

UNIVERZITA KARLOVA V PRAZE

Přírodovědecká fakulta
Katedra demografie a geodemografie



**Porovnání metod konstrukce oficiálních
úmrtnostních tabulek ve vybraných zemích**

Comparison of life tables construction in selected countries

Bakalářská práce

Ivan Godunov

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla použita k získání jiného nebo stejného akademického titulu.

V Praze, 15. 8. 2014

.....

Děkuji RNDr. Borisu Burcinovi, Ph.D. za ochotu, trpělivost, cenné rady a připomínky při tvorbě této práce. Dále bych chtěl poděkovat příslušným statistickým úřadům za poskytnutí dat o způsobech konstruování úmrtnostních tabulek a v neposlední řadě svým přátelům a rodičům za všeobecnou podporu.

Porovnání metod konstrukce oficiálních úmrtnostních tabulek ve vybraných zemích

Abstrakt

Cílem této práce je analýza metod, které se používají statistickými úřady vybraných zemí při konstrukci úmrtnostních tabulek. Byly prozkoumány základní rozdíly týkající se typů publikovaných životních tabulek a způsobů odhadu parametrů. Teoretická část práce nabízí přehled základních metodologických informací a popis přesné metodologie, která byla zjištěna při komunikaci s vybranými statistickými úřady. Během analýzy bylo zjištěno, že rozdíly při výpočtu jednotlivých funkcí úmrtnostních tabulek jsou minimální, proto je analytická část věnována především metodám odhadu pravděpodobnosti úmrtí ve věku 0 a vyrovnání křivky úmrtnosti ve starším věku. Získané postupy a metody byly aplikovány na data za ČR, což umožnilo samotné srovnání. Závěrečná část je pak celkovým zhodnocením dosažených výsledků spolu s komentáři k vybraným postupům a metodám.

Klíčová slova: úmrtnostní tabulka, srovnání, tabulkové funkce, pravděpodobnost úmrtí, vyrovnání křivky úmrtnosti

Comparison of life tables construction in selected countries

Abstract

The main goal of this work is to analyze the methods used by the selected statistical offices in the construction of mortality tables. This work explored fundamental differences regarding the type of published life tables and parameters estimation methods. The theoretical part provides an overview of basic methodological information and description of methodology used, which was acquired during communicating with the selected statistical offices. During the analysis it was found that the differences in calculating the various functions of life tables are minimal, so the analytical part is mainly devoted to methods of estimating the probability of death at age 0 and smoothing of probability of death in older age. Acquired procedures and methods were applied to the data for the Czech Republic, which allowed the comparison itself. The final part is the overall evaluation of achieved results, where can be also found commentaries on selected procedures and methods.

Keywords:

life table, comparison, life table components, probability of death, smoothing of probability of death

OBSAH

Přehled použitých zkratk	6
Seznam obrázků	7
Seznam tabulek	8
1 Úvod	9
1.1 Zdroje dat	10
1.2 Struktura práce	11
2 Charakteristika úmrtnostních tabulek	12
2.1 Klasifikace úmrtnostních tabulek	12
2.2 Funkce úmrtnostní tabulky	13
2.3 Odhad pravděpodobnosti úmrtí	16
2.4 Metody vyrovnání úmrtnostních tabulek	18
3 Popis oficiálních úmrtnostních tabulek vybraných zemí	24
3.1 Přehled vybraných zemí	24
3.1.1 Belgie	24
3.1.2 Česká republika	25
3.1.3 Kanada	27
3.1.4 Portugalsko	29
3.1.5 Singapur	30
3.1.6 Slovensko	31
3.1.7 Španělsko	33
3.1.8 Švédsko	35
3.1.9 Velká Británie	37
3.1.10 Estonsko, Grónsko a Irsko	40
4 Analýza a porovnání vybraných metod konstruování úmrtnostních tabulek	41
4.1 Kojenecká úmrtnost	41
4.2 Vyrovnání křivek úmrtnosti	42
5 Závěr	48
Seznam použité literatury	50
Seznam datových zdrojů	53

PŘEHLED POUŽITÝCH ZKRATEK

ČSÚ	Český statistický úřad
GBA	databáze obecných osobních záznamů
GM	Gompertz-Makeham
HP	Heligman-Pollard
HPC	úmrtnostní standard
KT	Kannisto-Thatcher
NUTS	nomenklatura územních statistických jednotek
ÚT	úmrtnostní tabulka

SEZNAM OBRÁZKŮ

Obr. 1: Grafické znázornění tabulkových funkcí, ČR, muži a ženy, 2010	17
Obr. 2: Modelované pravděpodobnosti úmrtí, Švédsko, ženy, 2013	37
Obr. 3: Porovnání křivek pravděpodobnosti úmrtí v nejvyšších věcích podle vybraných modelů, ČR, 2008, muži	42
Obr. 4: Porovnání křivek pravděpodobnosti úmrtí v nejvyšších věcích podle vybraných modelů, ČR, 2008, ženy	43
Obr. 5: Hodnoty naděje dožití ve věku 65 až 90 let při aplikaci prezentovaných modelů, ČR, 2008, muži	44
Obr. 6: Hodnoty naděje dožití ve věku 95 až 120+ let při aplikaci prezentovaných modelů, ČR, 2008, muži	45
Obr. 7: Hodnoty naděje dožití ve věku 65 až 90 let při aplikaci prezentovaných modelů, ČR, 2008, ženy	46
Obr. 8: Hodnoty naděje dožití ve věku 95 až 120+ let při aplikaci prezentovaných modelů, ČR, 2008, ženy	46
Obr. 9: Hodnoty indexu determinace při aplikaci prezentovaných modelů, ČR, 2008, muži a ženy	47

SEZNAM TABULEK

Tab. 1: Přehled nejpoužívanějších modelů vyhlazování a extrapolace křivky úmrtnosti	23
Tab. 2: Předpokládané průměrné věky při úmrtí do 1 roku, Velká Británie, 2013, chlapci a dívky.....	38
Tab. 3: Charakteristiky úmrtnostních tabulek za Estonsko, Grónsko a Irsko.....	40
Tab. 4: Porovnání pravděpodobnosti úmrtí ve věku 0 podle metod vybraných modelů, ČR, 2008, chlapci a dívky.....	42
Tab. 5: Hodnoty naděje dožití ve vybraných věcích při aplikaci prezentovaných modelů, ČR, 2008, muži.....	44
Tab. 6: Hodnoty naděje dožití ve vybraných věcích při aplikaci prezentovaných modelů, ČR, 2008, ženy.....	45
Tab. 7: Hodnoty indexu determinace při aplikaci prezentovaných modelů, ČR, 2008, muži a ženy.....	47

Kapitola 1

Úvod

Úmrtnostní tabulka¹ je statistická tabulka kvantitativně popisující proces vymírání populace, je jednou z nejpoužívanějších metod analýzy demografických dat (Hinde, 1998). Hlavní příčinou vzniku těchto tabulek byla snaha nalézt pravidelnosti ve struktuře přežívajících a zemřelých na základě agregovaných dat. Tabulky života jsou zvláštním nástrojem měření úrovně úmrtnosti především v rámci demografických studií, mohou být ale také použity v jiných studiích, které se zabývají populací. Tabulky života jsou nezbytným a užitečným nástrojem při analýze plodnosti, procesů reprodukce, migrace a struktury obyvatelstva: při odhadu nebo projekci velikosti obyvatelstva, jeho struktury a dalších ekonomických charakteristik populace.

V nejjednodušším případě tabulky života je celá tabulka generována z věkově specifických měr úmrtnosti a jejich výsledné hodnoty se používají k výpočtu ukazatele střední délky života. V jiných případech se míry úmrtnosti kombinují s jinými demografickými daty v rámci složitějších modelů, které vypočítají kombinovaný efekt mortality a jedné nebo více sociodemografických charakteristik. Příkladem je tabulka ekonomické aktivity, kde se kombinují míry úmrtnosti a míry ekonomické aktivity, a výsledkem této tabulky je jejich kombinovaný efekt na ekonomickou aktivitu celkem (Bryan, Heuser 2004). Standardní tabulka života se skládá ze šesti funkcí (tj. sloupců dat), včetně pravděpodobnosti úmrtí, pravděpodobnosti dožití, tabulkového počtu zemřelých, tabulkového počtu žijících, počtu zbývajících člověkolet, které má tabulková populace před sebou a naděje dožití, a to vše podle věku.

Jednou z hlavních výhod úmrtnostních tabulek oproti jiným metodám měření úmrtnosti je, že neodrážejí vliv věkové struktury skutečného počtu obyvatel a nevyžadují přijetí standardní populace pro přijatelné srovnání míry úmrtnosti v různých populacích. Další předností je, že úmrtnostní tabulka snadno povoluje vytváření úmrtnostních příspěvků pro věkové kohorty², což eliminuje těžké sestavování statistik úmrtí pro věkové kohorty z ročních úmrtnostních statistik podle věku (Bryan, Heuser 2004).

¹ V rámci práce jsou pojmy úmrtnostní tabulka a tabulka života považovány za synonyma.

² Termín kohorta v demografii nejčastěji označuje soubor osob narozených v určitém daném období. Jestliže toto období není určeno, jde zpravidla o jeden rok a pak je synonymem termínu ročník. Ve stejném smyslu se používá také termínu generace. Někdy rozlišujeme mužskou generaci a ženskou generaci. (Mnohojazyčný demografický slovník [český svazek], 2005, 713-1)

Během přípravy práce byl kromě studia odborné literatury kladen velký důraz na komunikaci s oficiálními statistickými úřady vybraných zemí, aby byly zjištěny potřebné informace, a na hledání vhodné metodické literatury. Jako pracovní hypotéza bylo ustanoveno tvrzení, že metodiky konstrukce úmrtnostních tabulek³ různých zemí mají určité rozdíly a lokální zvláštnosti.

Za cíl si tato práce stanovuje porovnání metod konstrukce úmrtnostních tabulek ve vybraných zemích světa. Toto porovnání je prováděno na základě různých metod výpočtu ukazatele pravděpodobnosti úmrtí a vyrovnání či extrapolace křivky úmrtnosti. Analýza je doplněna o mezinárodní srovnání periodicity, územního pokrytí a typů úmrtnostních tabulek vybraných zemí.

1.1 Zdroje dat

Nejdůležitějšími zdroji pro získání potřebných dat a informací jsou především oficiální publikace příslušných statistických úřadů, které jsou k dispozici buď v tištěné či elektronické podobě na příslušných webových stránkách. Seznam relevantních statistických úřadů a dalších organizací je k dispozici na stránkách Českého statistického úřadu⁴.

Komunikace se statistickými úřady probíhala ve formě elektronického dopisu, který obsahoval dotazník zahrnující 4 otázky ohledně metod konstruování úmrtnostních tabulek.

Tyto otázky se týkají následujících okruhů:

1. Jak často jsou úmrtnostní tabulky publikovány?
2. Jaký je nejvyšší věk nebo věkový interval v tabulkách života?
3. Jaká metoda se používá při konstruování pravděpodobnosti úmrtí q_x ?
4. Jaká metoda se používá při konstruování a případném vyrovnání křivky úmrtnosti ve starším věku?

Mimo elektronickou komunikaci jsou informace o metodách konstruování dostupné na příslušných webových stránkách jednotlivých statistických úřadů.

V případě Belgie poskytují některá data stránky belgického statistického úřadu (http://statbel.fgov.be/fr/statistiques/chiffres/population/decès_mort_esp_vie/tables/).

Z podobných stránek lze čerpat informace o ÚT, které publikuje statistický úřad Španělska: <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft20%2Fp319a&file=inebase&L=1> a http://www.ine.es/en/daco/daco42/mortalidad/metodo_9111_en.pdf). Tabulky života za Nizozemsko jsou dostupné pro přehled na níže uvedených internetových odkazech (<http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=37360NED&D1=a&D2=a&D3=a&D4=1&HD=140709-0942&HDR=G1,T&STB=G2,G3>, <http://www.cbs.nl/NR/rdonlyres/9394039F-005A-4CF0-A576C98C2C12AE8E/0/2007k3b15p66art.pdf>). ÚT pro Irsko jsou taky dostupné na webovém odkázu (<http://www.cso.ie/en/media/duplicatecsomedia/newmedia/releasespublications/documents/birthsdm/current/irishlife.pdf>). Metodologické poznámky k úmrtnostním tabulkám Portugalska jsou

³ Pod pojmem úmrtnostní tabulka se v práci rozumí pouze ty úmrtnostní tabulky, které jsou vytvářené a publikované oficiálními statistickými úřady ve vybraných zemích.

⁴ Viz http://www.czso.cz/csu/redakce.nsf/i/statisticke_urady_ve_svetu.

taktéž k dispozici online (<http://smi.ine.pt/DocumentacaoMetodologica/Detalhes/1239>). Údaje za Velkou Británii jsou dostupné v online podobě na příslušných stránkách britského statistického úřadu (<http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/demography/guide-to-calculating-national-life-tables/index.html>) a taky na odkazu <http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/pop-ests/calculating-estimates-of-the-very-elderly/index.html>). Zdrojem informací o úmrtnostních tabulkách Singapurů je internetová publikace Complete Life Tables 2003-2006 for Singapore Resident Population, kterou publikuje statistický úřad Singapurů (http://www.singstat.gov.sg/publications/publications_and_papers/births_and_deaths/ip-s13.pdf). Metodologické poznámky k vytvoření úmrtnostních tabulek pro Českou republiku jsou taky k dispozici na webu (http://www.czso.cz/csu/redakce.nsf/i/umrtnostni_tabulky_metodika). Pro ostatní neuvedené země je hlavním zdrojem informací elektronická komunikace s pracovníky příslušných statistických úřadů, kteří jsou zodpovědní za konstruování úmrtnostních tabulek.

1.2 Struktura práce

Tato práce je rozdělena na tři části. První z nich obsahuje základní informace týkající se dané bakalářské práce. V ní je stanovena pracovní hypotéza a cíl práce, stručně popsán proces sběru dat a informací, a také problémy s tím spojené. Úvodní část také poskytuje přehled využitých zdrojů a odkazů na příslušné dokumenty či statistické databáze. Druhá část se zabývá především metodikou. Tato kapitola popisuje základní informace o demografickém nástroji úmrtnostní tabulky, zahrnuje nezbytné informace o klasifikaci ÚT, důležitých ukazatelích a vztazích mezi nimi – resp. metody výpočtu pravděpodobnosti úmrtí q_x , a dává přehled metod vyrovnání ÚT. Tato část druhé kapitoly je nejvíc navázaná na odbornou literaturu týkající se daného tématu. Třetí část této práce je věnovaná analýze jednotlivých zemí, resp. popisu metod, které se používají při tvorbě úmrtnostních tabulek. Následující analytická část má za úkol porovnat metody konstrukce úmrtnostních tabulek vybraných statistických úřadů. V rámci této analýzy budou na data za ČR aplikovány různé metody, což umožní analýzu a porovnání různých postupů. V závěru této kapitoly jsou uvedeny výsledky, ke kterým bylo dospěno během zpracování této bakalářské práce.

Kapitola 2

Charakteristika úmrtnostních tabulek

V následující kapitole jsou rozebrány různé typy klasifikace úmrtnostních tabulek, popsány jednotlivé ukazatele, které jsou základem samotné úmrtnostní tabulky, včetně detailnějšího popisu metod konstrukce jednotlivých ukazatelů a různých způsobů vyrovnávání úmrtnostních tabulek. Popis těchto charakteristik je před samotnou analýzou metod konstruování úmrtnostních tabulek, které používají vybrané statistické úřady ve světě, nezbytný.

2.1 Klasifikace úmrtnostních tabulek

Jedním ze základních dělení úmrtnostních tabulek je členění podle typu sledované populace. Z tohoto pohledu se rozlišují dva typy úmrtnostních tabulek:

Prvním typem jsou průřezové úmrtnostní tabulky (current life table, period life table), které jsou určeny pro popis úmrtnosti na určitou kalendářní dobu, nejčastěji na jeden rok. Základem této tabulky je hypotetické sledování narozených v daném roce za předpokladu, že tato generace bude mít v každém věku stejnou míru úmrtnosti, jaká byla v daném roce pro každou věkovou skupinu. Například úmrtnostní tabulka pro rok 1990 předpokládá, že osoby, které se narodili v dané generaci v roce 1990, budou mít v každém věku od 0 do 100 let během celého života stejné míry úmrtnosti, které převládaly v každém věku od narození až do 100 let v roce 1990. Tento druh úmrtnostních tabulek je nejpoužívanější, navíc v praxi je často synonymem samotného pojmu životní tabulka (Pavlík, 1986).

Druhým typem je generační nebo kohortní úmrtnostní tabulka (cohort life table), kterou charakterizuje to, že míry úmrtnosti v příslušné kohortě jsou sledovány do okamžiku úmrtí posledního zástupce dané kohorty. Například roční pravděpodobnosti úmrtí osob narozených v roce 1900 lze sledovat až do roku 2000 a jejich míry úmrtnosti je možné získat v každém věku od narození až do věku 100 let. Pokud máme k dispozici tyto míry, můžeme vytvořit úmrtnostní tabulku za danou kohortu za předpokladu, že většina osob zemřela v roce 2000. Aby bylo možné vytvořit generační tabulku života, musí být úmrtnostní a migrační údaje sbírány po dlouhou dobu. V praxi lze tento postup použít jenom pro uzavřené populace bez uvažování migrace. Význam kohortní úmrtnostní tabulky je spíše historický, protože odráží intenzitu úmrtnosti jedinců

narozených dávno, kteří žili během svého života v odlišných podmínkách než pozorovatel v okamžiku přípravy tabulky.

Dalším základním rozdělením je rozřídění na úplné a zkrácené úmrtnostní tabulky. Tento rozdíl spočívá v délce věkového intervalu, ve kterém jsou uvedena data. Úplná neboli podrobná úmrtnostní tabulka obsahuje údaje za každý rok od narození do posledního platného věku, kdy sledovaný soubor osob zemře. Na druhé straně zkrácená úmrtnostní tabulka obsahuje údaje podle určitého věkového intervalu, zpravidla pětiletého (Hinde, 1998).

Pro země s nespolehlivou demografickou evidencí obyvatel nebo s absencí rozvinuté statistické evidence ve státě se obvykle připravuje spíše jednoduchá zkrácená úmrtnostní tabulka než komplikovanější úplná. Hodnoty pro pětileté nebo desetileté intervaly jsou dostatečně přesné pro většinu účelů a tvorba zkrácené tabulky je méně obtížná. Navíc použití zkrácené ÚT je často mnohem výhodnější varianta. Občas jsou základní hodnoty z úplné úmrtnostní tabulky uvedeny pouze pro každý pátý rok, aby se šetřilo na prostoru. (Siegel, 2004). Avšak údaje úplných úmrtnostních tabulek se mohou lišit od údajů ze zkrácených ÚT, a to zejména u starších věkových skupin, v důsledku rozdílů v metodách výpočtu.

2.2 Funkce úmrtnostní tabulky

Úmrtnostní tabulka sestává z řady ukazatelů kvantitativně popisujících vymírání sledované populace. Některé funkce úmrtnostní tabulky se vztahují k přesnému věku (l_x a q_x), další odkazují k dokončenému věku nebo k věkovým skupinám (q_x , d_x , L_x a T_x) (Siegel, 2004). Základním ukazatelem je pravděpodobnost úmrtí q_x , která je odvozena ze specifických měř úmrtnosti m_x . Vstupními daty pro úmrtnostní tabulku jsou P_x – počet žijících osob ve věku x a D_x – počet zemřelých osob ve věku x . Za P_x se zpravidla označuje střední stav obyvatelstva, tj. počet obyvatel podle věku v daném kalendářním roce, kteří se dožijí středu sledovaného období.

Ukazatel m_x vyjadřující specifickou míru úmrtnosti je počítán jako poměr zemřelých D_x k počtu obyvatel středního stavu P_x :

$$m_x = \frac{D_x}{P_x}. \quad (1)$$

Ukazatel q_x označuje pravděpodobnost úmrtí, tj. pravděpodobnost toho, že x -letá osoba se nedožije věku $x+1$, tedy že zemře před dosažením věku let $x+1$ let:

$$q_x = 1 - e^{-m_x}. \quad (2)$$

Funkce q_x je založena na pozorovaných věkově specifických mírach úmrtnosti a je nezávislá na ostatních funkcích tabulky a pozorované věkové struktuře. Všechny další funkce tabulky jsou závislé na funkci q_x (Siegel, 2004). Pro vyšší věky se pravděpodobnost úmrtí vyhlazuje pomocí Gompertz-Makehamova modelu a dalších možností vyrovnání křivky úmrtnosti.

Zvlášť se počítá pravděpodobnost úmrtí ve věku 0 (tzv. kojenecká úmrtnost), která je rovna podílu zemřelých ve věku 0 a živě narozených:

$$q_0 = \frac{D_0}{N^v} , \quad (3)$$

kde N^v je počet živě narozených osob v daném období a v dané populaci.

Doplňkem k pravděpodobnosti úmrtí q_x je pravděpodobnost dožití p_x neboli pravděpodobnost toho, že se osoba dožije věku $x+1$:

$$p_x = 1 - q_x , \quad (4)$$

Rovněž platí následující:

$$p_x = e^{-m_x}$$

$$p_x + q_x = 1.$$

Dalším ukazatelem úmrtnostních tabulek je tabulkový počet dožívajících označovaný jako l_x , tj. hypotetický počet osob, které se dožijí věku x let ze zvolené hodnoty počtu živě narozených. Tato hodnota se označuje jako l_0 a je tzn. kořenem tabulky, který nabývá hodnoty odmocniny z 10, zpravidla 100 000:

$$l_{x+1} = p_x * l_x . \quad (5)$$

Funkce l_x není ovlivněna úmrtností ve věku x a všech dalších věkových kategorií; je ovlivněna pouze mírami úmrtnosti před věkem x (Siegel, 2004).

Rovněž platí, že:

$$q_x = \frac{l_x + l_{x+1}}{l_x} \quad \text{a} \quad p_x = \frac{l_{x+1}}{l_x} .$$

Tabulkový počet jedinců, kteří zemřeli ve věku x , se označuje jako d_x – je počítán jako rozdíl dvou po sobě jdoucích tabulkových počtů dožívajících. Platí následující vztah:

$$d_x = l_x - l_{x+1} , \quad (6)$$

resp.:

$$d_x = l_x * q_x .$$

Další ukazatel L_x vyjadřuje tabulkový počet žijících, tj. hypotetický průměrný počet žijících v dokončeném věku x let (Hinde, 1998). Počítá se jako průměr ze dvou po sobě jdoucích tabulkových počtů dožívajících (kromě věku 0). Jedná se o tzv. celkový počet “člověkoroků”, které osoby ve věku x let do věku $x+1$ let prožijí:

$$L_x = \frac{l_x + l_{x+1}}{2}, \text{ kde } x \neq 0. \quad (7)$$

Tabulkový počet žijících ve věku 0 (L_0) je vypočítán z přesného rozložení zemřelých kojenců v daném období podle ročníku narození. Koeficient α (alfa) znamená, jaký podíl zemřelých ve věku 0 v daném roce (ve III. hlavním souboru událostí) pochází z generace narozených daného roku:

$$L_0 = l_0 - \alpha * d_0. \quad (8)$$

Funkce d_x a L_x jsou ovlivněny pouze měrami úmrtnosti od začátku tabulky života přes věk x (Siegel, 2004).

Ukazatel T_x vyjadřuje počet let života, které má tabulková generace (nikoliv jednotlivec) v daném věku ještě před sebou. Tento ukazatel se spočítá jako kumulace počtu žijících osob od nejvyššího věku v tabulce až po námi zjišťovaný věk:

$$T_x = T_{x+1} + L_x, \\ T_x = \sum_{\omega-1}^x L_x. \quad (9)$$

Funkce T_x je závislá na všech mírách úmrtnosti nad věkovou škálou, protože v sobě spojuje všechny L_x , z nichž každé je určeno pro všechny míry úmrtnosti podle věku x (Siegel, 2004).

Posledním ukazatelem úmrtnostní tabulky je e_x , který označuje střední délku života neboli naděje dožití, tj. průměrný počet let, kterých se ještě dožije jedinec ve věku x :

$$e_x = \frac{T_x}{l_x}. \quad (10)$$

Nejčastěji se používá pojem „střední délka života při narození“ nebo „naděje dožití při narození“, který vyjadřuje průměrnou délku života osoby právě narozené za předpokladu setrvání úmrtnostních poměrů platných v roce, ve kterém se osoba narodila, a pro který je konstruována úmrtnostní tabulka. Například je-li střední délka života při narození pro muže v roce 2002 v ČR 72,1 let, mají chlapi narození v tomto roce statistickou naději dožít se právě tohoto věku.

Naděje dožití se vzhledem k odlišnosti v úmrtnosti obou pohlaví zásadně uvádí zvlášť za muže a za ženy. Úmrtnost v předchozích věcích nemá žádný vliv na střední délku života (Siegel, 2004).

Co se týče zkrácených úmrtnostních tabulek, všechny funkce tabulky života lze rozdělit do dvou skupin: ty, které jsou stejné v obou úplných a zkrácených tabulkách života (l_x , T_x a e_x) a ty, které se liší (q_x , d_x a L_x) (Siegel, 2004).

Při konstruování zkrácené úmrtnostní tabulky lze postupovat dvěma způsoby. Buď se tabulka počítá přímo z dat pro n -leté intervaly anebo se konstruuje na základě tabulky úplné. V druhém případě lze neměnicí se funkce první skupiny jednoduše přepsat z jedné tabulky do druhé pro věky 0, 1, 5, 10, 15, ..., 95 a 100+. Funkce q_x , d_x a L_x musí být přejmenovány na ${}_nq_x$, ${}_nd_x$ a ${}_nL_x$ (resp. ${}_5q_x$, ${}_5d_x$ a ${}_5L_x$ v případě pětiletých věkových intervalů) aby mohly reprezentovat věkové skupiny, a jejich jednoleté hodnoty musí být kombinovány k získání požadovaných skupinových hodnot. Například:

$${}_5d_{55} = d_{55} + d_{56} + d_{57} + d_{58} + d_{59} ,$$

$${}_5q_{55} = \frac{{}_5d_{55}}{l_{55}} ,$$

$${}_5L_{55} = (L_{55} + L_{56} + L_{57} + L_{58} + L_{59})/5 .$$

V evropských zkrácených úmrtnostních tabulkách se L_x zpravidla násobí rozpětím intervalů při výpočtu stacionární populace.

2.3 Odhad pravděpodobnosti úmrtí

Při sestavování úmrtnostních tabulek lze postupovat buď pomocí přímé nebo nepřímé metody. Zásadním cílem je výpočet vstupní charakteristiky úmrtnostní tabulky – pravděpodobnosti úmrtí q_x . Přímá metoda zahrnuje tři možnosti výpočtu pravděpodobnosti úmrtí:

1) I. hlavní soubor událostí⁵:

$$q_x = \frac{{}_t^z D_x + {}_{t+1}^z D_x}{{}_x^z P_x + {}_t^z D_x}$$

Předpokladem pro následující výpočty je, že k dispozici jsou hlavní soubory zemřelých tříděné na elementární soubory dolní a horní, tedy nejen podle kalendářních let, ale také podle generace (roku narození).

2) II. hlavní soubor událostí⁶:

Tento způsob se týká jenom jedné generace a předpokládá znalost elementárních souborů zemřelých.

$$q_x = \frac{{}_x^z D_{x+1} + {}_x^z D_x}{{}_x^z P_x} ,$$

kde:

${}_x^z D_{x+1}$ je počet zemřelých generace z ve věku $x+1$,

${}_x^z D_x$ počet zemřelých generace z ve věku x a

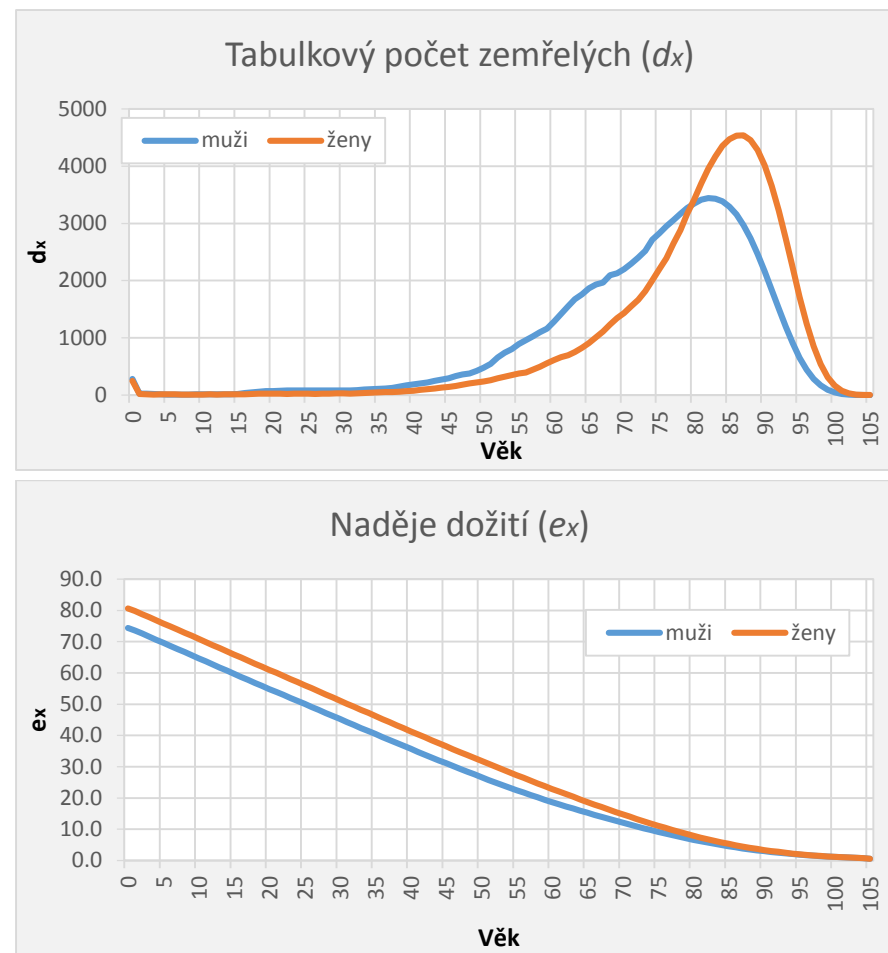
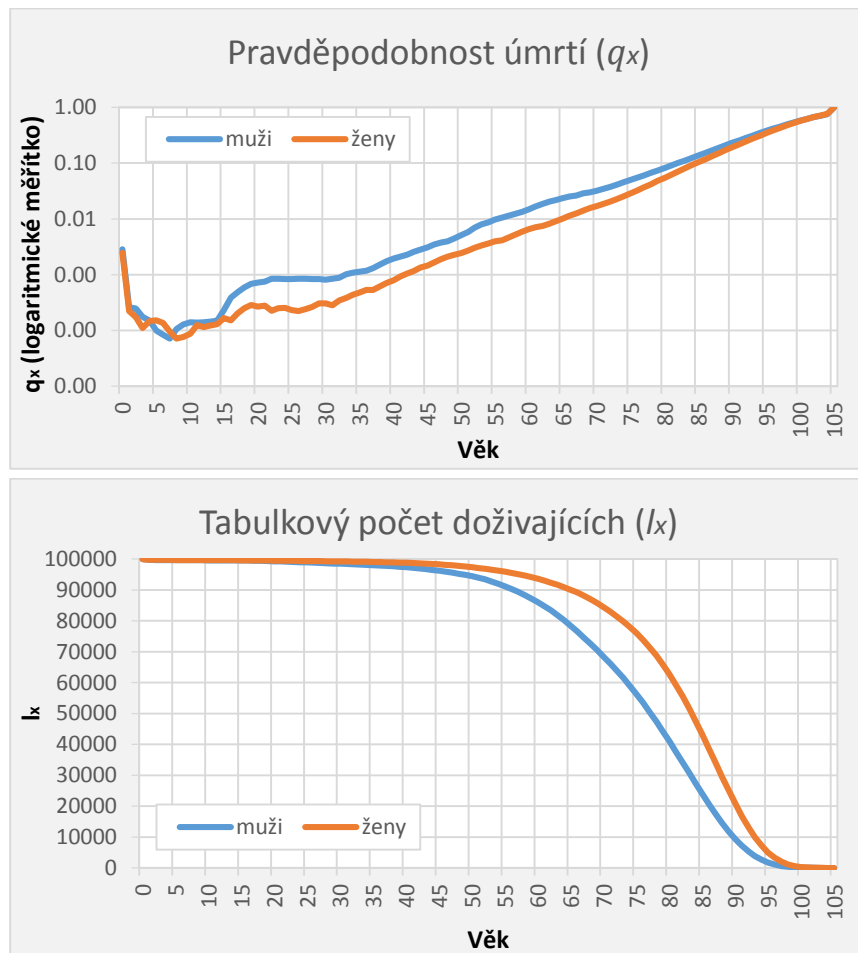
${}_x^z P_x$ počet obyvatel generace z ve věku x .

Jestliže jsou k dispozici údaje ze sčítání lidu a o počtu zemřelých následujícího roku, které již byli roztříděné do jednotlivých elementárních souborů, je tento způsob stejně přesný jako předcházející (Pavlík, 1986).

⁵ jsou to události, které se přihodily jedné generaci v průběhu jednoho roku věku x a dvou kalendářních let $t, t+1$.

⁶ je vymezen jedním rokem narození z , jedním kalendářním rokem t , avšak dvěma lety dokončeného věku x a $x+1$.

Obr. 1 – Grafické znázornění tabulkových funkcí, ČR, muži a ženy, 2010



Zdroj: ČSÚ, 2010, vlastní výpočty

3) III. hlavní soubor demografických událostí⁹:

$$q_x = \frac{{}^{z-1}D_x}{{}^{z-1}P_x} + \frac{{}^zD_x}{{}^zP_x} - \left(\frac{{}^zD_x}{{}^zP_x} \right) * \left(\frac{{}^{z-1}D_x}{{}^{z-1}P_x} \right),$$

kde ξ je přesný věk

Poslední násobek se vzhledem k malé velikosti obou jeho členů zpravidla zanedbává (Pavlík, 1986).

Nepřímá metoda je založena na použití tabulkových měr úmrtnosti m_x , které se často v praxi nahrazují mírami úmrtnosti podle věku u_x . Následně se z nich vypočítá q_x , přičemž výpočet tohoto ukazatele vyžaduje předpoklad buď rovnoměrného rozložení zemřelých v jednotlivých hlavních souborech demografických událostí nebo exponenciálního rozložení zemřelých celkem. Přestože první předpoklad zpravidla platí v praxi skoro vždycky, problém exponenciálního rozložení se nejčastěji projevuje při měření kojenecké úmrtnosti a úmrtnosti v nejstarších věkových kategoriích (Caselli, Vallin, 2005). Každopádně výpočet úmrtnosti nejstarších věkových kategorií způsobuje další problémy, které budou rozebrány v následující kapitole. Nepřímá metoda má tři výhody: data jsou často dostupná ve vhodném tvaru, týkají se jednoho kalendářního roku a zároveň identifikují každý rok věku (Caselli, Vallin, 2005). Pro tento postup je charakteristická menší přesnost, avšak jeho výhodou je, že nevyžaduje znalost o elementárních souborech demografických událostí zemřelých (Pavlík, 1986).

Platí-li $u_x \approx m_x$, potom:

$$q_x = \frac{2 \cdot u_x}{(2 + u_x)} = \frac{u_x}{(1 + 0,5 \cdot u_x)}.$$

Potom platí:

$q_x = 1 - \exp(-u_x)$, resp. ${}_nq_x = 1 - \exp(-n \cdot u_{x,x+n-1})$ pro zkrácenou tabulku, kde n je šířka intervalu.

2.4 Metody vyrovnání úmrtnostních tabulek

Pojem vyrovnání označuje techniky vyrovnávání neboli vyhlazování úmrtnostních tabulek. Občas může dojít k chybám při výběru dat, což vede k tomu, že odhadované hodnoty se silně liší od hodnot skutečných. Pomocí tohoto procesu se eliminují právě nesystematické nepravidelnosti a kolísání kolem reálných hodnot z odhadovaných pravděpodobnostních úmrtí. Vyrovnání spočívá v nalezení přijatelně křivky pravděpodobnosti úmrtí q_x , přičemž se využívají všechny statistické informace. Proces vyrovnání lze jednoduše popsat jako postup zmenšující náhodné chyby odhadnutých měr úmrtnosti.

V současné době je tato problematika stále více aktuální z důvodu prodlužování naděje dožití, resp. z důvodu toho, že se vysokého věku dožívá stále více osob. Proto se klade velký důraz na modelování úmrtnosti ve vyšších věcích (Pastor, 2007). Avšak vyrovnání může způsobit i

⁹ je vymezen jedním rokem stáří x , jedním kalendářním rokem t a dvěma generacemi $z, z+1$.

zkreslení skutečných odchylek v pozorovaných datech. Proto je úkolem analytika rozhodnout, zda by prioritou mělo být co nejsilnější zachování skutečných dat nebo reprezentace teoretického "správného" vzorce úmrtnosti (Siegel, 2004).

Následující popis a komentáře k jednotlivým metodám vyrovnání ÚT se týkají především nejpoužívanějších modelů. V praxi však může dojít k případům, kdy statistický úřad používá svou vlastní metodu, která je často založena na jednom z obecných modelů. To se týká například metody Martenelle, kterou používá statistický úřad Švédska a která je založena např. na metodě Gompertz-Makehama. Podobné zvláštní modely budou podrobněji rozebrány v příslušných podkapitolách části práce.

Základní model, který popisuje závislost úmrtnosti na věku, byl formulován B. Gompertzem v roce 1825 (Pavlík, 1986). Podle něj je síla úmrtnosti závislá na dvou faktorech, z nichž jeden nezávisí na věku a organismu, zatímco druhý je přímou funkcí věku. Tento model, jemuž se říká Gompertzův zákon, je omezen věkovým intervalem 10 – 60 let, neboť předpokládá, že v tomto období roste síla úmrtnosti geometrickou progresí v závislosti na věku.

Gompertzův model lze vyjádřit jako:

$$\mu_x = a * e^{b*x} \text{ neboli ve tvaru } \mu_x = B * C^x,$$

kde μ_x je síla úmrtnosti a a a b (resp. B a C) jsou parametry modelu.

V roce 1860 William Makeham doplnil tento vzorec o konstantu nezávislou na věku (např. vnější faktory jako války na epidemie). V takovém tvaru lze rovnici zapsat jako:

$$\mu_x = c + a * e^{b*x} \text{ neboli ve tvaru } \mu_x = A + B * C^x,$$

kde μ_x je síla úmrtnosti a a , b a c (resp. A , B a C) parametry modelu.

Přestože tento zákon má značný praktický význam do věku přibližně 80 let, ve vyšších věcích již nelze relativní přírůstky úmrtnosti s věkem považovat za konstantní.

Proto použití Gompertz-Makehamovy funkce může značným způsobem nadhodnocovat skutečnou intenzitu úmrtnosti.

Mezi další modely patří například:

Modifikovaný Gompertz-Makeham

$$\mu_x = a + b * c^{x_0 + \frac{1}{\gamma} * \ln[\gamma * (x - x_0) + 1]},$$

kde $x > x_0$, x_0 je věk, od kterého se provádí vyrovnání pomocí modifikovaného Gompertz-Makehamova modelu, a , b , c jsou parametry modelu a γ je parametr, který vyjadřuje pokles rychlosti nárůstu úmrtnosti s věkem. Navržený model vykazoval dobré výsledky pro data konce 19. a začátku 20. století. Avšak časem se postupně měnil charakter mortality, resp. zvýšila se úmrtnost v mladším a středním věku, což nebylo úplně vhodným prostředím pro používání Makehamova vzorce. Důsledkem těchto změn bylo zavedení nového modelu Heligman-Pollarda (Debon, 2005).

Heligman – Pollard

Původní vzorec představuje osmi-parametrický model. Rovnice obsahuje tři části, z nichž každá označuje výraznou složku mortality. První část odráží pokles úmrtnosti během dětství. Druhá část odráží nárůst úmrtností, který existuje ve věku 10 až 40 let. Tento nárůst zahrnuje zvýšenou úmrtnost na vnější příčiny u mužů, zvýšenou úmrtnost při nehodách obecně a mateřskou úmrtnost u žen. Třetí část rovnice odráží exponenciální model úmrtnosti dospělých ve starším věku (Yuen, 1997).

$$\frac{q_x}{(1-q_x)} = A^{(1+B)^C} + D * \exp\{-E(\ln x - \ln F^2)\} + GH^x,$$

kde A, B, C, D, E, F, G, H jsou parametry modelu, x je věk a q_x je pravděpodobnost úmrtí.

Zaměříme-li jen na nejvyšší věky, je možné využít jen jednu ze tří základních složek tohoto původního modelu, resp. jen část, která byla používána právě za účelem vyrovnání úmrtnosti v nejvyšších věcích. Po takové transformaci se počet parametrů v rovnici redukuje na dva, přičemž ty musí být odhadnuty z empirických dat (Burcin, Tesárková, Šídlo, 2010). Upravený model lze vyjádřit jako

$$q_x = \frac{b * e^{a * x}}{1 + b * e^{a * x}},$$

kde a , b jsou parametry modelu, x je věk a q_x je pravděpodobnost úmrtí.

K odhadu parametrů lze použít například metodu nejmenších čtverců. Nutné je taky zmínit podobnost vyjádření s následujícím modelem.

Himes – Preston – Condran

Návrhem tohoto modelu v roce 1994 bylo použití úmrtnostního standardu (standard mortality schedule, HPC), který reprezentoval typický model úmrtnosti ve starším věku na základě modelů pozorovaných v různých zemích a obdobích. Standard HPC byl konstruován pomocí prozkoumání měř úmrtnosti podle jednotlivých věků pro věkový interval 45-99 let pro 16 zemí s nízkou úmrtností¹⁰. Data pokrývala období 1948-1985. Výsledkem bylo zvolení standardu odvozeného z 82 různých plánů úmrtnosti pro obě pohlaví (Buettner, 2002).

Druhou komponentou je rozšíření tohoto standardního úmrtnostního profilu i pro věky od 95 do 115 let za předpokladu lineárního růstu logitově transformovaných věkově specifických měř úmrtnosti (Buettner, 2002).

$$\text{logit}(m_x) = a + \beta * x,$$

po úpravě dostáváme tvar:

$$m_x = \frac{b * e^{a * x}}{1 + b * e^{a * x}},$$

kde a , b jsou odhady parametrů modelu, x je věk a m_x je specifická míra úmrtnosti.

¹⁰ Mezi těchto 16 zemí patří: Austrálie, Rakousko, Belgie, Kanada, Dánsko, Anglie a Wales, Finsko, Maďarsko, Itálie, Japonsko, Nizozemsko, Nový Zéland, Norsko, Skotsko, Španělsko a Švédsko. Československo, Irsko a Severní Irsko byli vyloučeni z důvodu nedostatečné kvality dat. Data za Francii, Východní a Západní Německo a USA nebyla zahrnuta kvůli jejich rozporuplnosti.

Jak bylo zmíněno u předchozího modelu, lze vidět jeho podobnost vyjádření s modelem Himese, Prestona a Condrana. Prakticky jediný rozdíl najdeme na levé straně rovnice, zatímco v modelu Himese, Prestona a Condrana se vyrovnávání uvedenou funkcí vztahuje na míry úmrtnosti, u Heligmana a Pollarda jde o pravděpodobnosti úmrtí (Burcin, Tesárková, Šídlo, 2010).

Thatcher

Předpokladem tohoto modelu je logistický průběh křivky úmrtnosti. Níže uvedený vzorec představuje zjednodušený případ, který lze také zapsat v tříparametrovém tvaru, avšak výsledek obou vzorců je naprosto identický (Thatcher, 1999).

$$\mu_x = \frac{z}{1+z} + \gamma,$$

kde $z = \alpha * e^{\beta * x}$, α , β a γ jsou parametry modelu.

Kannistö

Model navržený v roce 1992 popisuje zvláštní případ logistické funkce, kde se logitová transformace vyjadřuje jako lineární funkce věku. Přestavený vzorec je dvouparametrovou variantou tohoto modelu (Coelho, Magalhaes, Bravo, 2007):

$$\mu_x = \frac{e [\theta_0 + \theta_1 * (x - 80)]}{1 + e^{[\theta_0 + \theta_1 * (x - 80)]}},$$

kde θ_0 , θ_1 jsou parametry modelu, které nabývají nezáporných hodnot, μ_x je intenzita úmrtnosti ve věku x .

Beard

Beardův tříparametrový logistický model ve své podstatě vychází ze systému logistických křivek, který navrhl Perkse v roce 1932 (Beard, 1959).

$$m_x = \frac{b * e^{a * x}}{1 + c * e^{a * x}},$$

kde a , b , c jsou parametry modelu a m_x je míra úmrtnosti ve věku x .

Kubický model

Tento model představuje rozšíření Gompertzova zákona, přičemž pro odhad parametrů je pak výhodnější jeho logaritmická transformace (Burcin, Tesárková, Šídlo, 2010). Transformované vyjádření tohoto modelu lze popsat jako:

$$\ln \mu_x = \ln B + x * \ln C + x^2 * \ln D + x^3 * \ln E, \text{ případně ve tvaru } \mu_x = B * C^x * D^{x^2} * E^{x^3},$$

kde x je věk a B , C , D , E jsou parametry modelu. K odhadu lze použít metodu nejmenších čtverců, přičemž vyrovnané hodnoty je třeba na závěr výpočtu odlogaritmovat (Burcin, Tesárková, Šídlo, 2010).

Coale – Kisker

Coale a Guo (1989) použili novou metodu vyrovnání tabulek života, která předpokládá, že exponenciální rychlost zvýšení mortality u velmi starších věků není konstantní jako v klasickém

modelu Gompertze, ale klesá lineárně. Podle obvyklé praxe metoda nejprve použitá Coalem a Guo a poté Coalem a Kiskeren byla označena jako metoda Coale-Kisker. (Buetter, 2002, str. 6). Proměnná k_x , která je definována jako logaritmus podílu dvou měr úmrtnosti pro dva po sobě jdoucí věky, byla vymezena na základě změn hodnot měr úmrtnosti podle věku (Boleslawski, Tabeau, 2001).

$$k_x = \ln \frac{m_x}{m_{x-1}},$$

kde m_x je míra úmrtnosti ve věku x . Průběh této veličiny od věku 85 let předpokládali jako lineární:

$$k_x = k_{85} - (x - 85) * s,$$

kde x je věk, k_{85} a s jsou parametry modelu.

Pro tento model platí dva předpoklady. Prvním je, že míry úmrtnosti okolo věku 85 let musí být ještě spolehlivé, a tím pádem může být parametr k_{85} určen přímo z empirických dat. Druhý předpoklad se váže k hodnotě míry úmrtnosti ve věku, který je předpokládán jako nejvyšší dosažitelný. Coale a Kisker ho stanovili na úrovni 110 let. Určení míry úmrtnosti pro tento věk pak umožní odhad parametru s . (Burcin, Tesárková, Šídlo, 2010). Hodnotu míry úmrtnosti ve věku 110 let pak striktně určili na 1,0 pro muže a 0,8 pro ženy. Obě hodnoty vychází z průběhu měr úmrtnosti ve Švédsku (Boleslawski, Tabeau, 2001).

Vypočtený model pak lze vyjádřit jako:

$$m_x = e^{a*x^2 + b*x + c},$$

kde a , b a c jsou parametry modelu.

Denuit – Goderniaux

Tato další vybraná metoda se používá přímo na hrubé pravděpodobnosti úmrtí a stanoví omezující věk pro úmrtnostní tabulku (Bravo, Malta, 2010). Obecně lze model vyjádřit ve tvaru:

$$\ln \hat{q}_x = a + b * x + c * x^2 + \varepsilon_x,$$

kde \hat{q}_x je odhad q_x , a , b , c jsou parametry modelu a ε_x je náhodná složka. Pro náhodnou složku by měly být splněny klasické předpoklady kladené na náhodné složky ve stochastických modelech, kde jejich hodnoty by měly odpovídat normálnímu rozdělení s nulovou střední hodnotou a konstantním rozptylem (Burcin, Tesárková, Šídlo, 2010).

Polynomické funkce

Polynomické funkce jsou zvláštní skupinou metod, které lze využít k odhadování vyrovnaných pravděpodobností úmrtí. Tyto funkce nejsou zákony formulované na základě nějakých vstupních předpokladů o možném vývoji úmrtnosti ve vyšších věcích, ale jsou to základní matematické funkce (Burcin, Tesárková, Šídlo, 2010.). Nejčastěji se využívají polynomické funkce druhého a třetího stupně, tedy:

$$m_x = a + b * x + c * x^2,$$

$$m_x = a + b * x + c * x^2 + d * x^3,$$

kde x je věk a a, b, c, d jsou parametry modelu.

Weibull

Toto rozdělení je známé především z techniky, kde slouží k odhadu životaschopnosti strojů a zařízení, bývalo používáno ale i pro modelování úmrtnosti (Boleslawski, Tabeau, 2001). Hlavní výhodou Weibullové metody je schopnost poskytovat přiměřeně přesné odhady u velmi malých vzorků (Abernethy, 1998).

Model lze vyjádřit jako:

$$m_x = b * c^a,$$

kde x je věk a a, b jsou parametry modelu.

Tab. 1 - Přehled nepoužívanějších modelů vyhlazování a extrapolace křivky úmrtnosti

Název modelu	Rok navržení	Typ	Počet parametrů	Rovnice
Gompertz-Makeham	1860	exponenciální	3	$\mu_x = c + a * e^{b*x}$
Modifikovaný Gompertz-Makeham	1980	exponenciální	4	$\mu_x = a + b * c^{x_0 + \frac{1}{\gamma} \ln[\gamma * (x - x_0) + 1]}$
Himes-Preston-Condarn	1994	logistický	2	$m_x = \frac{b * e^{a*x}}{1 + b * e^{a*x}}$
Heligman-Pollard	1980	logistický	2	$q_x = \frac{b * e^{a*x}}{1 + b * e^{a*x}}$
Thatcher	1999	logistický	3	$\mu_x = \frac{z}{1 + z} + \gamma$
Kannisto	1992	logistický	2	$\mu_x = \frac{e [\theta_0 + \theta_1 * (x - 80)]}{1 + e^{[\theta_0 + \theta_1 * (x - 80)]}}$
Beard	1959	logistický	3	$m_x = \frac{b * e^{a*x}}{1 + c * e^{a*x}}$
Kubický	2007	kubický	4	$\mu_x = B * C^x * D^{x^2} * E^{x^3}$
Coale-Kisker	1990	lineární	3	$k_x = \ln \frac{m_x}{m_{x-1}}$
Denuit-Goderniaux	2005	logistický	3	$\ln \hat{q}_x = a + b * x + c * x^2 + \varepsilon_x$
Polynomická funkce 2. stupně	...	matematický	3	$m_x = a + b * x + c * x^2 + d * x^3$
Polynomická funkce 3. stupně	...	matematický	4	$m_x = a + b * x + c * x^2 + d * x^3$
Weibull	1951	mocninový	2	$m_x = b * c^a$

Zdroj: Burcin, Tesárková, Šídlo, 2010

Kapitola 3

Popis oficiálních úmrtnostních tabulek vybraných zemí

3.1 Přehled vybraných zemí

Celkově bylo analyzováno 12 zemí: Belgie, Česká republika, Estonsko, Grónsko, Irsko, Kanada, Nizozemsko, Portugalsko, Singapur, Španělsko, Švédsko a Velká Británie. Kromě Kanady a Singapuru jsou vybrané země evropské. Je třeba zmínit, že ačkoli Irsko, Estonsko a Grónsko poskytli odpovědi na všechny zkoumané otázky v poslaném dotazníku, obsah těchto odpovědí, resp. chybějící podrobné vysvětlení metodik, bohužel nestačí, aby tyto země byly zaříděny do jednotlivých podkapitol. Proto informace týkající se těchto 3 zemí jsou zahrnuté do společné tabulky na konci této kapitoly.

3.1.1 Belgie

Belgický statický úřad publikuje úmrtnostní tabulky úplné a zkrácené. U zkrácených tabulek jsou také počítány jednoleté věkové intervaly pro věky 0 až 5. Doba mezi sběrem dat a publikováním aktuálních tabulek činí přibližně půl roku. Nejstarším věkovým intervalem pro úmrtnostní tabulky v dokončeném věku je 104+ let, zatímco nejstarší věková hranice pro ÚT v přesném věku je 105 let.

Výpočet pravděpodobnosti úmrtí q_x pro ÚT v dokončených věcích je založen na II. hlavních souborech demografických událostí, pro ÚT v přesných věcích jsou pravděpodobnosti úmrtí vypočteny nepřímou metodou, tj. odvozeny ze specifických měr (m_x) s použitím exponenciální transformace pravděpodobností.

Pravděpodobnost úmrtí pro dokončený věk se vypočítá jako:

$$q_x = \frac{d_x}{p_x}, \text{ nebo } q_x = \frac{d_x}{d_{x+1}}, \text{ pokud } p_x < d_x,$$

kde

d_x – počet zemřelých v dokončeném věku x

p_x – populace v dokončeném věku x .

Pravděpodobnost úmrtí pro přesný věk:

$$q_x = 1 - e^{-m_x},$$

$$q_0 = 1 - \left(1 - \frac{D_0^{z-1}}{{}_{1.1}P_0}\right) * \left(1 - \frac{D_0^z}{N}\right), \quad (11)$$

kde

D_0^{z-1} – zemřelí ve věku 0 generace z-1,

D_0^z – zemřelí ve věku 0 generace z,

P_0 – populace ve věku 0 k 1. 1. ve vybraném roce,

N – počty narozených ve vybraném roce.

Belgický statistický úřad nepoužívá žádné metody vyrovnávání a extrapolace křivky úmrtnosti ve starších věcích, používá se jenom skutečné pozorování.

3.1.2 Česká republika

Podrobné úmrtnostní tabulky za Českou republiku a její kraje jsou založeny na III. hlavních souborech demografických událostí. Vstupní pravděpodobnosti úmrtí jsou vypočteny nepřímou metodou, tj. odvozeny ze specifických měr (m_x). Tabulky jsou podrobné, tj. s jednoletým věkovým intervalem a spočítají se odděleně pro obě pohlaví. Krajské tabulky jsou zpracovány za dvouleté kalendářní období z důvodu vyloučení nahodilých výkyvů.

Vstupními daty pro výpočet pravděpodobnosti úmrtí pro podrobnou úmrtnostní tabulku jsou:

- počet zemřelých dle pohlaví a jednotek věku v daném roce (v rocích za jednotlivé kraje),
- počet zemřelých ve věku 0 v ČR podle ročníku narození,
- počet obyvatel dle pohlaví a jednotek věku v ČR k 1. 7. daného roku (resp. k 1. 1. daného roku v jednotlivých krajích) a
- počet živě narozených dle pohlaví v daném roce (rocích za jednotlivé kraje).

Tabulky se konstruují dle následujícího postupu:

Z empirických dat se vypočítá pro věk $x \geq 1$ specifická míra úmrtnosti (m_x), tj. podíl zemřelých daného věku a pohlaví a stavu obyvatel daného věku k 1. 7. a pohlaví v dané územní jednotce:

$${}_t m_x^{\text{ČR}} = \frac{{}_t D_x}{{}_{1.7.t} P_x},$$

$${}_{t-1,t} m_x^{\text{kraj}} = \frac{{}_{t-1} D_x + {}_t D_x}{{}_{1.1.t} P_x} / 2.$$

Dalším krokem je výpočet pravděpodobnosti úmrtí, která je založena na vzorcí (2). Pravděpodobnost úmrtí ve věku 0 je rovna tzv. kojenecké úmrtnosti neboli podílu zemřelých ve věku 0 a živě narozených v daném období.

$${}_t q_0^{\text{ČR}} = \frac{{}_t D_0}{{}_t N^v},$$

$${}_{t-1,t}q_0^{kraj} = \frac{{}_{t-1}D_0 + {}_tD_0}{{}_{t-1}N^v + {}_tN^v}.$$

Pro odstranění náhodných výkyvů jsou hodnoty pravděpodobnosti úmrtí od věku 4 let je potřeba vyrovnání dle formule:

$$q_x^{vyrovn} = [105 * q_x + 90 * (q_{x-1} + q_{x+1}) + 45 * (q_{x-2} + q_{x+2}) - 30 * (q_{x-3} + q_{x+3})]315].$$

Vzhledem k malé velikosti souboru zemřelých ve vyšším věku pravděpodobnost úmrtí přibližně od věku 80 let odvozuje (extrapoluje) pomocí výše uvedeného Gompertz-Makehamova vzorce. Vstupní charakteristikou pro Gompertz-Makehamovu formuli je přirozený logaritmus pravděpodobnosti dožití odvozené z vyrovnané hodnoty pravděpodobnosti úmrtí:

$$\ln p_x^{vyrovn} = \ln(1 - q_x^{vyrovn}).$$

ČSÚ používá King-Hardyho metodu, při které extrapolace vychází ze soustavy rovnic (R_1, R_2, R_3) pro tři stejně dlouhé po sobě jdoucí intervaly, kde d je délka intervalu a x_0 věk na počátku prvního intervalu (zde $x_0 = 60$ a $d = 8$).

$$R_1 = \sum_{i=x_0}^{x_0+d-1} \ln p_i^{vyrovn}$$

$$R_2 = \sum_{i=x_0+d}^{x_0+2d-1} \ln p_i^{vyrovn}$$

$$R_3 = \sum_{i=x_0+2d}^{x_0+3d-1} \ln p_i^{vyrovn}.$$

Konstanty a, b, c obsažené v Gompertz-Makehamově formuli popisují sílu úmrtnosti a vypočítají se podle vztahů:

$$c^d = \frac{R_3 + R_2}{R_2 + R_1}$$

$$c = \sqrt[d]{c^d}$$

$$b = \frac{(c - 1) * (R_2 + R_1)}{c^{x_0} * (c^d - 1)^2}$$

$$a = \frac{\left[R_1 = \frac{(R_2 - R_1)}{(c^d - 1)} \right]}{d}.$$

Pomocí dosazení a, b, c do Gompertz-Makehamova vzorce lze vypočítat pro věk $x \geq 71$ modelové pravděpodobnosti dožití $r_x = \exp(a + b * c_x)$ a najít věk y ($y \geq 75$), pro který nabývá odchylka $|p_x^{vyrovn} - r_x|$ minimální hodnoty. Od věku y pak pravděpodobnost úmrtí nejlépe vystihuje funkce q_x^{GM} , která je doplňkem funkce r_x do jedné. Přechod na extrapolované hodnoty je upraven vyrovnáním hodnot pro věk $z = (y - 4), \dots, (y + 4)$:

$$q_z^{GM} = 1 - \left[\left(1 - \frac{z - y + 5}{10} \right) * p_z^{\text{vyrovn}} + \frac{z - y + 5}{10} * r_z \right].$$

Pro výpočet úmrtnostní tabulky jsou tedy vstupními hodnotami následující pravděpodobnosti úmrtí:

- pro věk 0 - kojenecká úmrtnost
- pro věk 1, 2, 3 - pravděpodobnosti odvozené ze specifických měř úmrtnosti
- pro věk 4 až $(y - 5)$ - pravděpodobnosti vyrovnané
- pro věk $(y - 4)$ až $\omega - 1$ - pravděpodobnosti vyrovnané a extrapolované

3.1.3 Kanada

Statistický úřad Kanady publikuje úplné a zkrácené (podle pětiletých věkových skupin) úmrtnostní tabulky pro Kanadu, provincie a teritoria. Tabulky jsou publikovány od roku 1939 v nepravidelných intervalech, nejčastěji v souladu s roky konání na sčítání lidu (Statistics Kanada, 2012).

Pro období 2005 až 2007 byly zavedeny dvě zásadní změny v metodice. První změna se týká způsobu odhadu úmrtnosti v starším věku. Kvadratický model Coale-Kiskera (1990), který byl poprvé použit při konstruování tabulek pro období let 1995 až 1997, byl nahrazen logistickým modelem založeným na práci Kannista (1992). Tato změna byla provedena z toho důvodu, že ze studií zemí s vysokou kvalitou dat se ukázalo, že nárůst úmrtnostních poměrů mezi věky má tendenci sledovat logistickou křivkou.

Druhá změna se týkala způsobu vyrovnání věkově specifických pravděpodobností úmrtí. Předchozí metoda používaná od roku 1970, která se spoléhala mimo jiné na klíčové věky, byla nahrazena metodou založenou na spline křivkách. Tato metoda je mnohem flexibilnější a nespolečá se na libovolné klíčové věky (Statistics Kanada, 2012).

Statistický úřad Kanady také provedl další změny týkající se zveřejňování úmrtnostních tabulek. Úmrtnostní tabulky se staly ročním produktem, aby lépe vyhovovaly potřebám mnoha klientů. Tyto tabulky zahrnují tři referenční roky, počínají obdobím 2005 až 2007, pak 2006 až 2008, a tak dále.

Vstupními daty pro sestavování tabulek života jsou:

- Pozorovaný počet úmrtí podle věku (5 až 109 let a pro otevřenou věkovou skupinu 110 let a více) pro kalendářní roky (1. ledna až 31. prosince) $a-1$, a a $a+1$
- Pozorovaný počet úmrtí podle věku (0 až 4 let) a rok narození pro kalendářní roky $a-1$, a a $a+1$
- Odhad populace k 1. červenci podle věku (5 až 110 let) pro kalendářní roky $a-1$, a a $a+1$
- Odhady populace k 1. lednu podle věku (0 až 4 let) pro kalendářní roky $a-1$, a , $a+1$ a $a+2$

Při konstruování úplných ÚT se nejdříve vypočítají pozorované míry úmrtnosti pro věk 5 až 109 let a pro otevřenou věkovou skupinu 110+ let dle vzorce:

$${}^nM_x = \frac{\sum_{t=a-1}^{a+1} D_{x,t}}{\sum_{t=a-1}^{a+1} P_{x,t}}, \quad (11)$$

kde

${}_nM_x$ – pozorované míry úmrtnosti mezi věky x a $x+n$ (v případě úplné ÚT $n=1$)

$\sum_{t=a-1}^{a+1} D_{x,t}$ – celkový počet zemřelých mezi věky x a $x+n$ pro kalendářní roky $a-1$, a a $a+1$

$\sum_{t=a-1}^{a+1} P_{x,t}$ – celková populace k 1. 7. mezi věky x a $x+n$ pro kalendářní roky $a-1$, a a $a+1$.

Dalším krokem je modelování pozorovaných měr úmrtnosti pro věkovou kategorii 95 až 109 let a pro otevřenou věkovou skupinu 110 let a více pomocí formule pro model Kannisto z podkapitoly 2.5. Aby bylo zajištěno, že logistický model správně upravuje pozorovaná data pro starší věk, minimálně 15 pozorovaných měr úmrtnosti ve věku 80 až 110+ let, které byly spočítány v kroku 1, musí být vyčíslitelné (ale nepřičítají se) pro odhad modelu. Pokud by tento práh nebyl dosažen, model nemůže být odhadnout a místo podrobné ÚT se vypočítá zkrácená (Statistics Kanada, 2012).

Následující výpočet pravděpodobností úmrtí pro věk 5 až 110+ let se provádí při použití vzorce (2), resp. nepřímé metody. Po věku 110 a více let má pravděpodobnost úmrtí hodnotu 1. Mezi věky 0 až 4 let jsou pravděpodobnosti úmrtí odhadovány přímo.

Pro věk 0:

$${}_1q_0 = 1 - \left(\frac{P_{0,a+1}^*}{E_{0,a}} * \frac{E_{1,a}}{P_{0,a}^*} \right), \quad (12)$$

kde hodnota $E_{0,a}$ je součet populace ve věku 0 k 1. 1. roku $a+1$ a počtu dětí zemřelých ve věku 0 v průběhu roku a . Hodnota $E_{1,a}$ vyjadřuje rozdíl mezi počtem obyvatel ve věku 0 k 1. 1. roku a a počtu zemřelých ve věku 0 v průběhu roku a

Pro tříletý časový interval, který používá statistický úřad Kanady při konstruování ÚT, je však potřeba výše uvedený vzorec upravit:

$${}_1q_0 = 1 - \left(\frac{\sum_{t=a-1}^{a+1} P_{0,t+1}}{\left(\sum_{t=a-1}^{a+1} P_{0,t+1} + \sum_{t=a-1}^{a+1} D'_{0,t} \right)} * \frac{\left(\sum_{t=a-1}^{a+1} P_{0,t} - \sum_{t=a-1}^{a+1} D''_{0,t} \right)}{\sum_{t=a-1}^{a+1} P_{0,t}} \right),$$

Kde:

${}_1q_0$ – pravděpodobnost úmrtí ve věku 0,

$\sum_{t=a-1}^{a+1} P_{0,t+1}$ – součet počtů obyvatel ve věku 0 k 1. lednu a , $a+1$ a $a+2$ let, což jsou poslední dva roky sledovaného období a rok, který následuje,

$\sum_{t=a-1}^{a+1} D'_{0,t}$ – součet zemřelých ve věku 0 let během referenčního období pro děti zemřelé ve stejném roce jako byl rok jejich narození,

$\sum_{t=a-1}^{a+1} P_{0,t}$ – součet počtů obyvatel ve věku 0 k 1. lednu v průběhu referenčního období,

$\sum_{t=a-1}^{a+1} D''_{0,t}$ – součet zemřelých ve věku 0 let během referenčního období pro děti narozené v průběhu roku, který předchází roku jejich úmrtí.

Pro věk 1 až 4:

$${}_1q_0 = 1 - \left(\frac{\sum_{t=a-1}^{a+1} P_{x,t+1}}{\left(\sum_{t=a-1}^{a+1} P_{x,t+1} + \sum_{t=a-1}^{a+1} D'_{x,t} \right)} * \frac{\left(\sum_{t=a-1}^{a+1} P_{x,t} - \sum_{t=a-1}^{a+1} D''_{x,t} \right)}{\sum_{t=a-1}^{a+1} P_{x,t}} \right)$$

V případě, že pro věk 1 až 4 let nebyla pozorována žádná úmrtí v průběhu referenčního období, pravděpodobnosti úmrtí se interpolují stejně jako v následujícím kroku.

Aby bylo zajištěno, že se pravděpodobnosti úmrtí vyvíjí důsledně z jednoho věku do druhého, a bylo možno v případě potřeby odhadnout chybějící pravděpodobnosti ve věku 1 až 4 let, používá se metoda vyrovnání pomocí B-spline. Výhodou této vyrovnávací techniky je její flexibilita, což znamená poskytnutí uživatelům mnoho možností tak, aby data vyhovovala stanoveným požadavkům nejlepším možným způsobem. Při využití výše uvedené metody je na časové ose konstruováno několik pozic neboli uzlů, kde dochází ke spojení. Čím větší je počet uzlů, tím více odpovídají vyrovnané křivky původní křivce pravděpodobností úmrtí podle věku. Naopak malý počet uzlů vede k silnějšímu vyrovnání. Výsledkem je, že náhodné výkyvy jsou eliminovány ve prospěch křivky s pravidelnějším tvarem. B-spline vyrovnání pravděpodobností úmrtí ve věku 1 až 94 let v aktuálních tabulkách života Kanady bylo vytvořené procedurou TRANSREG v statistickém softwaru SAS (Statistics Kanada, 2012).

Výpočet tabulkových funkcí probíhá stejným způsobem jako bylo uvedeno v kapitole 2.2, avšak s několika výjimkami:

Počet let prožitých populací mezi věky x a $x+n$ se počítá dle vzorce:

$${}_nL_x = n * l_{x+n} + ({}_n d_x * {}_n f_x),$$

kde pro podrobné tabulky

${}_1 f_x$ – separační faktor,

$${}_1 f_x = \left\{ \begin{array}{l} 1 - \left(\frac{\sum_{t=x}^{x+a-1} D'_{x,t}}{\sum_{t=x}^{x+a-1} (D'_{x,t} + D''_{x,t})} \right) \text{ pro } x = 0 \text{ až } 4 \text{ let, pokud čítec a jmenovatel jsou větší, než } 0 \\ a \\ 0.5 \text{ pro } x \geq 5 \text{ let} \end{array} \right\}.$$

Pokud je buď čítec nebo jmenovatel roven 0, při výpočtu ${}_1 f_x$ pro věk 0 až 4 let je hodnota ${}_1 f_x$ vypočtena na základě přístupu, který byl použit ve vzorci (11).

$$L_{110+} = l_{110} * e_{110} \text{ pro otevřenou věkovou skupinu } 110+ \text{ let, kde } e_{110} = \frac{1}{M_{110}}$$

Zkrácené ÚT za Kanadu jsou konstruovány pro menší územní celky, kde použití podrobných tabulek není optimální z důvodu chybějících úmrtnostních údajů zejména pro mladší věky. V některých případech jako při výpočtu pravděpodobnosti úmrtí ve věku 0 jsou metody zkrácené a podrobné tabulky totožné. Nicméně není potřeba použít model pro odhad úmrtnosti ve starším věku, protože zkrácená úmrtnostní tabulka končí otevřeným věkovým intervalem 90+ let. Také se nepoužívá žádná metoda vyrovnání pravděpodobnosti úmrtí, protože náhodné výkyvy jsou minimální (Statistics Kanada, 2012).

3.1.4 Portugalsko

Statistický úřad Portugalska vydává dva typy úmrtnostních tabulek: národní úmrtnostní tabulky a krajské úmrtnostní tabulky. Jedná se o kompletní úmrtnostní tabulky s referenčním obdobím tří po sobě jdoucích let. Metodika pro oba druhy je stejná, ale při konstruování krajských úmrtnostních tabulek se používá metoda vyrovnání. Navíc se pro úmrtnostní tabulky dle

klasifikace NUTS III (z nichž se zveřejňují pouze střední délka života při narození a ve věku 65 let) aplikuje relační model.

Tabulky života jsou zveřejňovány ročně. Úplná úmrtnostní tabulka pro Portugalsko se publikuje na konci května. Úplné úmrtnostní tabulky podle NUTS II a střední délka života pro NUTS III jsou obvykle zveřejněny v listopadu. Nejstarším věkem je 115 let, ale data jsou publikována pouze do věku 100 let (Statistics Portugal, 2012).

Pravděpodobnost úmrtí se odhaduje podle formule:

$$\hat{q}_x = \frac{d_{x,t,g}^L + d_{x,t+1,g}^U + d_{x,t+1,g+1}^L + d_{x,t+2,g+1}^U}{P_{x,t+1} + d_{x,t,g}^L + P_{x,t+2} + d_{x,t+1,g+1}^L},$$

kde

$d_{x,t,g}^L$ – zemřelí v dolním souboru událostí Lexisova diagramu pro věk x , kalendářní rok t a generaci g ,

$d_{x,t,g+1}^U$ – zemřelí v horním souboru událostí Lexisova diagramu pro věk x , kalendářní rok t a generaci $g+1$,

$P_{x,t}$ – počet obyvatel mezi věky x a $x+1$.

Pro věk 0:

$$\hat{q}_0 = \frac{d_{0,t,g}^L + d_{0,t+1,g}^U + d_{0,t+1,g+1}^L + d_{0,t+2,g+1}^U}{NV_t + NV_{t+1}}, \quad (14)$$

kde NV_t – počet živě narozených v roce t .

Pro všechny úmrtnostní tabulky používá portugalský statistický úřad metodu Denuit a Goderniaux (2005) pro uzavření a extrapolaci pravděpodobnosti úmrtí pro starší věkové kategorie. Kromě této metody se pro regionální úmrtnostní tabulky podle NUTS II používá Gompertz-Makehamova metoda. Pro NUTS III se namísto metody Gompertz-Makehama používá Brassův relační model založený na horním regionu NUTS II (Statistics Portugal, 2012).

3.1.5 Singapur

Statistický úřad Singapur sestavuje ročně od roku 2003¹¹ podrobné úmrtnostní tabulky, které jsou odvozené z pravděpodobnosti úmrtí (q_x), která se vypočítá z měr úmrtnosti (m_x) pro každý rok věku x . Při výpočtu měr úmrtnosti se používá stav populace k 1.7. Vzhledem k tomu, že Singapur má jen malý počet obyvatel, je počet úmrtí řídko distribuovaný v jednotlivých letech. Výsledkem jsou míry úmrtnosti pro vybraný věk odvozené přímo z hrubých dat s významnými náhodnými výkyvy. Za účelem snížení vlivu náhodných fluktuací jsou míry úmrtnosti pro věk 0, věkovou skupinu 1 až 4 let a pětileté věkové skupiny 5-9, 10-14 let atd. nejprve odvozeny dělením

¹¹ Před rokem 2003 byly publikovány jen zkrácené tabulky

průměrného tříletého počtu zemřelých ve vybrané věkové skupině¹². Míry úmrtnosti ve věku x jsou pak odvozeny s využitím lineární interpolace (Singapore Department of Statistics, 2008).

Pro zajištění hladké křivky úmrtnosti s rostoucím věkem se odvozené m_x pro každý věk x vyrovnávají pomocí metody Whittaker-Hendersona. Vyrovnané m_x jsou pak vybrané tak, aby minimalizovaly následující rovnici:

$$J = F + hS \equiv \sum_{x=0}^n (m_x^{vyrovn} - m_x)^2 + h \sum_{x=0}^{n-3} (\Delta^3 m_x^{vyrovn})^2,$$

kde

m_x – počáteční odhad úmrtnosti ve věku x ,

m_x^{vyrovn} – vyrovnaná míra úmrtnosti ve věku x ,

h – nezáporný parametr,

Δ – parametr následujícího rozdílu.

Míra úmrtnosti ve věku 0 je výrazně větší než míra úmrtnosti ve věku 1 až 4 let. Chceme-li zajistit, aby vyrovnané míry úmrtnosti odrážely základní míry úmrtnosti ve věku 1 až 4 let, používá se fiktivní m_0 (vypočítá se jako $m_1 * \frac{m_1}{m_2}$), která nahrazuje původní m_0 během procesu vyrovnání.

Pro vyrovnání měr úmrtnosti ve starším (85-99 let) věku se používá metoda Coale-Kiskera. Vyrovnané míry úmrtnosti pro věk 0-99 let jsou pak použity při konstruování podrobných úmrtnostních tabulek (Singapore Department of Statistics, 2008).

Pravděpodobnost úmrtí se odhaduje nepřímou metodou dle vzorce:

pro věk 0 $q_0 = \frac{m_0}{1+(1-f_0)m_0}$, kde f_0 je separační faktor odvozený z historických dat,

pro věk 1 až 99 $q_x = \frac{m_x}{1+0.5m_x}$,

pro věk 100+ let $q_{100} = 1$.

Tabulkový počet dožívajících se počítá dle vzorce:

$$l_x = l_{x-1} - (l_{x-1} * q_{x-1})$$

Ostatní tabulkové funkce jsou vypočítané dle formulí (1) až (10) uvedených v kapitole 2.3.

3.1.6 Slovensko

Metodika, kterou používá statistický úřad Slovenska je velmi podobná metodice pro Českou republiku. Úmrtnostní tabulky počítané pro jednotky věku, tj. s jednoletým věkovým intervalem, se označují jako tabulky podrobné. Vstupem pro jejich výpočet jsou počty zemřelých v daném věku a počet žijících v daném věku neboli střední stav za sledovaný rok (Mészáros, 2000).

¹² Pro sledovaný rok t , kde úmrtnostní data pro rok $t+1$ nejsou k dispozici, jsou předběžné míry úmrtnosti odvozeny na základě průměrného počtu zemřelých v roce t a $t-1$.

Pravděpodobnost úmrtí se vypočítá pomocí nepřímé metody a pro odstranění náhodných výkyvů pro věk 6 až 85 let se používá stejný vzorec jako pro ČR, resp.:

$$q_x^{vyrovn} = [105 * q_x + 90 * (q_{x-1} + q_{x+1}) + 45 * (q_{x-2} + q_{x+2}) - 30 * (q_{x-3} + q_{x+3})] / 315.$$

Tento způsob výpočtu pravděpodobnosti úmrtí je spolehlivý do věku přibližně 80 let. Nad tímto věkem vykazují počty zemřelých velké nepravidelnosti, protože jsou zatíženy velkými náhodnými odchylkami z důvodu jejich malého počtu. Proto se pro vyšší věkové ročníky používá extrapolace hodnot pravděpodobnosti úmrtí pomocí Gompertz-Makehamova vzorce (metoda King-Hardy) (Mészáros, 2000).

Úmrtnostní tabulka se ukončuje věkem 100 a více let, a to tak, že se hodnoty L_x a T_x nahradí hodnotou výrazu $l_{100} - 0,5 * l_{100} * q_{100}$.

Zkrácené úmrtnostní tabulky jsou odvozené z podrobných úmrtnostních tabulek a počítají se pro pětileté věkové skupiny následujícím způsobem:

Nejdříve se vypočítají zbývající roky života ve věkové skupině y dle vzorce:

$$T_y^s = \sum_{j \geq y} \sum_{i \in j} L_i \text{ pro všechny věkové skupiny } y.$$

Následně se odhaduje počet žijících ve věkové skupině y :

$$L_y^s = \frac{\sum_{i \in j} L_i}{n_y} \text{ taky pro všechny věkové skupiny } y.$$

Zvolí se hodnoty $l_0 = 100000$, $d_0 = (l_0 - L_0) / a_0$ a $l_1 = l_0 - d_0$, kde a_0 je zjištěné při

výpočtu podrobné ÚT pro SR.

V dalším kroku se zvolí hodnoty $l_0 = 100000$, $d_0 = (l_0 - L_0) / a_0$ a $l_1 = l_0 - d_0$, kde a_0 je zjištěno při výpočtu podrobné ÚT.

Postupně se vypočítají hodnoty:

$$l_y = L_{y-1}^s * 2 - l_{y-1} \text{ pro } y \geq 2$$

$$d_y = l_y - l_{y+1} \text{ pro } y \geq 1$$

$$q_y = 1 - l_{y+1} / l_y \text{ pro všechny } y$$

$$e_y = T_y^s / l_y \text{ pro všechny } y.$$

Za poslední řádek tabulky se zvolí věková skupina 85 a více let. Tento řádek se upravuje následně:

$$q_{85+} = 1 - \exp\left(\frac{-D_{85+}}{P_{85+}}\right),$$

kde D_{85+} je počet zemřelých ve věku 85 a více let, P_{85+} je počet žijících (střední stav) ve věku 85 a více let.

3.1.7 Španělsko

Úmrtnostní tabulky pro Španělsko znázorňují výskyt úmrtnosti bydlícího obyvatelstva v dané zemi v průběhu referenčního roku. Zároveň simulují úmrtnostní chování buď v rámci kohorty, nebo v rámci fiktivní generace osob (Instituto Nacional de Estadística, 2013). Specifické míry úmrtnosti dle věku m_x se odhadují na základě předpokladu o rovnoměrném rozložení narozených, kteří nezemřeli v průběhu roku, a na rovnoměrném rozložení přistěhovalých a odstěhovalých během sledovaného období dle formule:

$$m_x = \frac{D(t, x, s)}{\frac{(P(t, x, s) - D_2(t, x, s))}{2} + \sum_{i=1}^{D_2(t, x, s)} b_2(t, x, s, i) + \frac{P(t+1, x, s)}{2} + \sum_{i=1}^{D_1(t, x, s)} b_1(t, x, s, i)},$$

$x = 0, 1, \dots, 99$.

kde:

t – sledovaný rok, resp. období,

x – dosažený věk pro $x = 0, 1, \dots, 99$,

s – pohlaví,

$P(t, x, s)$ – bydlící obyvatelstvo k 1. 1. ve věku x , za rok t , pohlaví s ,

$D_1(t, x, s)$ – zemřelí v roce t ve věku x , pohlaví s , kteří dosáhli věku x během roku t ,

$D_2(t, x, s)$ – zemřelí v roce t ve věku x , pohlaví s , kteří dosáhli věku x během roku $t-1$,

$b_1(t, x, s, i)$ – rozdíl (v počtu let) mezi úmrtím a narozením (během roku t) pro každého jedince i , pohlaví s , který zemřel během roku t ve věku x a dosáhl věku x během roku t . Je třeba vzít v úvahu, že toto množství je prožitý čas (počet let) ve věku x každého jedince, který zemřel v tomto věku v roce t a který dosáhnul věku x v tomto roce.

$b_2(t, x, s, i)$ – rozdíl (v počtu let) mezi úmrtím a 1. lednem roku t pro každého jedince i , pohlaví s , který zemřel během roku t ve věku x a dosáhl věku x během roku $t-1$. Je třeba vzít v úvahu, že toto množství je prožitý čas (počet let) ve věku x každého jedince, který zemřel v tomto věku v roce t a který dosáhnul věku x v roce $t-1$.

Pro otevřenou věkovou skupinu 100+ se ukazatel m_x se odhaduje dle formule:

$$m_{100+} = \frac{D(t, 100+, s)}{\frac{(P(t, 100+, s) - D_2(t, 100+, s))}{2} + \sum_{i=1}^{D_2(t, 100+, s)} b_2(t, 100+, s, i) + \frac{P(t+1, 100+, s)}{2} + \sum_{i=1}^{D_1(t, 100+, s)} b_1(t, 100+, s, i)},$$

kde:

t – sledovaný rok, resp. období,

s – pohlaví,

$D(t, 100+, s) = \sum_{x \geq 100} D(t, x, s)$ – součet zemřelých ve věku 100+ let v roce t , pohlaví s ,

$D_1(t, 100+, s)$ – počet zemřelých ve věku 100 let v roce t , kteří dosáhli věku 100 let během roku t ,

$D_2(t, 100+, s)$ – počet zemřelých ve věku 100+ let v roce t , kteří dosáhli věku 100 let během roku $t-1$,

$P(t, 100+, s) = \sum_{x \geq 100} P(t, x, s)$ – bydlící obyvatelstvo k 1. 1. roku t ve věku 100+ let, pohlaví s ,

$b_1(t, 100, s, i)$ – rozdíl (v počtu let) mezi úmrtím a narozením (během roku t) pro každého jedince i , pohlaví s , který zemřel během roku t ve věku 100 let a dosáhl věku 100 let během roku t . Je třeba vzít v úvahu, že toto množství je prožitý čas (počet let) ve věku 100 let pro každého jedince, který zemřel v tomto věku v roce t a který dosáhnul věku x v tomto roce.

$b_2(t, 100+, s, i)$ – rozdíl (v počtu let) mezi úmrtím a 1. lednem roku t pro každého jedince i , pohlaví s , který zemřel během roku t ve věku 100+ a dosáhnul věku 100 během roku $t-1$. Je třeba vzít v úvahu, že toto množství je prožitý čas (počet let) ve věku 100 let pro každého jedince, který zemřel v tomto věku v roce t a který dosáhnul věku 100 let v roce $t-1$.

Porovnáváme-li odhadované hodnoty specifických měr úmrtnosti studovaného obyvatelstva s mírami, které odpovídají specifickým mírám úmrtnosti v každém věku x fiktivní kohorty 100000 jednotlivců, pravděpodobnost neboli riziko úmrtí ve věku x - q_x výše uvedené kohorty jedinců prezentuje stejný výskyt úmrtnosti u každého věku populace v referenčním období a počítá se podle vzorce:

$$q_x = \frac{m_x}{1 + (1 + a_x) * m_x},$$

$$x = 0, 1, \dots, 99,$$

kde

a_x – je průměr let prožitých v posledním roce života těchto zemřelých jedinců fiktivní kohorty, kteří dosáhli věku x .

Výše uvedená funkce a_x je vypočítaná s ohledem na krátkodobé výskyty úmrtnosti v populaci v referenčním roce s použitím průměrné prožité doby jednotlivců ve věku x , kteří zemřeli v uvedeném věku v průběhu tohoto roku (Instituto Nacional de Estadística, 2013):

$$a_x = \frac{\sum_{i=1}^{D(t,x,s)} a(t, x, s, i)}{D(t, x, s)},$$

$$x = 0, 1, \dots, 99,$$

kde

$a(t, x, s, i)$ – je doba, prožita jednotlivcem i pohlaví s ve sledované populaci, který zemřel ve věku x v průběhu referenčního roku t .

Pro poslední otevřenou věkovou skupinu 100+ let platí:

$$q_{100+} = 1,$$

$$a_{100+} = \frac{1}{m_{100+}}.$$

Tabulkové ukazatele l_x a d_x jsou vypočítány dle formulí (1) až (10) uvedených v podkapitole 2.3. Počet „člověkolet“ jednotlivců fiktivní generace pro pohlaví s , kteří se dožili věku x , je odvozen z výrazu:

$$L_x = l_{x+1} + a_x * d_x, \text{ pro } x = 0, 1, \dots, 100.$$

Naděje dožití ve věku x pro pohlaví s se počítá dle vzorce:

$$e_x = \frac{\sum_{y \geq x} L_y}{l_x}, \text{ pro } x = 0, 1, \dots, 99, 100.$$

3.1.8 Švédsko

Švédský statistický úřad ročně publikuje úplné a zkrácené úmrtnostní tabulky s pětiletými věkovými intervaly. Nejstarší věkovou jednotkou je 110 let (Johansson, 2014). Základní ukazatel pravděpodobnosti úmrtí se počítá dle formule:

$$q_x = \frac{\bar{d}_x^t + d_{-x}^t}{\frac{1}{2}(K_x^{t-1} + K_x^t) + \bar{d}_x^t} = \frac{D_x^t}{\frac{1}{2}(K_x^{t-1} + K_x^t) + \bar{d}_x^t},$$

kde

D_x^t – počet jednotlivců, kteří dosáhli věku x a zemřeli ve stejném roce t ,

\bar{d}_x^t – počet jednotlivců, kteří zemřeli ve věku x v roce t před svými narozeninami,

K_x^t – počet jednotlivců na konci roku t ve věku x , nikoli $x+1$.

Pro věk 0 platí:

$$q_0 = \frac{D_0}{N}, \tag{15}$$

kde:

D_0 – počet zemřelých ve věku 0,

N – počet narozených.

Co se týče vyrovnání křivky úmrtnosti, do roku 1986 používal statistický úřad Švédska Wittsteinovou metodu, která nadhodnocovala pravděpodobnost úmrtí v starším věku. Z tohoto důvodu se od roku 1987 používala metoda, kterou vytvořil Sten Martinelle ze švédského statistického úřadu. Nicméně tato metoda, kde Martinellovy modelované pravděpodobnosti úmrtí nahrazovaly pozorované pravděpodobnosti pro věkové kategorie 90 a výše let, silně spoléhala na údaje o úmrtnosti za Švédsko a některé další země a nebrala v úvahu, že pozorované pravděpodobnosti musí být vyrovnané.

Proto se v průběhu času kvalita dat postupně zhoršila, což mělo za následek systematické podhodnocení pravděpodobnosti úmrtí a nadhodnocení naděje dožití pro novorozence e_0 . V roce 2012, kdy se výpočty v úmrtnostních tabulkách přezkoumaly, bylo rozhodnuto, že metoda Martinelle by měla být aktualizovaná, což vedlo k vytvoření nové metody, u které se vyrovnávání pravděpodobnosti úmrtí lépe hodilo na pozorovaná data (Johansson, 2014).

Martinellův model je založen na formulích Gompertz-Makehama a Perkse (1932):

$$\mu_x = A + Be^{kx}, \text{ resp. } \mu_x = \frac{A+B^{kx}}{1+Dx^{kx}},$$

kde

$$A > 0, B > 0, D > 0, k > 0.$$

Martinellova námitka proti použití Perksova vzorce spočívala v tom, že ve vysokém věku se síla úmrtnosti přibližuje konstantní hodnotě B/D . Jeho argumentem bylo, že neexistuje žádný empirický důkaz pro existenci “plošiny“ úmrtnosti u stoletých osob. Proto Martinelle mění předpoklad použitý ve vzorci Perkse a nahrazuje předpoklad gama rozdělení parametru B předpokladem parametru B s generalizovaným gama rozdělením, které bylo posunuté doprava, tedy za předpokladu, že je možné, aby parametr B mohl nabývat hodnot velmi blízko k nule.

V případě, kdy parametr $A = 0$, což dobře odráží situaci ve Švédsku v současné době, Martinellův model tvrdí, že síla úmrtnosti pro jednotlivce s křehkostí¹³ z je podle Gompertz-Makehamova zákona (Johansson, 2014):

$\mu_{(x|z)} = ze^{kx}$ a hodnota křehkosti z má posunuté gama rozdělení s hustotou

$$g(t) = \begin{cases} \frac{b^a(t-c)^{a-1}}{\Gamma(a)} e^{-b(t-c)} & \text{pro } t > c \\ 0 & \text{pro } t \leq c \end{cases},$$

kde $(a, b, c > 0)$, pak průměrná síla úmrtnosti je dána

$$\mu(x) = \left(c + \frac{\eta}{1 + \eta\alpha^2 \int_{x_0}^x e^{kt} dt} \right) e^{kx},$$

kde η a a jsou průměr a relativní směrodatná odchylka proměnných $z-c$.

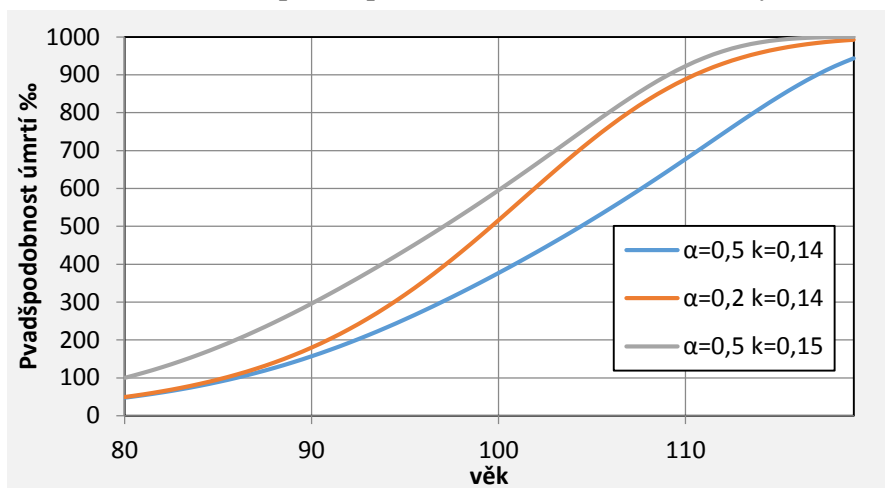
¹³ Konstanta B neboli schopnost jednotlivce odolávat smrti

Na základě rovnice $q_x = 1 - e^{-\mu(x)}$ dostáváme:

$$q_x = 1 - e^{\left(c + \frac{\eta}{1 + \eta a^2 \int_{x_0}^x e^{kt} dt} \right) e^{kt}},$$

kde a , k , c a η jsou parametry modelu.

Obr. 2 – Modelované pravděpodobnosti úmrtí, Švédsko, ženy, 2013¹⁴



Zdroj: Švédský statistický úřad, 2013

U velmi vysokých věků pokles α vede k prudšímu nárůstu pravděpodobnosti úmrtí mezi věky a zvýšení k vede k téměř konstantnímu zvýšení pravděpodobnosti na širokém intervalu vysokých věkových kategorií. Výše uvedený model se používá švédským úřadem od roku 2012 (Johansson, 2014).

3.1.9 Velká Británie

Průřezové tabulky života pro Velkou Británii jsou publikovány ročně. Jednotlivé tabulkové funkce se vypočítají podle vzorců uvedených v kapitole 2.2, avšak je nezbytné podrobně probrat metody odhadu pravděpodobnosti kojeneckého úmrtí a vyrovnání křivky úmrtnosti v nejstarších věcích (Office of National Statistics, 2014).

Úmrtnostní tabulky zahrnují úmrtí kojenců pro věk <4 týdny, 1-2 měsíce, 3-5 měsíců, 6-8 měsíců a 9-11 měsíců, které jsou počítány odděleně pro muže a ženy pro období od roku t do roku $t+2$ včetně.

<4 týdny:

$$\frac{N_{\text{prosinec}(t-1)}}{2} + \sum_{i=\text{leden}(t)}^{i=\text{listopad}(t+2)} N_i + \frac{N_{\text{prosinec}(t+2)}}{2}$$

1-2 měsíce:

¹⁴ Pro $c = 1.363 \cdot 10^{-7}$ a $\eta = 5.535 \cdot 10^{-7}$

$$\frac{(N_{říjen(t-1)} + N_{listopad(t+2)})}{4} + \frac{3 * (N_{listopad(t-1)} + N_{říjen(t+2)})}{4} + \sum_{i=prosinec(t-1)}^{i=září(t+2)} N_i$$

3-5 měsíců:

$$\frac{(N_{červenec(t-1)} + N_{září(t+2)})}{6} + \frac{(N_{srpen(t-1)} + N_{srpen(t+2)})}{2} + \frac{5 * (N_{září(t-1)} + N_{červenec(t+2)})}{6} + \sum_{i=říjen(t-1)}^{i=červenec(t+2)} N_i$$

6-8 měsíců:

$$\frac{(N_{duben(t-1)} + N_{leden(t+2)})}{6} + \frac{(N_{květen(t-1)} + N_{květen(t+2)})}{2} + \frac{5 * (N_{červenec(t-1)} + N_{duben(t+2)})}{6} + \sum_{i=červenec(t-1)}^{i=březen(t+2)} N_i$$

9-11 měsíců:

$$\frac{(N_{leden(t-1)} + N_{březen(t+2)})}{6} + \frac{(N_{únor(t-1)} + N_{únor(t+2)})}{2} + \frac{5 * (N_{březen(t-1)} + N_{leden(t+2)})}{6} + \sum_{i=duben(t-1)}^{i=prosinec(t+1)} N_i$$

kde $N_{x(t)}$ - narození v měsíci x kalendářního roku t .

Každá z celkových skupin zemřelých se dělí na odpovídající populace vybraných věkových skupin, které se vypočítají podle výše uvedených vzorců. Výsledky jsou pak sečteny pro odhad q_0 (Office of National Statistics, 2014).

Ukazatel m_0 se vypočítá z q_0 dle vzorce:

$$m_0 = \frac{2q_0}{2 - q_0}.$$

Pomocí stanovení průměrného věku úmrtí u každého období kojenecké úmrtnosti lze vypočítat a_0 . Tento ukazatel se pak používá při odhadu L_0 podle vzorce:

$$L_0 = a_0 l_0 + (1 - a_0) l_1$$

Tab. 2 - Předpokládané průměrné věky při úmrtí do 1 roku, Velká Británie, 2013, chlapci a dívky

Věk při úmrtí	Předpokládaný průměrný věk při úmrtí (v měsících)
<4 týdny	0.2*
1-2 měsíce	1.5
3-5 měsíců	4
6-8 měsíců	7
9-11 měsíců	10

*na základě dat o úmrtnosti pro věk menší než 1 měsíc pro Anglii a Wales

Zdroj: Office for National Statistics, 2013

Pro vyrovnání křivky úmrtnosti ve věku 90 a více let používá statistický úřad Velké Británie metodu Kannisto-Thatcher. Odhad populace ve vybraném věku se konstruuje pomocí poměrů počtu přeživších v rámci kohorty, kteří jsou stále naživu, ku počtům z této kohorty, kteří zemřeli v posledních několika letech. Předpokladem je, že v nejvyšším věku, kdy každý v dané kohortě zemře, je možné vytvořit algoritmus pomocí těchto poměrů přeživších, které poskytují odhady počtu lidí žijících v mladším věku pro každou kohortu. To znamená, že metoda KT používá tzv. "věku při úmrtí" ("age-at-death") k vybudování distribučních profilů počtu starších lidí v předchozích letech. Například pokud někdo zemře v roce 2006 ve věku 105 let, znamená to, že byl naživu ve věku 104 let v roce 2005 a ve věku 103 let v roce 2004 atd. Shromáždění dat za "věk při úmrtí" pro řadu let umožňuje provést odhad počtu žijících osob daného věku v libovolném vybraném roce a vytvořit tak příslušné profily věkové distribuce za předpokladu, že migrace v nejstarších věcích je minimální. Při vytvoření odhadů pro současný rok není možné použít data o úmrtnosti, protože sledujeme populaci, která v tomto nebo velmi nedávném okamžiku byla naživu (Office of National Statistics, 2014).

Metoda KT také používá průměr věku při úmrtí za posledních pět let pro odhad počtu přeživších v aktuálním roce. Odhady jsou pak omezeny na odhad středního stavu obyvatelstva k 1. červenci z celkového počtu lidí ve věku 90 + v daném roce.

Tuto metodu lze pro všechny roky předcházející vybranému roku popsat následujícím způsobem:

$$P_x^t = P_{x+1}^{t+1} + D_x^t,$$

kde P je populace, D jsou zemřelí ve věku x k 1. lednu, x je věk a t je kalendářní rok.

Pro vybraný rok platí:

$$P_x^T = (D_{x-1}^{T-1} + D_{x-2}^{T-2} + D_{x-3}^{T-3} + D_{x-4}^{T-4} + D_{x-5}^{T-5}) * S_x^T * c,$$

kde c je korekční faktor, S je poměr přeživších a T je vybraný rok.

Poměr přeživších S se odhaduje podle formule:

$$S_x^T = \sum_{T-1}^{T-5} P_x / (\sum_{T-2}^{T-6} D_{x-1} + \sum_{T-3}^{T-7} D_{x-2} + \sum_{T-4}^{T-8} D_{x-3} + \sum_{T-5}^{T-9} D_{x-4} + \sum_{T-6}^{T-10} D_{x-5}).$$

Kromě toho se předpokládá, že nikdo se nedožije věku 120 let. C se vypočítá podle:

$$\sum_{x=90}^{x=120} P_x^T = E,$$

kde E je oficiální odhad populace ve věku 90 a více let k 1.7.

Jedním z výsledků této metody je, že se každým rokem odhady na základě předchozích let stávají více přesnější jakmile jsou k dispozici nová data o úmrtí. Poté, co všichni členové vybrané kohorty zemřou, poskytují jejich zaznamenaná data narození a úmrtí dostatek informací k rekonstrukci obyvatel, kteří byli naživu ke každému datu v minulosti, alespoň pro věky, kde vliv migrace může být ignorován. Na základě této metody, která je známá jako metoda "vymřelých generací" nebo

"zaniklých kohort", mohou být vylepšené odhady z předchozích generací vypočítány retrospektivně (Office of National Statistics, 2014).

3.1.10 Estonsko, Grónsko a Irsko

Vybrané země byly seskupeny do jedné tabulky z důvodu malého rozsahu poskytnutých dat. Zajímavý je případ Grónska, který dobře ilustruje situaci, ve které u země s malým počtem obyvatel se nepoužívá žádná metoda extrapolace křivky úmrtnosti.

Tab. 3 – Charakteristiky úmrtnostních tabulek za Estonsko, Grónsko a Irsko

Název země	Typy ÚT	Periodicita	Nejstarší věk / věkový interval	Metoda výpočtu pravděpodobnosti úmrtí	Metoda vyrovnání křivky úmrtnosti
Estonsko	podrobné / zkrácené	ročně	100 / 85+ let	nepřímá*	Heligman-Pollard**
Grónsko	podrobné***	ročně	99 let	nepřímá	není****
Irsko	podrobné / zkrácené	jednou za 5 let	105 / 85+ let	nepřímá	King's Method of Osculatory Interpolation (1911)

* s přičítáním počtu osob, které zemřeli po svých narozeních v posledním roce života

** v kombinaci s metodou klouzavého průměru

*** tabulky života pokrývají pětiletý časový interval, a to od roku 1977

**** z důvodu 5letých intervalů a skutečnosti, že celková populace zahrnuje 56300 osob a roční počet zemřelých je přibližně 4500, se nepoužívají žádné metody vyrovnání křivek úmrtnosti

Zdroj: elektronická komunikace

Kapitola 4

Analýza a porovnání vybraných metod konstruování úmrtnostních tabulek

Následující kapitola se zabývá srovnáním různých metod, které používají vybrané statistické úřady při konstruování úmrtnostních tabulek. První podkapitola se věnuje rozdílům při výpočtu pravděpodobnosti úmrtí ve věku 0 na základě vzorců uvedených v kapitole 3. Druhá podkapitola analyzuje vybrané metody vyrovnání křivky úmrtnosti a zároveň konstruuje ukazatele střední délky života podle jednotlivých modelů. Co se týče výpočtu ostatních funkcí úmrtnostní tabulky, vybrané statistické úřady používají stejné metody, které již byly uvedeny v podkapitole 2.2.

4.1 Kojenecká úmrtnost

V porovnání s takovými zeměmi jako Belgie, Kanada, Nizozemsko, Portugalsko, Singapur a Španělsko používá ČSÚ při výpočtu kojenecké úmrtnosti jednoduchou formuli (3), kde se počet zemřelých dělí počtem živě narozených ve věku 0. Podobnou formuli používá při sestavování úmrtnostních tabulek i statistický úřad Švédska (15) avšak s tím rozdílem, že počet zemřelých se dělí jen celkovým počtem narozených. Velmi shodné výsledky (viz Tab. 3). se objevují při použití metodiky statistického úřadu Portugalska (14), která ale vyžaduje znalosti o počtech zemřelých podle generací pro dva následující roky.

Značné rozdíly (viz Tab. 4) vykazují hodnoty pravděpodobnosti úmrtí, které byly vypočítané podle metod statistických úřadů Belgie (11) a Kanady (12). Zatímco výsledkem použití belgického modelu je několikrát menší hodnota pravděpodobnosti (0.000068 u chlapců a 0.000071 u dívek) oproti 0.002381 u chlapců a 0.001305 u dívek při použití českého modelu, výsledek vypočítaný dle kanadské metody je naopak několikrát větší. Navíc výpočet dle metodiky kanadského statistického úřadu vyžaduje znalosti generačních úmrtnostních dat za následující rok, stejně jako za rok předchozí.

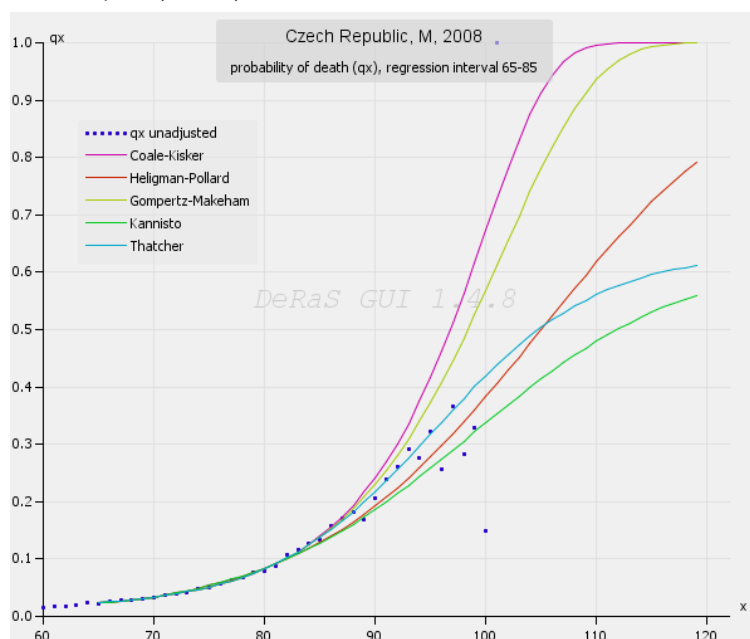
Tab. 4 - Porovnání pravděpodobnosti úmrtí ve věku 0 podle metod vybraných modelů, ČR 2008, chlapci a dívky

	Chlapci	Dívky
Belgie	0.000068	0.000071
ČR	0.002381	0.001305
Kanada	0.036440	0.526827
Portugalsko	0.002117	0.001385
Švédsko	0.002375	0.001302

Zdroj: vlastní výpočty

4.2 Vyrovnání křivky úmrtnosti

Aplikace modelů vyrovnání pravděpodobnosti úmrtí ve starším věku zmíněných v kapitole 3 probíhala pomocí programu DeRaS na data ČR za rok 2008 za obě pohlaví (viz Obr. 3 a 4). Při konstruování jednotlivých funkcí úmrtnostních tabulek používá DeRaS metodu, která je totožná metodě Českého statistického úřadu. Při výběru věkového intervalu, který vstupuje do výpočtu parametrů modelů, byl použit interval 65 až 85 let.

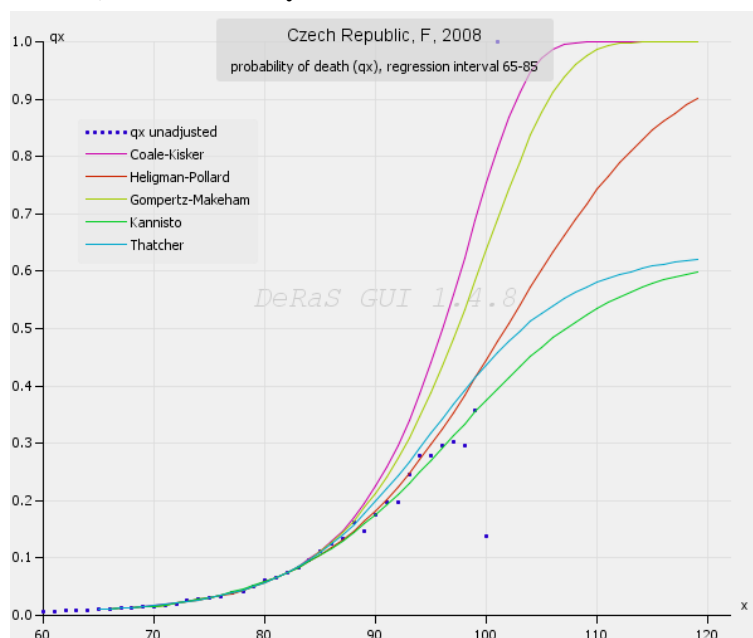
Obr. 3 - Porovnání křivek pravděpodobnosti úmrtí v nejvyšších věcích podle vybraných modelů, ČR, 2008, muži

Zdroj: DeRaS, vlastní výpočty

Na základě dosažených výsledků, lze vybrané modely rozdělit do tří skupin. První skupina zahrnuje metody Gompertz-Makehama a Coale-Kiskera, přičemž jejich charakteristickým rysem je poměrně rychlý nárůst pravděpodobnosti úmrtí, která se blíží hodnotě 1.0 v relativně nižším věku cca 110 let u metody CK a cca 115 let u GM u mužské populace. U žen je situace velmi

podobná, avšak s tím rozdílem, že obě křivky jsou od sebe minimálně vzdálené, a dosahují maximální hodnoty q_x dříve než u mužů (cca 107 let u CK a 112 let u GM).

Obr. 4 - Porovnání křivek pravděpodobnosti úmrtí v nejvyšších věcích podle vybraných modelů, ČR, 2008, ženy



Zdroj: DeRaS, vlastní výpočty

K další skupině patří jen model Heligmana-Pollarda, kde hodnota pravděpodobnosti úmrtí konstantně roste, avšak menším tempem, než u modelů předchozí skupiny. Ve věku 120 let činí hodnota q_x mužů cca 0.9, přičemž dosažení hodnoty 1.0 lze očekávat ve věku 135 let, pokud tempo růstu zůstane stejné. Modelovaná křivka úmrtnosti ženské populace ukazuje mnohem rychlejší tempo růstu, kde hodnota pravděpodobnosti úmrtí 0.9 je dosažená již ve věku 119 let. Dosažení maximální hodnoty lze očekávat ve věku přibližně 130 let.

Zbývající modely Kannisto a Thatcherova jsou poslední třetí skupinou. Vývoj vyrovnané křivky úmrtnosti těchto modelů se od předchozích skupin významně odlišuje. Heligman-Pollardův model popisuje v své podstatě velmi podobný vývoj jako první skupina jen s pomalejším tempem a modelovaná křivka úmrtnosti u modelů třetí skupiny se ve věku 120 let blíží hodnotě pravděpodobnosti úmrtí cca 0.6 u Kannisto a cca 0.55 u Thatcherova u mužů. Vývoj vyrovnané pravděpodobnosti úmrtí žen probíhá velmi podobně, avšak obě proložené křivky leží mnohem blíže k sobě a pohybují se kolem hodnoty 0.6 ve věku 120. Dosažení maximální hodnoty q_x u obou modelů lze očekávat v extrémně vysokých věcích. V tomto případě ukazují modely poslední skupiny neoptimističtější variantu vývoje úmrtnosti ve starších věcích, zatímco metoda Coale-Kiskera naopak patří k variantám nejpesimističtější. Metodu Gompertz-Makehama, kterou používá ČSÚ, lze taky začlenit do této pesimističtější skupiny.

Vývoj modelované střední délky života mužské populace (viz. Tab. 5) je skoro stejný u všech modelů do věku přibližně 65 let. V nejstarších věcích jsou rozdíly mnohem více viditelné, což dobře zobrazuje obrázky 11 a 12. Nejvyšší hodnoty e_x ukazují metody Kannisto a Thatcher, zatímco výsledky při použití metod Coale-Kisker a Gompertz-Makeham jsou minimální.

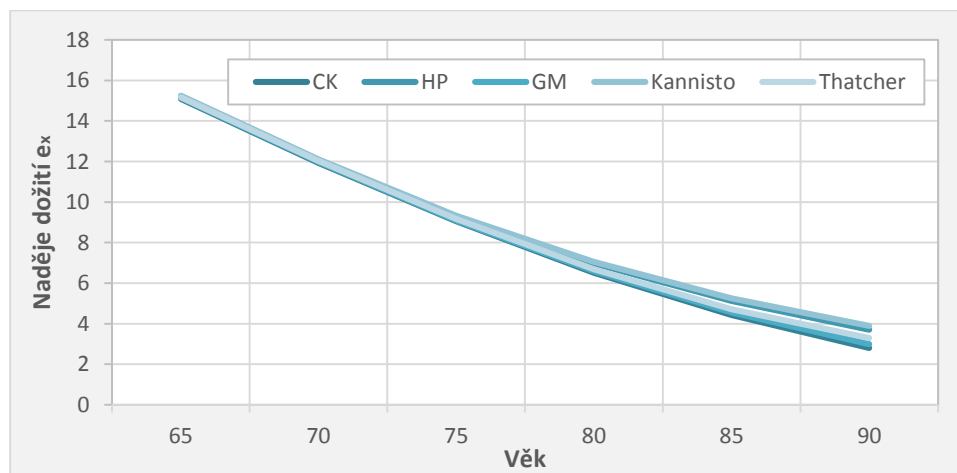
Zajímavé je, že hodnota naděje dožití při použití metody Heligman-Pollarda v nejstarších věcích je větší než u metody Thatcher, avšak jen do věku 105 let, kde se situace obrací (viz Obr. 16).

Tab. 5 - Hodnoty naděje dožití ve vybraných věcích při aplikaci prezentovaných modelů, ČR, 2008, muži

věk	CK	HP	GM	Kannisto	Thatcher
0	73.93	74.03	73.95	74.05	73.99
5	69.22	69.33	69.25	69.35	69.28
10	64.26	64.36	64.28	64.38	64.32
15	59.32	59.42	59.34	59.44	59.38
20	54.47	54.58	54.5	54.6	54.53
25	49.69	49.79	49.72	49.81	49.75
30	44.9	45.01	44.93	45.03	44.96
35	40.11	40.22	40.14	40.24	40.17
40	35.39	35.49	35.41	35.51	35.45
45	30.79	30.89	30.82	30.92	30.85
50	26.38	26.49	26.41	26.51	26.44
55	22.27	22.38	22.3	22.4	22.33
60	18.5	18.62	18.53	18.64	18.57
65	15.09	15.22	15.13	15.25	15.17
70	11.93	12.08	11.98	12.1	12.03
75	9.05	9.29	9.09	9.33	9.16
80	6.52	6.97	6.59	7.05	6.69
85	4.43	5.12	4.55	5.25	4.72
90	2.8	3.7	2.99	3.89	3.29
95	1.65	2.66	1.89	2.92	2.35
100	0.94	1.92	1.18	2.26	1.77
105	0.59	1.42	0.77	1.82	1.44
110	0.5	1.09	0.57	1.54	1.26
115	0.5	0.87	0.51	1.36	1.17
120+	0.5	0.73	0.5	1.25	1.12

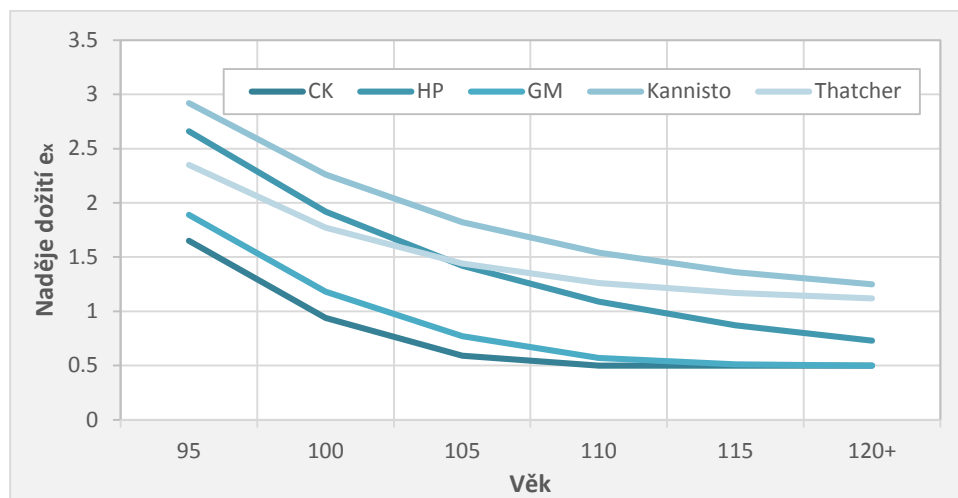
Zdroj: DeRaS, vlastní výpočty

Obr. 5 - Hodnoty naděje dožití ve věku 65 až 90 let při aplikaci prezentovaných modelů, ČR, 2008, muži



Zdroj: DeRaS, vlastní výpočty

Obr. 6 - Hodnoty naděje dožití ve věku 95 až 120+ let při aplikaci prezentovaných modelů, ČR, 2008, muži



Zdroj: DeRaS, vlastní výpočty

Tab. 6 - Hodnoty naděje dožití ve vybraných věcích při aplikaci prezentovaných modelů, ČR, 2008, ženy

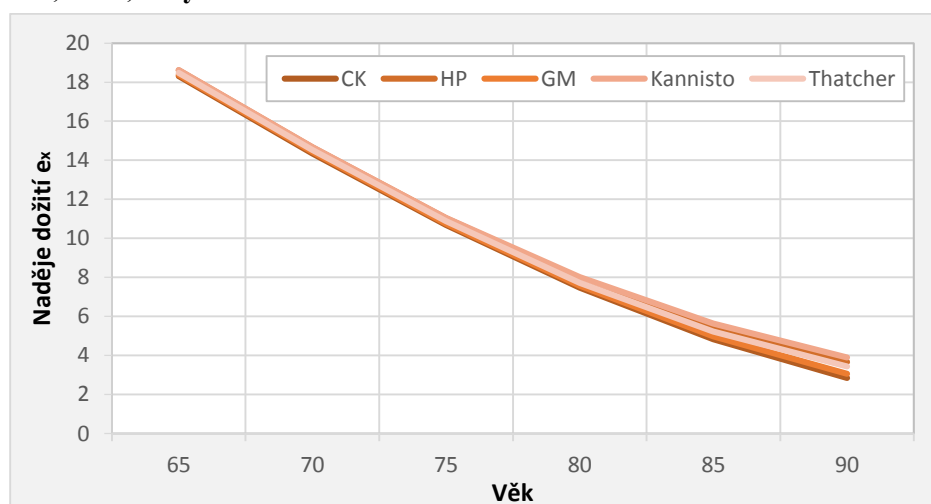
věk	CK	HP	GM	Kannisto	Thatcher
0	80,06	80,33	80,13	80,34	80,22
5	75,30	75,57	75,37	75,58	75,46
10	70,33	70,60	70,40	70,61	70,49
15	65,36	65,63	65,43	65,64	65,52
20	60,43	60,70	60,50	60,71	60,59
25	55,51	55,78	55,57	55,78	55,66
30	50,58	50,86	50,65	50,86	50,74
35	45,69	45,96	45,75	45,96	45,84
40	40,82	41,09	40,89	41,10	40,98
45	36,04	36,32	36,11	36,32	36,20
50	31,36	31,64	31,43	31,65	31,52
55	26,84	27,12	26,90	27,12	27,00
60	22,49	22,78	22,56	22,78	22,65
65	18,31	18,61	18,38	18,62	18,48
70	14,34	14,66	14,42	14,67	14,53
75	10,69	11,03	10,76	11,05	10,88
80	7,48	7,91	7,57	8,02	7,72
85	4,84	5,48	4,98	5,64	5,22
90	2,85	3,66	3,06	3,89	3,43
95	1,53	2,40	1,77	2,72	2,31
100	0,80	1,60	1,01	1,99	1,68
105	0,53	1,11	0,63	1,57	1,36
110	0,50	0,83	0,51	1,34	1,21
115	0,50	0,68	0,50	1,22	1,14
120+	0,00	0,60	0,50	1,15	1,10

Zdroj: vlastní výpočty

Modelovaná střední délka života ženské populace potvrzuje (viz. Tab. 6) zásadní demografický předpoklad, že se ženy dožijí staršího věku než muži (přibližně 80 let ve věku 0 u žen proti 74 let u mužů). Stejně jako u mužské populace jsou rozdíly mezi modely pozorovatelné až na věkovém intervalu 65 a více let. V nejstarších modelovaných věcích ukazují největší hodnoty střední délky života modely Kannisto a Thatcher (1,15 a 1,10). Modely Gompertz-Makeham a Heligman-Pollard mají naopak nejmenší odhady (0,5 a 0,6), přičemž hodnota e_{120+} metody GM je podle odhadů stejná u obou pohlaví.

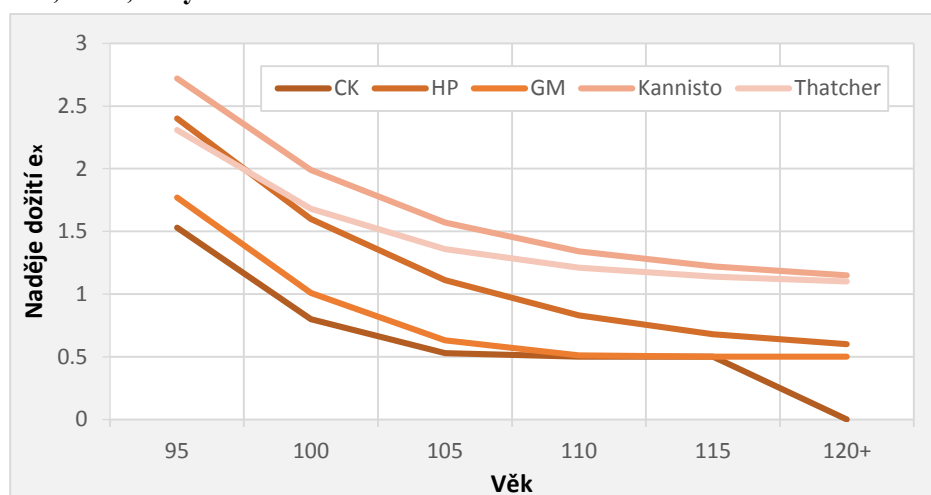
Zajímavým faktem je, že Coale-Kisker je jedinou metodou, která má v nejstarším věku 120+ let hodnoty naděje dožití 0,0 (viz Obr. 8). Podle obrázku 13 lze taky vidět, že se opakuje stejný případ s metodami Thatcher a HP, kde hodnota e_x se podle odhadů Thatchera stává větší než e_x dle HP již ve mladším věku 97 let.

Obr. 7 - Hodnoty naděje dožití ve věku 65 až 90 let při aplikaci prezentovaných modelů, ČR, 2008, ženy



Zdroj: vlastní výpočty

Obr. 8 - Hodnoty naděje dožití ve věku 65 až 90 let při aplikaci prezentovaných modelů, ČR, 2008, ženy



Zdroj: vlastní výpočty

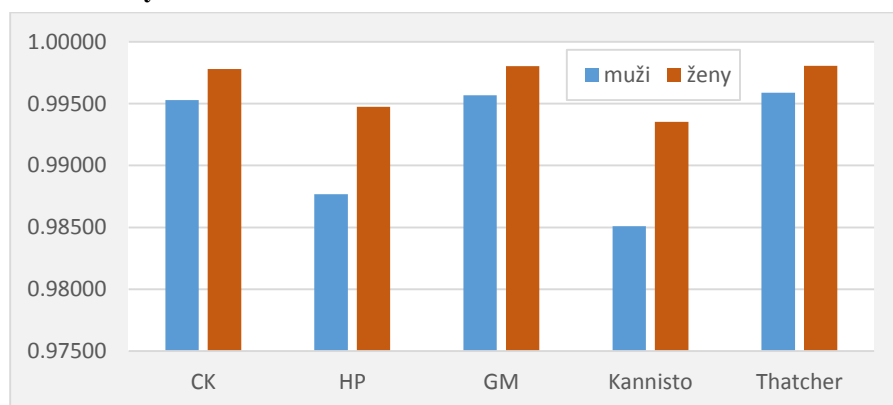
Kromě porovnání grafických výstupů lze výše uvedené modely srovnávat pomocí koeficientu determinace, který nabývá hodnot od 0 do 1, a který znázorňuje, kolik procent rozptylu vysvětlované proměnné je vysvětleno modelem a kolik zůstalo nevysvětleno.

Tab. 7 - Hodnoty indexu determinace při aplikaci prezentovaných modelů, ČR, 2008, muži a ženy

	CK	HP	GM	Kannisto	Thatcher
muži	0.99530	0.98767	0.99568	0.98509	0.99588
ženy	0.99779	0.99474	0.99803	0.99353	0.99807

Zdroj: vlastní výpočty

Obr. 9 - Hodnoty indexu determinace při aplikaci prezentovaných modelů, ČR, 2008, muži a ženy



Zdroj: vlastní výpočty

Z porovnání (viz Tab.7 a Obr. 9) lze vidět, že index determinace u všech modelů nabývá velmi vysokých hodnot a blíží se 1. Nejmenší hodnoty vykazuje model Kannisto u mužů, kde se tato hodnota rovná 0.985, avšak stále zůstává velmi vysoká. Další zvláštností vybraného modelu je největší rozdíl hodnot indexu mezi oběma pohlaví. Podobnou situaci lze pozorovat i u modelu HP, zatímco zbývající modely CK, GM a Thatcher vykazují nejmenší rozdíl. Nejvyšší hodnoty indexu determinace (0.9958 u mužů a 0.998 u žen) má Thatcherův model, ale je nutné zmínit, že rozdíl mezi ním a modely CK a GM je minimální (v řádech tisícinných míst).

Kapitola 5

Závěr

Během zpracování této bakalářské práce bylo provedeno srovnání metodologických postupů při konstruování úmrtnostních tabulek. Potřebné informace byly zjištěny prostřednictvím komunikace se statistickými úřady vybraných zemí pomocí vytvořeného dotazníku. Tento dotazník obsahoval otázky ohledně základních metod tvorby životních tabulek. Vybrané otázky se týkaly periodicity publikovaných tabulek, posledního věkového intervalu, který se používá při jejich konstruování, metod výpočtu pravděpodobnosti úmrtí a vyrovnání křivky pravděpodobnosti ve starších věcích. Následné výsledky analýzy byly porovnány v analytické části, která obsahovala grafické a numerické zobrazení rozdílů vybraných metod a postupů.

Také je nutné zmínit problémy, které se objevily při zjištění potřebných informací, týkajících se metod konstruování životních tabulek. Komunikace probíhala přímo s pracovníky příslušných statistických úřadů, avšak reakce ne vždy dávaly podrobný přehled metod konstrukce. Částou formou odpovědi bylo pouze poskytnutí odkazu na metodologickou publikaci, bez ohledu, jestli tato publikace obsahovala potřebné informace nebo nikoliv. Potíže také vyskytovaly při překládání metodologických poznámek z jiných jazyků, zejména z méně rozšířených. Jako příklad můžeme uvést Vietnam, který nebyl analyzován v této bakalářské práci. Další problém způsobily země, které sice představují velice zajímavý objekt pro zkoumání, avšak nemají jednotný statistický systém, tj. získávání potřebných informací bylo obtížné. Příkladem takové země může být Japonsko, kde sestavování úmrtnostních tabulek má na starosti ministerstvo zdravotnictví, které odmítlo poskytnout přesnou metodiku výpočtu pravděpodobnosti úmrtí.

Bylo taky zjištěno, že existuje několik faktorů, které ovlivňují proces sestavování tabulek. Prvním z nich, který ovlivňuje tento proces, je populační struktura země. Nejpodrobnější informace o populaci poskytují podrobné úmrtnostní tabulky, které publikují všechny prozkoumané země. Tabulky se obvykle publikují jednou ročně, což umožňuje podrobnou analýzu vybraných populací. Avšak pro menší územní celky s malým počtem obyvatel je zkrácená úmrtnostní tabulka výhodnější variantou. Důvodem je menší fluktuace v populační struktuře. V tomto případě je příprava podrobných ročních životních tabulek zbytečná, protože dostatečné informace můžou poskytnout i zkrácené ÚT zpracované za několikaleté období, příkladem čehož mohou sloužit zkrácené tabulky, které publikuje statistický úřad Kanady. Nejužitečnějším typem je průřezová podrobná úmrtnostní tabulka, zatímco kohortní tabulka má spíš historický význam, a většina prozkoumaných zemí nemá tyto tabulky veřejně dostupné.

Dalším důležitým faktorem, který má vliv na proces konstruování úmrtnostních tabulek, jsou metody výpočtu tabulkových funkcí. Během přípravy této bakalářské práce bylo zjištěno, že vybrané statistické úřady používají velmi podobnou metodu konstrukce tabulkových funkcí a rozdíly mezi jednotlivými metodikami jsou minimální. Výjimkou jsou metody výpočtu pravděpodobnosti úmrtí pro věk 0 a metody vyrovnávání křivky úmrtnosti ve starších věcích. Nejpopulárnější metodou výpočtu pravděpodobnosti úmrtí je nepřímá metoda, která poskytuje méně přesné výsledky, avšak je také méně citlivá na dostupnost dat pro jednotlivé soubory demografických událostí. Tuto metodu odhadu pravděpodobnosti úmrtí používají všechny vybrané statistické úřady. Výpočet pravděpodobnosti úmrtí pro věk 0 ve většině případech vyžaduje přesná data o souborech událostí ve věku 0, případně údaje o následujících generacích. Tento postup je vhodný při retrospektivní analýze, ale způsobuje problémy při zpracování dat za aktuální rok, protože údaje následujících generací zatím nejsou dostupné. Proto některé statistické úřady (jako např. ČSÚ nebo statistický úřad Švédska) při výpočtu tohoto ukazatele používají jednoduché vzorce.

Kromě odhadu pravděpodobnosti úmrtí ve věku 0 se práce zabývala analýzou metod vyrovnání křivky úmrtnosti zvláště ve starších věcích. V průběhu práce bylo zjištěno, že nejužívanějším typem metod vyrovnání ve vybraných zemích je logistický model, avšak musí být brán v úvahu relativně malý počet analyzovaných zemí. Co se týče jednotlivých metod, dá se říci, že modely Gompertz-Makehama, který používá ČSÚ, a Coale-Kiskera, významně nadhodnocují úmrtnost ve starším věku. Vzhledem k tomu, že podíl seniorů v současné populaci stále narůstá, může použití tohoto modelu vést ke značnému zkreslení situace v budoucnu. Ale modely Kannisto a Thatcherova také nejsou úplně vhodné pro použití, jelikož jejich odhady pravděpodobnosti úmrtí ve starším věku naopak podceňují úmrtnost v této věkové skupině a odpovídají spíše vzdálené situaci, kdy lidé budou schopni dožít se staršího věku v porovnání se současnou situací. V tomto případě vhodnou alternativní variantou může být využití modelu Heligmana-Pollarda. Je nutné také zmínit, že analýza metod vyrovnání křivek úmrtnosti byla částečně omezena možnostmi programu DeRaS, který zatím umožňuje porovnání jenom některých z nejpoužívanějších metod extrapolace křivky úmrtnosti.

Vybrané téma lze rozšířit celosvětovým srovnáním metod konstruování úmrtnostních tabulek. Zajímavá by mohla být analýza vývojových zemí, např. země Latinské Ameriky neboli států bývalého Sovětského svazu a porovnání jejich metod s metodami, které používají Evropské a další západní statistické úřady nebo mezinárodní organizace.

SEZNAM POUŽITÉ LITERATURY

- ABERNETHY, R. The New Weibull Handbook, Chapter 1: An overview of Weibull analysis. Society of Automotive Engineers, 3rd edition, 1998, ISBN 0965306208.
- BEARD, R. E. 1959. Note on some mathematical mortality models. In Wolstenholme G. E. W. – O'Connor, M. (eds.). *The Lifespan of Animals*. Ciba Foundation Colloquium on Ageing. Boston: Little, Brown, pp. 302–311. Dostupné z: <<http://longevity-science.org/Beard-1959.pdf>>.
- BOLESŁAWSKI, L.; TABEAU, E. 2001. Comparing Theoretical Age Patterns of Mortality Beyond the Age of 80. In Tabeau, E – van den Berg Jeths, A. – Heathcote, Ch. (eds.). *Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective*. Springer Netherlands. ISBN 978-0-7923-6833-5. Dostupné z: <<http://www.springerlink.com/content/vv65332142w06u2h/>>
- BRAVO, M.; MALTA, J. 2007. Estimating life expectancy in small population areas. Constructing assumptions for mortality: data, methods and analysis [online]. Conference of European statisticians. 23 April 2010. [cit. 27.7.2014]. Dostupné z: <<http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.11/2010/wp.10.e.pdf>>.
- BRYAN, T.; HEUSER, R. *The Methods and Materials of Demography*, Academic Press, 2nd edition, 2004, ISBN:0126419558.
- BUETTNER, T. 2002. 2007. Approaches and experiences in projecting mortality patterns for the oldest old. *North American Actuarial Journal* [online]. 2002, vol. 6, no. 3, [cit. 27.7.2014]. Dostupné z: <https://www.soa.org/Library/Journals/NAAJ/2002/july/naaj0207_2.aspx>.
- BURCIN, B.; TESÁRKOVÁ, K.; ŠÍDLO, L. 2010. *Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci*. Demografie. 2010, roč. 52, č. 2, str. 77-89.
- CASELLI, G.; VALLIN, J.; WUNSCH, G. *Demography: Analysis and Synthesis*. Academic Press; 1 edition, 2005, ISBN-10: 012765660X .

- COELHO, E. – MAGALHAES, M. G. – Bravo, J. M. 2007. *Mortality and Longevity Projections for the Oldest-Old in Portugal*. Proceedings of the Joint Eurostat/UNECE (United Nations Statistical Commission and Statistical Office of the Economic Commission for Europe). Work Session on Demographic Projections, Bucharest, Romania, pp. 117-132. Dostupné z: <<http://epc2008.princeton.edu/download.aspx?submissionId=80105>>.
- DEBON, A.; MONTES, F.; SALA, R. 2005. *A comparison of parametric models for mortality graduation. Application to mortality data for the Valencia Region (Spain)* [online]. July-December 2005, str. 269-288 [cit. 27.7.2014]. Dostupné z: <<http://www.idescat.cat/sort/sort292/29.2.7.debon-et-al.pdf>>.
- HINDE, K. 1998. *Demographic methods*. 1st edition: Hodder Arnold Publication , 1998. - ISBN 978-0340718926
- KALIBOVÁ, K.; PAVLÍK, Z. Mnohojazyčný demografický slovník [český svázek]. 2. vyd. Praha: Česká demografická společnost, 2005. ISBN 80-239-4864-4.
- MÉSZÁROS, J. 2000. *Metodický materiál. Výpočet úmrtnostních tabulek* [online]. INFOSTAT, 2000 [cit. 10.8.2014]. Dostupné z: <http://www.infostat.sk/vdc/pdf/metodika_ut.pdf>.
- PASTOR, K. 2007. A Cubic Model for Graduation of Life Tables. *Journal of the Applied Mathematics, Statistics and Informatics* [online]. 2007, vol. 3, no. 2, [cit. 27.7.2014]. Dostupné z: <http://jamsi.fpv.ucm.sk/docs/v03n02_12_2007/v03n02_08_PASTOR.pdf>.
- PAVLÍK, Z.; RYCHTAŘÍKOVÁ, J.; ŠUBRTOVÁ, A. *Základy demografie*. 1. vyd. Praha : Academia , 1986. - ISBN 80-200-0437-8.
- THATCHER, A. R. 1999. The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, Vol. 162, No. 1, 1999, pp. 5-43. Dostupné z: < <http://www.jstor.org/stable/2680465> >.
- YUEN, K. C. 1997. Comments on some parametric models for mortality tables. *Journal of Actuarial Practice* [online]. 1997, vol. 5, no. 2, [cit. 27.7.2014]. Dostupné z: <http://jofap.org/documents/vol5/v5no2_yuen.pdf>.
- PRINS, K.; KUIJPER, H. 2007. Bevolkingsstatistieken onder het persoonskaartenstelsel en het GBA-stelsel: overeenkomsten en verschillen. *Bevolkingstrends* 55(1), str. 14-33.
- BANABIC, Dorel. *Advanced Methods in Material Forming* [online]. Springer, 2007 [cit. 13.6.2011]. Dostupné z: <http://lib.myilibrary.com?ID=86355>
- Statistics Canada. *Methods for Constructing Life Tables for Canada, Provinces and Territories* [online]. Minister of Industry, 2012 [cit. 27.7.2014]. Dostupné z: <<http://www.statcan.gc.ca/pub/84-538-x/84-538-x2013001-eng.pdf>>.

Singapore Department of Statistics. *Complete Life Tables 2003-2006 for Singapore Resident Population* [online]. Singapore Department of Statistics, February 2008 [cit. 27.7.2014].

Dostupné z:

<http://www.singstat.gov.sg/publications/publications_and_papers/births_and_deaths/ip-s13.pdf>.

SEZNAM DATOVÝCH ZDROJŮ

Instituto Nacional de Estadística. June 2013. *Life Tables Methodology* [online]. [cit. 25.7.2014]. Dostupné z: <http://www.ine.es/en/daco/daco42/mortalidad/metodo_9111_en.pdf>.

JOHANSSON, Tomas. *Swedish lifetables* [elektronická pošta]. Message to: information@scb.se. 8. července 2014 [cit. 25.7.2014]. Osobní komunikace.

LEFEBER, Algie. *Population [#37058]* [elektronická pošta]. Statistics Netherlands. Message to infoservice@cbs.nl. 9. července 2014 [cit. 25.7.2014]. Osobní komunikace.

HENNESSY, Carol Anne. *Life tables methodology question* [elektronická pošta]. Central Statistics Office of Ireland. Message to CarolAnne.Hennessy@cso.ie. 9. července 2014 [cit. 25.7.2014]. Osobní komunikace.

Statistics Portugal. Sending an explanation/link - PED-218878966 [elektronická pošta]. Message to info@ine.pt. 9. července 2014 [cit. 25.7.2014]. Osobní komunikace.

Office of National Statistics: *Guide to Calculating National Life Tables*. [online]. Office of National Statistics [cit. 25.7.2014]. Dostupné z: <<http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/demography/guide-to-calculating-national-life-tables/index.html>>.

Office of National Statistics: *Calculating population estimates of very old*. [online]. Office of National Statistics [cit. 25.7.2014]. Dostupné z: <<http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/pop-ests/calculating-estimates-of-the-very-elderly/index.html>>.

PLAKK, Marie. *Life table methodology question* [elektronická pošta]. Population and Social statistic department Statistics Estonia. Message to mari.plakk@stat.ee. 11. července 2014 [cit. 25.7.2014]. Osobní komunikace.