ako v roku 1993 nie je silno ovplyvnená kolísaním mužskej úmrtnosti vo vysokých vekoch.

Graf č. 17 - Krivka úmrtnosti mužov, žien a pri zanedbaní pohlavia v roku 2005



Zdroj: vlastné spracovanie

Tabuľka č. 10 zachytáva miery presnosti modelu prostredníctvom hodnoty *MSE* pre 25 ročné vekové intervaly, a tiež pre celý sledovaný vekový interval od 0 po 100 rokov. Grafické porovnanie modelov je zobrazené v grafoch č. 18 až 21.

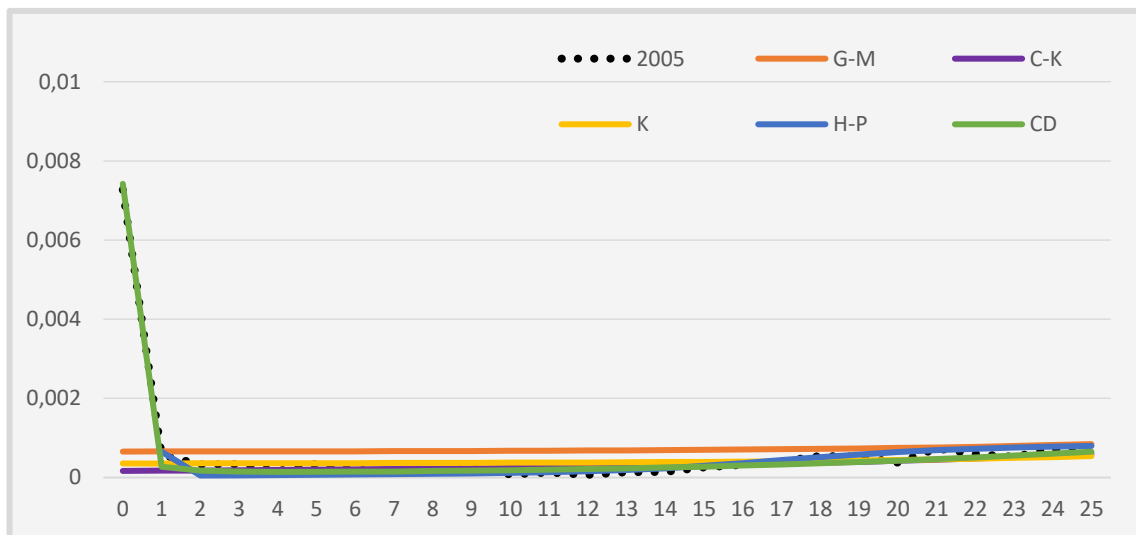
Tabuľka č. 10 - Hodnoty MSE pre rôzne vekové intervaly pre rok 2005

MSE	G-M	C-K	K	H-P	CD
0-25	1,8365E-06	1,96183E-06	1,86956E-06	2,05316E-06	1,68477E-08
26-50	9,09181E-08	3,7277E-08	6,7211E-08	4,81809E-08	8,55551E-08
51-75	1,1482E-06	3,28768E-05	1,09436E-06	1,53335E-05	8,82423E-06
76-100	0,000783705	0,000175619	0,000143764	0,000141899	0,000320478
0-100	0,000194766	5,21222E-05	3,6354E-05	3,94594E-05	8,1536E-05

Zdroj: vlastné spracovanie

Na vekovom intervale do 25 rokov dáva najlepšie výsledky pri modelovaní úmrtnosti rovnako ako v roku 1993 CoDe model, keďže ako jediný dokáže modelovať úmrtnosť do jedného roku života. Priemerný súčet štvorcov rezíduí $MSE = 1,68477 \cdot 10^{-8}$.

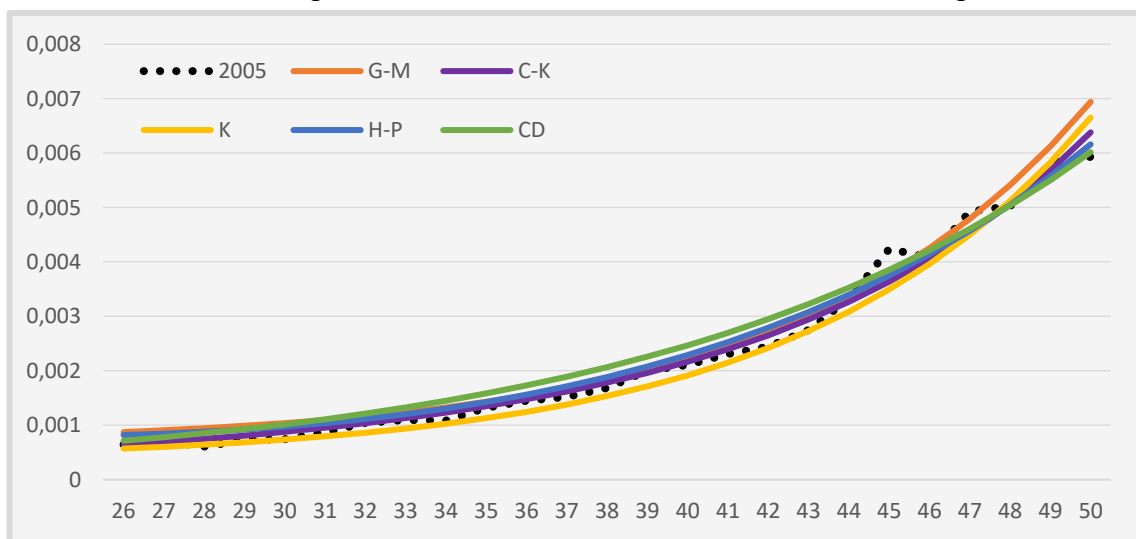
Graf č. 18 - Grafické porovnanie modelov od narodenia do 25 veku života pre rok 2005



Zdroj: vlastné spracovanie

Na ďalšom vekovom intervale je takisto zhoda s rokom 1993, kde najkvalitnejšie vyrovnanie patrí Coaleovmu-Kiskerovmu modelu s hodnotou $MSE = 3,7277 \cdot 10^{-8}$. Ostatné modely podávajú porovnateľne dobré výsledky, čo je viditeľné aj na grafe č. 19. Rozdiely medzi nimi sú zanedbateľné.

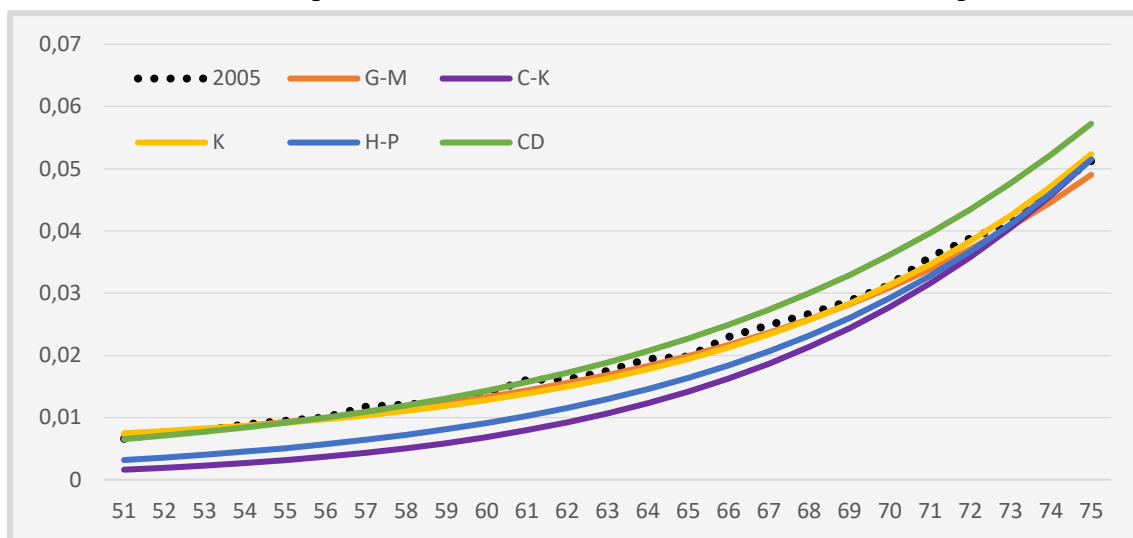
Graf č. 19 - Grafické porovnanie modelov medzi vekmi 26 až 50 rokov pre rok 2005



Zdroj: vlastné spracovanie

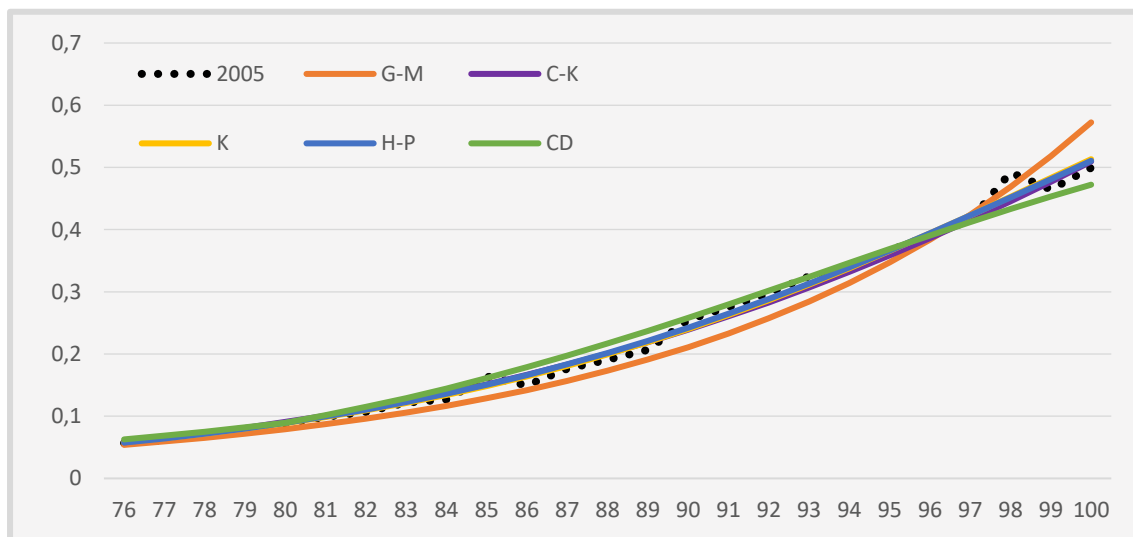
V ďalších intervaloch sa poradie trochu zmenilo. Pre interval od 51 do 75 rokov sa pre rok 2005 za najlepší model považuje Kannistov model s $MSE = 1,09436 \cdot 10^{-6}$ (Graf č. 20). Kannistov model vykazuje výborné výsledky aj medzi vekmi 76 až 100 rokov, avšak na tomto intervale ho o kúsok predbehol Helligmanov-Pollardov model (Graf č. 21).

Graf č. 20 - Grafické porovnanie modelov medzi vekmi 51 až 75 rokov pre rok 2005



Zdroj: vlastné spracovanie

Graf č. 21 - Grafické porovnanie modelov medzi vekmi 76 až 100 rokov pre rok 2005



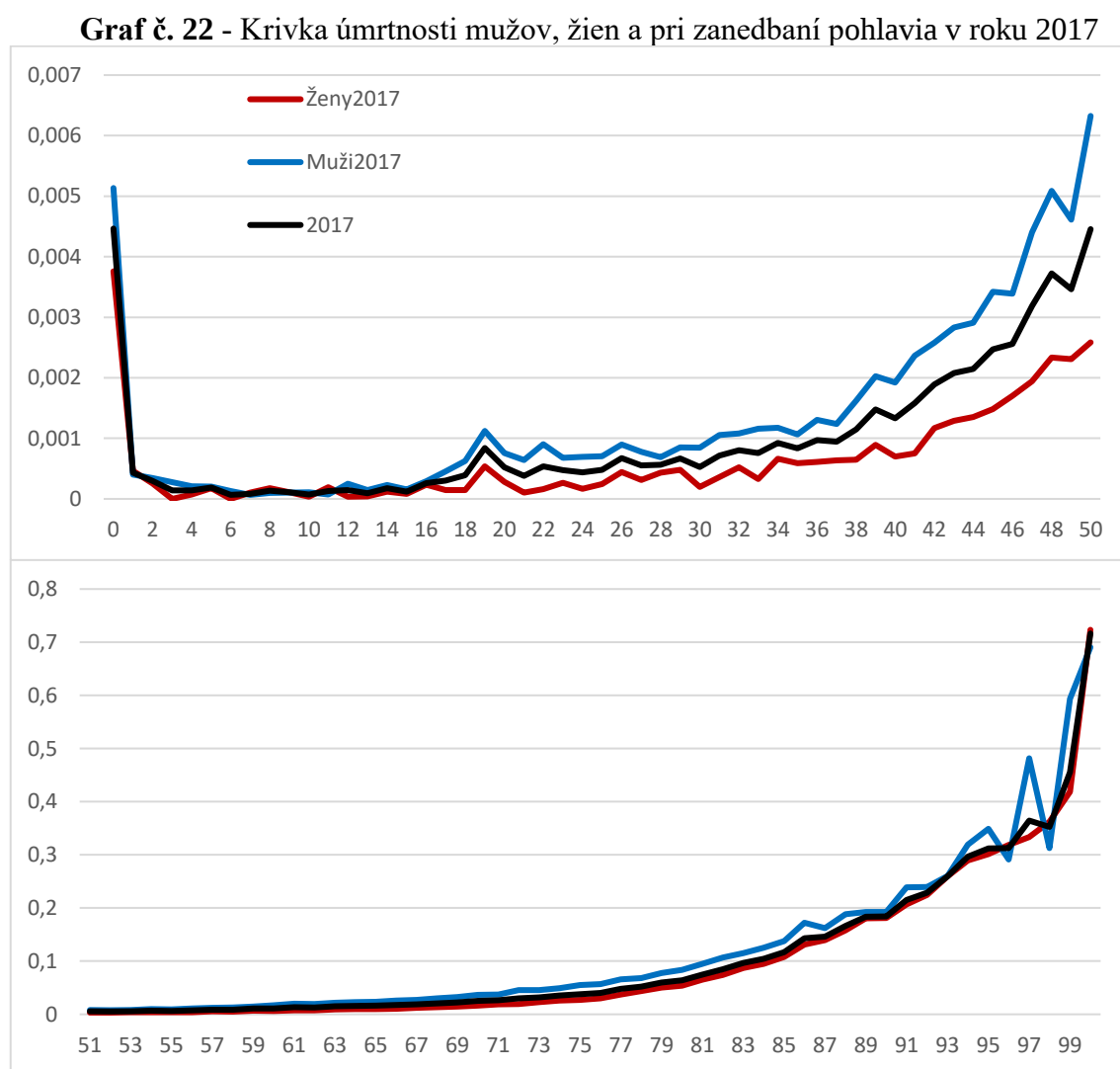
Zdroj: vlastné spracovanie

Ak sledujeme modely počas celého života populácie Slovenska v roku 2005, za najvhodnejší model môžeme považovať znova Kannistov model. Druhé a tretie miesto sa oproti roku 1993 vymenilo. Druhým najlepším sa pre tento rok stal Helligmanov-Pollardov

model. O niečo horšie výsledky vykazoval Coaleov-Kiskerov model. Treba poznamenať, že okrem Gompertzovho-Makehamovho modelu všetky modely dosiahli vynikajúce výsledky a sú vyhovujúce na modelovanie úmrtnosti v roku 2005.

4.2.3 Analýza v roku 2017

V časti 1.2.1 sme uviedli, že rozdiely úmrtnosti medzi mužmi a ženami sa stále viac približujú. Hoci to bol fakt, ktorý nezohľadňoval vek, na grafe č. 22 vidíme, že miery úmrtnosti mužov a žien sa naozaj približujú aj keď berieme do úvahy vekové rozloženie úmrtí. Pozorovateľné je to najmä medzi vekmi 51 až 100 rokov. Do budúcnosti je možné, že rozdiely sa budú natoľko zmenšovať, že faktor pohlavia naozaj nebude relevantný pri modelovaní úmrtnosti populácie.



Zdroj: vlastné spracovanie

Tabuľka č. 11 obsahuje výsledné hodnoty priemerného súčtu štvorcov rezíduí piatich modelov. V prvých štyroch riadkoch sú údaje uvedené pre intervaly s dĺžkou 25 rokov a v poslednom riadku sú uvedené hodnoty pre celý sledovaný vekový interval pri zanedbaní pohlavia ako faktora modelovania úmrtnosti. Grafické porovnanie vybraných modelov je zobrazené v grafoch č. 23 až 26.

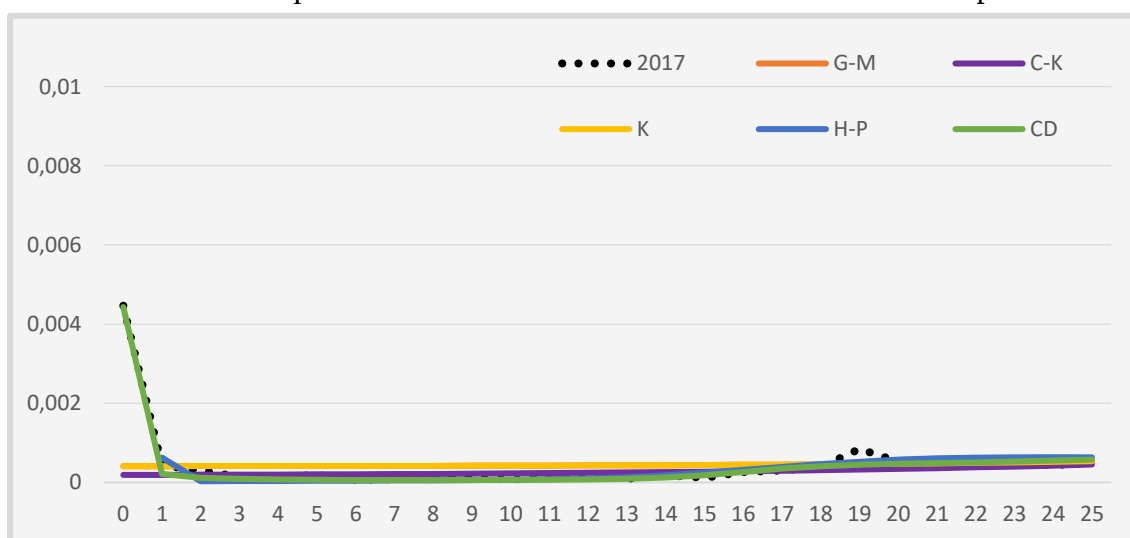
Tabuľka č. 11 - Hodnoty MSE pre rôzne vekové intervaly pre rok 2017

MSE	G-M	C-K	K	H-P	CD
0-25	6,86914E-07	7,24718E-07	6,86887E-07	7,843E-07	1,2569E-08
26-50	1,84799E-08	1,91987E-08	1,84603E-08	1,974E-08	3,1608E-08
51-75	6,36436E-06	2,05197E-05	3,80842E-05	7,71294E-05	1,4099E-05
76-100	0,001709342	0,001505952	0,002195066	0,002300848	0,004040649
0-100	0,000424861	0,000378031	0,000552941	0,000588815	0,001003662

Zdroj: vlastné spracovanie

Vo veku do 25 rokov je po tretíkrát najspoľahlivejší CoDe model pre $MSE = 1,2569 \cdot 10^{-8}$, a teda tento model možno považovať za najvhodnejší na tomto intervale. Na ostatných vekových intervaloch nevykazuje rovnako dobré výsledky v dôsledku jeho zložitosti, ktorá spôsobila problémy pri hľadaní optimálnych hodnôt parametrov. Z tohto dôvodu je už len na poistovni či využije model ako celok alebo ho zjednoduší a bude ním modelovať len úmrtnosť po obdobie dospelosti.

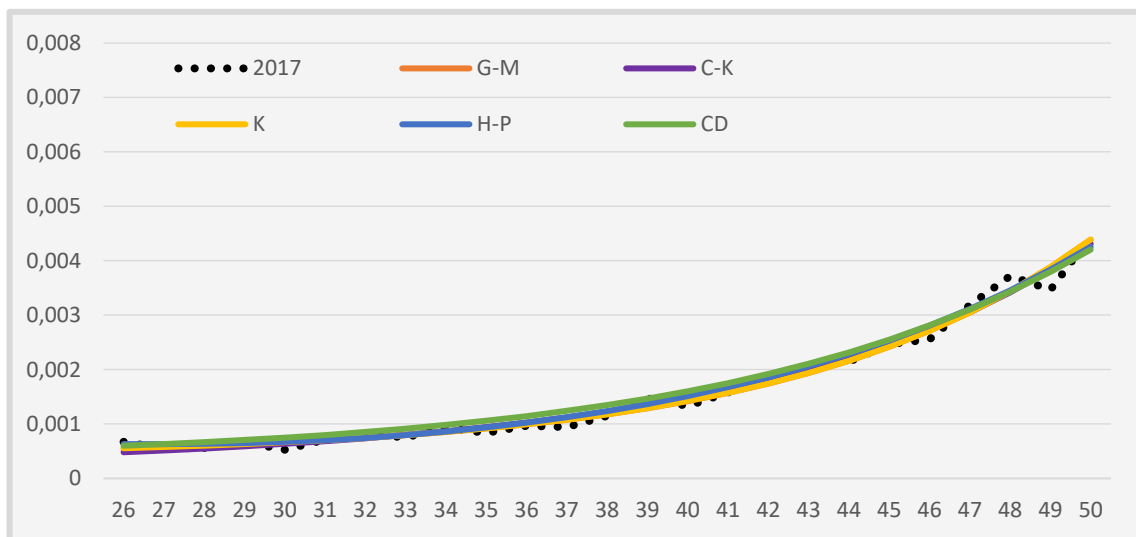
Graf č. 23 - Grafické porovnanie modelov od narodenia do 25 veku života pre rok 2017



Zdroj: vlastné spracovanie

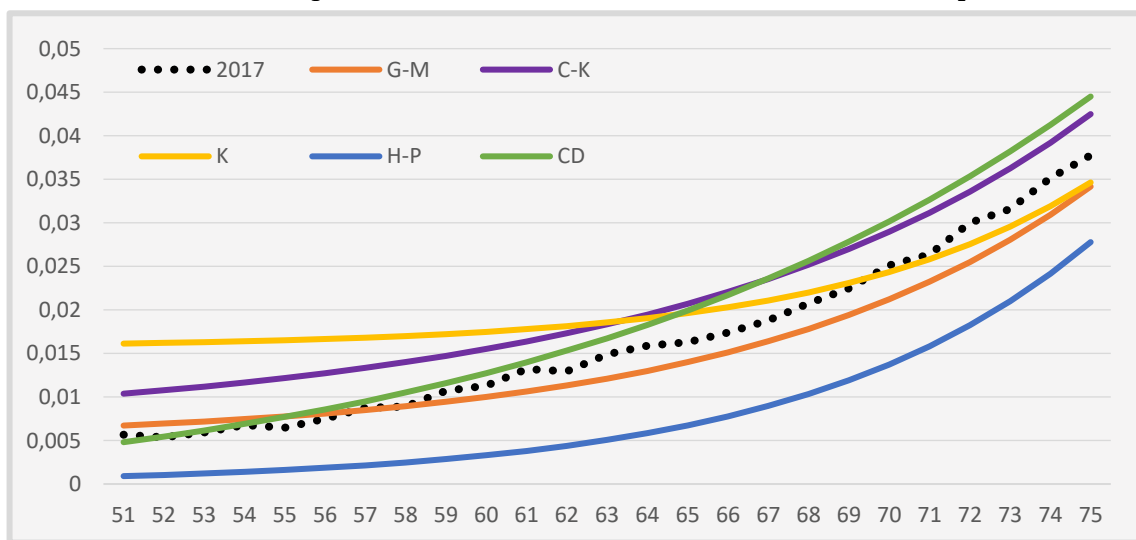
Medzi vekmi 26 až 75 rokov sa úmrtnostné správanie najlepšie podarilo opísať pomocou Gompertzovho-Makehamovho zákona úmrtnosti, ktorý sa považuje za najjednoduchší a najpoužívanejší z vybraných zákonov úmrtnosti. Ukazovateľ MSE má na intervale $\langle 26,50 \rangle$ hodnotu $1,84799 \cdot 10^{-8}$ a na intervale od 51 do 75 rokov hodnotu $6,36436 \cdot 10^{-6}$.

Graf č. 24 - Grafické porovnanie modelov medzi vekmi 26 až 50 rokov pre rok 2017



Zdroj: vlastné spracovanie

Graf č. 25 - Grafické porovnanie modelov medzi vekmi 51 až 75 rokov pre rok 2017

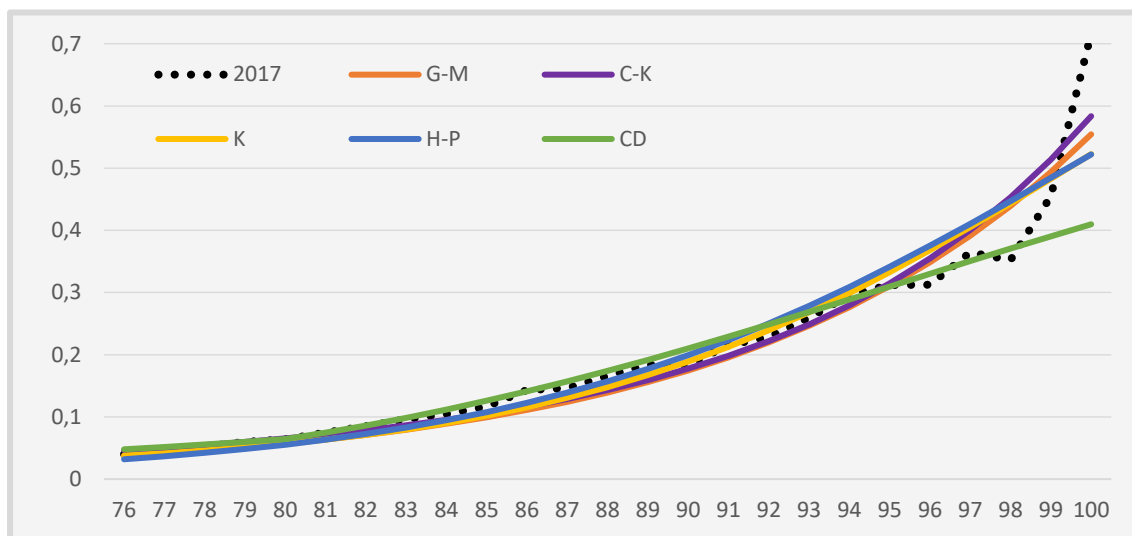


Zdroj: vlastné spracovanie

Vo vysokých vekoch mal najlepšie výsledky MSE Coaleov-Kiskerov model, konkrétne 0,001505952, čo veľkou mierou prispelo k jeho víťazstvu na celom vekovom

intervale. Predbehol Kannistov model úmrtnosti, ktorý sa tentokrát umiestnil až na treťom mieste. Prekvapivo ho predbehol najjednoduchší Gompertzov-Makehamov model, čo nás len utvrdzuje v tom, prečo hoci patrí k najstarším modelom, je stále veľmi často používaný.

Graf č. 26 - Grafické porovnanie modelov medzi vekmi 76 až 100 rokov pre rok 2017



Zdroj: vlastné spracovanie

Záver

Úmrtnostné správanie populácie v závislosti od veku je pre životnú poisťovňu nesmierne dôležitým poznatkom, na základe ktorého vykonáva mnohé aktuárske výpočty, ako napríklad kalkulácia poistného a technických rezerv. Čím presnejšie dokáže odhadnúť pravdepodobnosti úmrtia svojich poistencov, tým menšie riziko nesprávneho nastavenia poistného, respektíve nesprávneho uloženia finančných prostriedkov musí znášať.

Cieľom diplomovej práce bolo predovšetkým modelovanie úmrtnostného správania populácie Slovenskej republiky pomocou vybraných zákonov úmrtnosti v rokoch 1993, 2005, 2017 a ich vzájomné porovnanie.

V prvej kapitole sme sa snažili čitateľa voviesť do problematiky týkajúcej sa úmrtnosti v oblasti poisťovníctva. Oboznámili sme ho so základnými pojmami a charakteristikami životného poistenia. Ďalej sme bližšie opísali úmrtnosť, jej ukazovatele a analyzovali sme vývoj úmrtnosti v rokoch 1993 až 2017. Priblížili sme si najprielomovejšie smernice Európskej únie so zameraním na Gender smernicu, ktorá sa zásadne dotýka práve oblasti úmrtnosti, keďže pohlavie je významným faktorom jej modelovania.

Na začiatku 3. kapitoly sme upriamili pohľad na detailný opis zákonov úmrtnosti. Ukázali sme si ich základné vzťahy, prednosti, ale aj nedostatky, ktoré mali potom vplyv na presnosť modelov v praktickej časti. Venovali sme sa dôkladnej charakteristike metód nelineárnej regresnej analýzy, ktoré sa používajú na odhad parametrov nelineárnych regresných modelov. Vybrali sme si Gaussovu-Newtonovu metódu, Gradientnú metódu a Levenbergovu-Marquardtovu metódu, ktoré sú súčasťou analytického softvéru SAS Enterprise Guide. Za najefektívnejšiu môžeme považovať Levenbergovu-Marquardtovu metódu, ktorá spája prednosti ostatných dvoch spomínaných metód. Výsledné modely sme hodnotili na základe mier kvality modelov a presnosti ich prognóz, konkrétne priemerného súčtu štvorcov rezíduí MSE a štandardnej odchýlky rezíduí $RMSE$. Softvérové rozhranie SAS Enterprise Guide sme okrem základného opisu demonštrovali na Gompertzovom-Makehamovom zákone úmrtnosti spolu s výstupmi procedúry nelineárnej regresie.

Štvrtá a zároveň posledná kapitola je pre čitateľov najviac prínosná, pretože obsahuje aplikáciu zákonov úmrtnosti na reálne dáta o úmrtnosti populácie Slovenskej republiky v rokoch 1993, 2005 a 2017 na základe teoretických poznatkov z predchádzajúcich kapitol. Predmetom modelovania je vekový interval od 0 do 100 rokov, ktorý sme rozdelili na 2 časti

(od 0 do 50 rokov, od 51 do 100 rokov), čo prispelo k spresneniu výsledného modelu. Pre dosiahnutie ešte spoľahlivejších výsledkov odporúčame krivku rozdeliť na viac častí, kde prvá časť by modelovala kojeneckú úmrtnosť, druhá časť úmrtnosť do 20 rokov života, tretia časť takzvanú normálnu úmrtnosť pracujúcej populácie a štvrtá časť by sa zaoberala úmrtnosťou vo vyšších rokoch.

V prvom rade sme zohľadnili faktor pohlavia pri modelovaní úmrtnosti, vďaka čomu sme poukázali na rozdiely medzi úmrtnosťou mužov a žien. U mužov bola analýza náročnejšia než u žien, a to z dôvodu menšieho počtu dát a väčšej volatility dát najmä vo vysokých vekoch. Na rozlíšených dátach podľa pohlavia sme si ukázali správanie jednotlivých zákonov úmrtnosti a graficky sme si zvýraznili ich nedostatky. Hlavným nedostatkom starších modelov bolo zanedbanie dojčenskej úmrtnosti. CoDe model bol ako jediný schopný spoľahlivo odhadnúť hodnotu úmrtnosti do 1 roku života. 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 sme dosahovali lepšie výsledky na dátach v roku 1993, kedy tento hrb nebol taký výrazný. Na záver sme sa rozhodli rešpektovať ustanovenie Európskej únie o zákaze používania faktora pohlavia pri aktuárskych výpočtoch a modelovali sme miery úmrtnosti bez ohľadu na pohlavie. Zanedbanie pohlavia prispelo k zjemneniu výkyvov, čím sa zjednodušilo odhadovanie mier úmrtnosti, ale samozrejme stalo sa menej presnejším. Na intervale od 0 do 25 rokov dosahoval najlepšie výsledky CoDe model v každom zo sledovaných rokov. Na ostatných intervaloch sa poradie modelov menilo. V roku 1993 sa modelom, ktorý najpresnejšie opisuje úmrtnostné správanie populácie Slovenska stal Kannistov model. Najviac k tomu prispel fakt, že vo vysokých vekoch mal najpresnejšie výsledky aproximácie, čo bolo najproblémovejšou oblasťou pre ostatné modely. Svoje umiestnenie si Kannistov model obhájil aj v roku 2005. Nakoniec, v roku 2017 mal najlepšie výsledky Coaleov-Kiskerov model.

O žiadnom modeli sa však nedá prehlásiť, že by bol najlepší, alebo že by bol pre účely modelovania nevhodný na tomto vekovom intervale. Modely si navzájom veľmi dobre konkurovali. Otázkou pre poisťovňu zostáva, či stojí za to modelovať úmrtnosť pomocou zložitých zákonov úmrtnosti, keď tie jednoduché modely spoľahlivo opisujú úmrtnostné správanie populácie.

Zoznam obrázkov, tabuliek a grafov

Obrázok č. 1 - Dôsledky odstránenia pohlavia ako faktora z výpočtu poistného.....	20
Obrázok č. 2 - Schéma morálneho hazardu.....	21
Obrázok č. 3 - Schéma nepriaznivého výberu.....	22
Obrázok č. 4 - Časová os zákonov úmrtnosti.....	26
Obrázok č. 5 - Nelineárna regresná analýza v SAS Enterprise Guide.....	45
Obrázok č. 6 - Výstup softvéru SAS Enterprise Guide 1.....	49
Obrázok č. 7 - Výstup softvéru SAS Enterprise Guide 2.....	50
Obrázok č. 8 - Výstup softvéru SAS Enterprise Guide 3.....	50
Obrázok č. 9 - Výstup softvéru SAS Enterprise Guide 4.....	50
Obrázok č. 10 - Výstup softvéru SAS Enterprise Guide 5.....	50
Obrázok č. 11 - Výstup softvéru SAS Enterprise Guide 6.....	51
Tabuľka č. 1 – Základné parametrické funkcie ľudskej úmrtnosti.....	27
Tabuľka č. 2 - Odhad parametrov G-M pre rok 1993.....	55
Tabuľka č. 3 - Odhad parametrov G-M pre rok 2005.....	56
Tabuľka č. 4 - Odhad parametrov G-M pre rok 2017.....	57
Tabuľka č. 5 - Odhad parametrov C-K pre rok 1993.....	58
Tabuľka č. 6 - Odhad parametrov K pre rok 2005.....	60
Tabuľka č. 7 - Odhad parametrov H-P pre rok 2017.....	62
Tabuľka č. 8 - Odhad parametrov CoDe pre rok 2017.....	64
Tabuľka č. 9 - Hodnoty MSE pre rôzne vekové intervaly pre rok 1993.....	67
Tabuľka č. 10 - Hodnoty MSE pre rôzne vekové intervaly pre rok 2005.....	70
Tabuľka č. 11 - Hodnoty MSE pre rôzne vekové intervaly pre rok 2017.....	74
Graf č. 1 - Úmrtnosť na Slovensku podľa pohlavia v rokoch 1993 – 2017.....	16
Graf č. 2 - Úmrtnosť v SR podľa pohlavia a veku v rokoch 1993 a 2017.....	16
Graf č. 3 - Miery úmrtnosti mužov a žien v roku 2017.....	52
Graf č. 4 - Gompertzov-Makehamov model úmrtnosti mužov a žien v roku 1993.....	54
Graf č. 5 - Gompertzov-Makehamov model úmrtnosti mužov a žien v roku 2005.....	56

Graf č. 6 - Gompertzov-Makehamov model úmrtnosti mužov a žien v roku 2017.....	57
Graf č. 7 - Coaleov-Kiskerov model úmrtnosti mužov v roku 1993 pri ponechaní veku 0...58	
Graf č. 8 - Coaleov-Kiskerov model úmrtnosti mužov a žien v roku 1993.....	59
Graf č. 9 - Kannistov model úmrtnosti mužov a žien v roku 2005.....	61
Graf č. 10 - Helligmanov-Pollardov model úmrtnosti mužov a žien v roku 2017.....	63
Graf č. 11 - CoDe model úmrtnosti mužov a žien v roku 2017.....	65
Graf č. 12 - Krivka úmrtnosti mužov, žien a pri zanedbaní pohlavia v roku 1993.....	66
Graf č. 13 - Grafické porovnanie modelov od narodenia do 25 roku života pre rok 1993.....	67
Graf č. 14 - Grafické porovnanie modelov medzi vekmi 26 až 50 rokov pre rok 1993.....	68
Graf č. 15 - Grafické porovnanie modelov medzi vekmi 51 až 75 rokov pre rok 1993.....	68
Graf č. 16 - Grafické porovnanie modelov medzi vekmi 76 až 100 rokov pre rok 1993.....	69
Graf č. 17 - Krivka úmrtnosti mužov, žien a pri zanedbaní pohlavia v roku 2005.....	70
Graf č. 18 - Grafické porovnanie modelov od narodenia do 25 veku života pre rok 2005.....	71
Graf č. 19 - Grafické porovnanie modelov medzi vekmi 26 až 50 rokov pre rok 2005.....	71
Graf č. 20 - Grafické porovnanie modelov medzi vekmi 51 až 75 rokov pre rok 2005.....	72
Graf č. 21 - Grafické porovnanie modelov medzi vekmi 76 až 100 rokov pre rok 2005.....	72
Graf č. 22 - Krivka úmrtnosti mužov, žien a pri zanedbaní pohlavia v roku 2017.....	73
Graf č. 23 - Grafické porovnanie modelov od narodenia do 25 veku života pre rok 2017.....	74
Graf č. 24 - Grafické porovnanie modelov medzi vekmi 26 až 50 rokov pre rok 2017.....	75
Graf č. 25 - Grafické porovnanie modelov medzi vekmi 51 až 75 rokov pre rok 2017.....	75
Graf č. 26 - Grafické porovnanie modelov medzi vekmi 76 až 100 rokov pre rok 2017.....	76

Zoznam použitej literatúry

- [1] BILÍKOVÁ, M.: *Spojité metódy v poistnej matematike*. Bratislava: Vydavateľstvo EKONÓM, 2003. 96 s. ISBN 80-225-1698-8
- [2] BILÍKOVÁ, M. – JOHANESOVÁ, M.: *Aktuárske výpočty pre rôzne druhy poistenia m-tice osôb*. Bratislava: Vydavateľstvo EKONÓM, 2008. 86 s. ISBN 978-80-225-2530-5
- [3] CIPRA, T.: *Pojistná matematika - teorie a praxe*. Praha: EKOPRESS, 1999. 398 s. ISBN 80-86119-17-3
- [4] SIVAŠOVÁ, D.: *Aktuárska demografia v prostredí konkurenčného poistného trhu*. Bratislava: Vydavateľstvo EKONÓM, 2008. 98 s. ISBN 978-80-225-2509-1
- [5] PODMANICKÁ, Z. a kol.: *Štatistika v súvislostiach: Hlavné trendy populačného vývoja v SR v roku 2017*. Bratislava: Štatistický úrad SR, 2018. 125 s. ISBN 978-80-8121-670-1
- [6] MÉSZÁROS, J.: *Požiadavky modernizácie demografickej štatistiky vo vzťahu k novým metódam graduácie pravdepodobností úmrtia. Testovanie, aplikácia a zhodnotenie nových metód graduácie*. Bratislava: Výskumné demografické centrum INFOSTAT, 2016. 35 s. (Pracovný materiál)
- [7] WILLEMSE, W. J.: *Computational Intelligence: Mortality models for the actuary*. Delft. DUP Science, 2001. 188 s. ISBN 90-407-2160-2
- [8] OECD: *Mortality Assumptions and Longevity Risk: Implications for pension funds and annuity providers*. OECD Publishing, 2014. 194 s. ISBN 978-92-64-22274-8
- [9] SEBER, G. A. F. – Wild, C. J.: *Nonlinear Regression*. New Jersey: John Wiley & Sons, Inc, 2003. 768 s. ISBN 978-0-471-47135-6

- [10] ŠOLTÉS, E.: *Regresná a korelačná analýza s aplikáciami*. Bratislava: Iura Edition, 2008. 287 s. ISBN 978-80-8078-163-7
- [11] JANKOVIČ, M.: *Antidiskriminácia ohrozuje poisťovníctvo*. In *Poistné rozhľady*. Bratislava: Slovenská asociácia poisťovní, 2011. č. 5. ISSN 1335-1044
- [12] MOLIK, J.: *Jednotné poistné: hrozba alebo príležitosť*. In *Poistné rozhľady*. Bratislava: Slovenská asociácia poisťovní, 2012, č. 2. ISSN 1335-1044
- [13] SLASPO: *Stanovisko SLASPO k vplyvu antidiskriminačnej agendy Európskej únie na poisťovací priemysel*. In *Poistné rozhľady*. Bratislava: Slovenská asociácia poisťovní, 2011, č.5. ISSN 1335-1044
- [14] Dennis, J. E., Jr.: *Nonlinear least squares and equations*. New York: Academic Press, 1977. s. 269-312
- [15] MARQUARDT, D.: *An Algorithm for Least Squares Estimation of Nonlinear Parameters*. Journal of the Society for Industrial and Applied Mathematics, 1963. s. 431-441.
- [16] LEVENBERG, K.: *A method for the solution of certain non-linear problems in least squares*. Quarterly of Applied Mathematics, 1944. s. 164-168. ISSN 1552-4485

Elektronické zdroje:

- [17] GAVRILOVA, N. S. – GAVRILOV, L. A.: *Stárnutí a dlouhověkost: Zákony a prognózy úmrtnosti pro stárnoucí populace*. [online] Demografie, 2011. s. 109–128. [cit. 20.1.2019] Dostupné na: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4167024/>
- [18] DE BEER, J. – JANNSEN, F: *A new parametric model to assess delay and compression of mortality*. [online] Population Health Metrics, 2016. 21 s. [cit.

20.1.2019] Dostupné na: <https://pophealthmetrics.biomedcentral.com/articles/10.1186/s12963-016-0113-1>

- [19] Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). [dáta stiahnuté 2.2.2019] Dostupné na: www.mortality.org
- [20] SAS Documentation. [online] Dostupné na: <https://support.sas.com/documentation/onlinedoc/guide/>
- [21] SAS Documentation for Nonlinear regression procedure. [online] Dostupné na: <https://support.sas.com/documentation/onlinedoc/stat/141/nlin.pdf>
- [22] OXERA: *Equality in insurance pricing: the effect of the gender ban*. 2013. [online] Dostupné na: <https://www.oxera.com/agenda/equality-in-insurance-pricing-the-effect-of-the-gender-ban/>
- [23] OXERA: *Gender and insurance: impact on EU consumers of a ban on the use of gender*. 2011. [online] Dostupné na: <http://www.oxera.com/Latest-Thinking/Agenda/2011/Gender-and-insurance-impact-on-EU-consumers-of-a.aspx>
- [24] OXERA: *Gender and insurance: unintended consequences of unisex insurance pricing*. 2011. [online] Dostupné na: <http://www.oxera.com/Latest-Thinking/Agenda/2011/Gender-and-insurance-unintended-consequences-of-u.aspx>
- [25] OXERA: *The impact of a ban on the use of gender in insurance*. 2011. [online] Dostupné na: <https://www.oxera.com/publications/the-impact-of-a-ban-on-the-use-of-gender/>
- [26] 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>