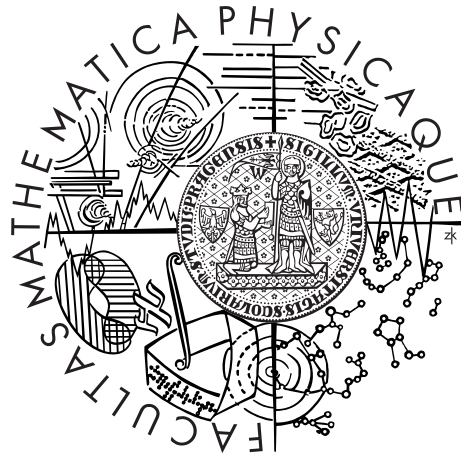


Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## DIPLOMOVÁ PRÁCE



Bc. Kateřina Štádlarová

## Projekce úmrtnosti podle příčin úmrtí

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Mgr. Jan Kořístka

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2013

Ráda bych poděkovala svému vedoucímu diplomové práce, Mgr. Janu Kořistkovi, za ochotu, cenné připomínky a rady.

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V ..... dne .....

Podpis autora

Název práce: Projekce úmrtnosti podle příčin úmrtí

Autor: Bc. Kateřina Štádlerová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Mgr. Jan Kořistka

Abstrakt: Práce se zabývá projekcemi úmrtnosti podle příčin úmrtí. Součástí práce je také aplikace těchto poznatků na data české populace. Projekce úmrtnosti se v dnešní době používají čím dál častěji kvůli stárnutí obyvatelstva. Výsledky této práce by mohly být zajímavé jak pro finanční instituce jako pojišťovny, tak například i pro penzijní plánování v daných oblastech politiky. Dosud o této problematice není napsáno mnoho článků v českém jazyce ani dosud nejsou zveřejněny výsledky takovýchto projekcí pro česká data.

Klíčová slova: projekce, úmrtnost, Lee-Carter, příčiny

Title: Mortality projections by cause

Author: Bc. Kateřina Štádlerová

Department: Department of probability and mathematical statistics

Supervisor: Mgr. Jan Kořistka

Abstract: The thesis focuses on the mortality projections by causes of death. The thesis includes also the application of such knowledge on the data of the Czech population. The mortality projections are used nowadays more and more often due to the population ageing. The results of this thesis may be interesting both for financial institutions such as insurance companies and for the purposes of certain areas of government policy with regards to the pension planning. So far not many articles have been published in the Czech language, nor are there any published results of similar projections using the data derived from the Czech environment.

Keywords: projection, mortality, Lee-Carter, causes

# Obsah

Úvod	3
<b>1 Vlivy na úmrtnost ve velkých zemích během 20. století</b>	<b>6</b>
1.1 Kanada	6
1.2 Anglie a Wales	6
1.3 Francie	7
1.4 Německo	7
1.5 Japonsko	7
1.6 Švédsko	8
1.7 USA	8
<b>2 Úvod do metod projekce měr úmrtnosti</b>	<b>9</b>
2.1 Historie	9
<b>3 Extrapolační metody</b>	<b>11</b>
3.1 Jednoduché extrapolační metody	11
3.1.1 Lineární změna	12
3.1.2 Geometrická změna	12
3.1.3 Exponenciální změna	12
3.2 Podílové extrapolační metody	13
3.2.1 Constant-share	13
3.2.2 Shift-share	13
3.2.3 Share-of-growth	13
3.3 Složité extrapolační metody	13
3.3.1 Jednofaktorové modely	14
3.3.2 Dvoufaktorové modely	15
<b>4 Jiné metody projekce úmrtnosti</b>	<b>20</b>
4.1 Procesové metody	20
4.2 Výkladové metody	20
4.3 Metody očekávání	20
<b>5 Rozdělení podle příčin úmrtí</b>	<b>22</b>
5.1 Vývoj hlavních příčin úmrtí ve 20. století	22
5.1.1 Vývoj v Anglii a Walesu	22
5.1.2 Vývoj v Japonsku	23
5.1.3 Vývoj v USA	23
5.2 Rizikové faktory podle příčin úmrtí	23
5.2.1 Ischemická choroba srdeční a cévní příhoda mozková	23
5.2.2 Rakovina	24
5.2.3 Nemoci dýchací soustavy: chronická obstrukční plicní nemoc (COPD)	24
5.2.4 Diabetes	24
5.3 Projekce podle příčin úmrtí	25

<b>6</b>	<b>Projekce úmrtnosti podle příčin úmrtí pro českou populaci</b>	<b>26</b>
6.1	Data . . . . .	26
6.1.1	Skupiny příčin úmrtí . . . . .	26
6.1.2	Věkové skupiny . . . . .	27
6.1.3	Nulové hodnoty . . . . .	28
6.2	Aplikace Lee-Carterovy metody . . . . .	29
6.2.1	Věkově-specifické parametry . . . . .	29
6.2.2	Odhad indexu úmrtnosti . . . . .	30
6.2.3	Projekce indexů úmrtnosti . . . . .	31
6.2.4	Projekce centrálních měř úmrtnosti . . . . .	32
6.2.5	Chybová složka . . . . .	35
6.2.6	Agregace . . . . .	36
6.2.7	Střední doba života . . . . .	38
	<b>Závěr</b>	<b>42</b>
	<b>Seznam použité literatury</b>	<b>44</b>
	<b>Seznam tabulek</b>	<b>46</b>
	<b>Přílohy</b>	<b>47</b>

# Úvod

Pro projekce úmrtnosti je zásadní uvědomit si, co vše vlastně lidskou úmrtnost ovlivňuje. V životě každého jedince existují rizikové faktory, které ovlivňují délku života. Do rizikových faktorů patří například stáří, prodělané nemoci, rizika v povolání, přístup k léčbě nebo kouření. Ne vždy však máme pro rizikové faktory potřebná data. Místo samotného rizikového faktoru se v takovém případě nabízí alternativa v podobě použití nějakého vhodného zástupce. Asi nejvýznamnějším zástupcem je rok narození.

Matematik Benjamin Gompertz v roce 1825 vypořádal, že míra úmrtnosti roste exponenciálně s věkem, a že šance na úmrtí se přibližně zdvojnásobují každých osm let života. Věk nás také může informovat o expozici vůči rizikovým faktorům v nějaké skupině lidí. V článku [17] je například uvedeno, že jedinci narození v Německu a Japonsku ve třicátých letech měli tu výhodu, že v sedmdesátých letech zažili protikuřácké kampaně a také pokrok v léčbě vysokého tlaku a cholesterolu.

Projekce úmrtnosti se také musí vypořádat se stále rostoucí délkou života. Podle [17] významně roste počet lidí dožívajících se sedmdesátého a osmdesátého roku života. Všimnout si toho můžeme například u švýcarských mužů narozených v roce 1900, kteří se průměrně dožili věku 51 let. O století později se tento věk předpokládal už o více než 35 let vyšší.

Růst délky života můžeme vysvětlit následujícími faktory, rozdělenými do tří skupin jako v [17]:

- Sociální faktory – do této skupiny patří například pokles počtu kuřáků, za který mohou především zákazy kouření na veřejných místech a rostoucí daň z tabákových výrobků. Dalším faktorem je zvyšující se povědomí o zdravém životním stylu, který dodržuje stále více lidí. Důsledkem může být pokles ischemických a srdečních onemocnění.
- Pokroky v léčbě – přirozeně zvyšují šanci pacienta na přežití. Mezi největší pokroky v USA v letech 2000–2010 patří například pokles srdečních onemocnění o 40 %, laboratorní průlomů ve výzkumu kmenových buněk, rozšíření cílených terapií u rakovin a nové léky nebo využívání minimální invaze a robotických technik u chirurgických zákroků.
- Preventivní metody – je lepší věnovat se prevenci, než léčbě chronických onemocnění. Je to výhodnější jak z hlediska nákladů, tak spotřeby zdrojů. Díky prevenci se dá předejít mnohým koronárním případům a díky lékům se dá redukovat počet úmrtí na srdeční onemocnění až o 42 %. American Heart Association v této souvislosti stanovila několik bodů, které by měli lidé dodržovat, aby si udrželi dobré zdraví. Tento seznam je znám pod názvem LifeSimple 7 a jeho položky jsou následující:
  - Být aktivní – fyzická aktivita alespoň 150 minut týdně s mírnou intenzitou.
  - Kontrolovat cholesterol – koncentrace cholesterolu by měla být menší než 200 mg/dL.

- Lépe se stravovat – 4 až 5 složek zdravé výživy.
- Kontrolovat krevní tlak – krevní tlak by měl být pod 120/80 mmHg.
- Zhubnout – BMI mezi 18,5 a 25.
- Snížit hladinu cukru v krvi – koncentrace glukózy by měla být menší než 100 mg/dL.
- Přestat kouřit.

Rostoucí délka života má dopad i po finanční stránce. Téměř všichni musí brát tento růst v potaz, aby mohli správně oceňovat budoucí finanční toky. Z tohoto hlediska je správná projekce velmi důležitá. V opačném případě by mohlo docházet například k nedostatečnému rezervování penzijních fondů a ke vzniku závazků, se kterými se, na základě špatné projekce, v plánech nepočítalo. Pojistitelé ale nejsou jediní, kdo by se měl touto problematikou zabývat. Podle [17] jsou hlavními skupinami, kterých se toto týká také vláda, zaměstnavatelé i jednotlivci:

- Vláda musí předpovídat dlouhověkost kvůli stanovování penzijního připojištění a důchodového věku, který se v mnoha industrializovaných zemích neustále zvyšuje, právě z důvodu růstu délky života lidí. Projekce se využívají také pro stanovení správné daně nebo zdravotních programů.
- Zaměstnavatelé musí správně financovat penze a výhody, které přislíbili svým zaměstnancům.
- Pojistitelé musí vyhovět požadavkům klientů poskytovaním anuit a životních pojištění, musí správně ocenit své produkty a ukázat, že jsou schopni plnit své závazky. Solventnost II uznává jako vhodný prostředek na snižování rizika dlouhověkosti zajišťování.
- S rostoucí délkou života souvisí také stárnutí populace, kvůli kterému se stále více očekává od jednotlivců, že si kromě penzijního připojištění vytvoří další oddělené zajištění, ze kterého budou financovat svůj důchod.

Pochopení úmrtnosti je samo o sobě velmi složitým problémem a existuje mnoho přístupů používaných k jejímu co nejlepšímu osvětlení. Uvést můžeme například tzv. multidisciplinární přístup, který zahrnuje poznatky z mnoha profesí, jakými jsou například aktuarii, demografové, lékařští experti, epidemiologové, farmakologové nebo gerontologové.

Tato práce se zabývá projekcí úmrtnosti podle příčin úmrtí. Jejím cílem je takovou projekci pro českou populaci vytvořit a výsledky porovnat s projekcí pro českou populaci získanou z celkových měř úmrtnosti. Tímto porovnáním lze zjistit, který z těchto dvou postupů je pro projekce centrálních měř úmrtnosti pro českou populaci vhodnější. Dále je možné pozorovat rozdíly mezi výsledky získanými těmito dvěma způsoby a zkoumat případné závislosti.

První kapitola práce se věnuje různým vlivům na úmrtnost ve dvacátém století v Japonsku, Francii, Kanadě, Švédsku, západním Německu, USA a v Anglii a Walesu. Je zde popsáno, z jakých důvodů jsou zvoleny právě tyto země, čím se liší a v čem jsou si vzájemně podobné. Pro každou zemi jsem zvlášť rozepsala hlavní vlivy na úmrtnost v daném století. Ve většině zemí jde o změny související s počtem kuřáků, světovými válkami či lékařskou péčí.



Druhá kapitola je věnována úvodu do metod projekce měř úmrtnosti. Jejím účelem je především získání obecného povědomí o projekčních metodách. V této části jsem popsala, jakými různými způsoby a do jakých skupin lze metody projekce měř úmrtnosti dělit. Kapitola rovněž obsahuje krátkou historii projekčních metod, jejich vývoj, hlavní důvody jejich vzniku a především jejich rozsáhlé potřeby v současnosti.

Ve třetí kapitole podrobněji popisují extrapoláční metody a celou kapitolu jsem rozdělila na tři podkapitoly – jednoduché extrapoláční metody, podílové extrapoláční metody a složité extrapoláční metody. V každé podkapitole uvádím několik metod, které spadají do jednotlivých skupin, nejvíce se však věnuji složitým extrapoláčním metodám. Do této skupiny totiž patří Lee-Carterova metoda, kterou v poslední kapitole používám pro projekci podle příčin úmrtí i pro projekci celkových měř úmrtnosti pro českou populaci. Právě tuto metodu jsem tedy v této kapitole popsala nejpodrobněji.

V další kapitole jsem pak stručněji popsala další projekční metody, které se používají.

Pátá kapitola je rozdělena do tří podkapitol. První z nich pojednává o vývoji hlavních příčin úmrtí ve dvacátém století. Nabízí tedy podobný náhled na země, jakým se zabývám již v úvodu, zde jsem se však více zaměřila na vývoj příčin úmrtí v jednotlivých zemích, než na vývoj různých vlivů na úmrtnost. Druhá podkapitola se zabývá rizikovými faktory podle příčin úmrtí. Popsala jsem zde, co může vést k úmrtí na ischemické choroby srdeční a cévní příhodu mozkovou (mrtvici), rakoviny, nemoci dýchací soustavy a diabetes. V poslední podkapitole této kapitoly jsem zmínila některé výhody a nevýhody projekcí úmrtnosti podle příčin úmrtí.

V poslední kapitole už se zabývám projekcí měř úmrtnosti podle příčin úmrtí pro českou populaci. Nejprve popisují, jak jsem získala data, jaké úpravy jsem na nich následně provedla a z jakého důvodu jsem vůbec data upravovala. Poté se zabývám samotnými projekcemi, které jsem vytvářela pomocí Lee-Carterovy metody. Podrobně jsem zde popsala celý postup projekce včetně popisu řešení problémů, na které jsem v průběhu práce narazila. Popisují zde získané výsledky a text jsem obohatila o tabulky a grafy ukazující některé zajímavé výsledky. Dále jsem v této kapitole popsala výsledky získané agregací centrálních měř úmrtnosti podle příčin úmrtí do celkových měř úmrtnosti a jejich následným porovnáním s projekcemi celkových centrálních měř úmrtnosti. Nakonec v této kapitole ještě uvádím výpočty a výsledky pro střední dobu života.

# 1. Vlivy na úmrtnost ve velkých zemích během 20. století

V průběhu 20. století se mnoho věcí změnilo. Ve většině industrializovaných zemí došlo k výrazným změnám v zastoupení kuřáků ve společnosti, k pokrokům ve zdravotnictví a v životním stylu obyvatel, v některých zemích došlo ke změnám ve zdravotním systému a nemalý vliv na země měly samozřejmě také světové války. Stručně se zaměřím na změny, ke kterým v průběhu 20. století došlo v sedmi velkých zemích, a to v Japonsku, Francii, Kanadě, Švédsku, západním Německu, USA a Anglii a Walesu. Změny, ke kterým v těchto zemích docházelo, jsou podrobněji popsány v článku [12]. Tyto země mají, s výjimkou Švédska, společnou příslušnost ke skupině G8, v mnoha ohledech se ale významně liší. Podíváme-li se na kohortní efekt v těchto zemích, zjistíme, že USA ho má oproti ostatním velmi slabý, ne-li žádný, zatímco data Japonska, Anglie a Walesu vykazují kohortní efekt silný. USA se oproti Kanadě liší také v přístupu ke zdravotní péči. Francie a západní Německo jsou do přehledu vybrány především z důvodu odlišností v samotných zemích. Ve Francii je velký rozdíl mezi ženskou a mužskou úmrtností, v Německu docházelo k výrazným změnám v průběhu století, a to hlavně v sociální a ekonomické sféře. Nakonec je zařazeno Švédsko kvůli dlouhé historii kvalitních dat.

## 1.1 Kanada

Míry zlepšení úmrtnosti zde nevykazují kohortní efekt. Populační úmrtnost v jakémkoli čase je totiž více ovlivněna přílivem imigrantů z různých zemí, než změnou životních standardů původního obyvatelstva. Právě imigrace je hlavním sociálním faktorem ovlivňujícím kanadskou populaci během dvacátého století. Zastoupení narozených v jiné zemi než Kanadě se zde na začátku století pohybovalo okolo 22 %, na konci stále ještě okolo 16 %. Dále lze o Kanadě říci, že má velmi silné sociální a zdravotní systémy. Ministerstvo zdravotnictví se dokonce zavazuje k tomu, aby země patřila mezi nejzdravější na světě, měřeno dlouhověkostí, životním stylem a efektivním využitím systému veřejné zdravotní péče. Tento přístup může být viděn například ve vztahu ke kouření. Úspěch Programu kontroly tabáku (spuštěný v roce 1999) je vidět na celkové redukci zastoupení kuřáků z 25 % v roce 1999 na 18 % v roce 2006.

## 1.2 Anglie a Wales

V průběhu dvacátého století byly země ovlivněny především změnami ve zdravotním systému a oběma světovými válkami. Co se týče zdravotního systému, v roce 1948 došlo k založení Národní zdravotní služby, která poskytuje zdravotní péči všem občanům UK založenou na potřebě, nikoli na schopnosti platit, jako to bylo do té doby. Hlavní jevy, které přispěly ke zlepšení úmrtnosti v důchodovém věku, potom zahrnují kohortní efekt, menší zastoupení kuřáků, rozšíření rozdílů mezi sociálními třídami, kde u osob z vyšších sociálních tříd došlo k většímu zlepšení

očekávané doby života, než u osob ze slabších sociálních tříd a dále tzv. „stárnutí zlepšení úmrtnosti“, což znamená zrychlení zlepšení úmrtnosti ve vysokých věcích.

### 1.3 Francie

Mezi lety 1945 a 1974 došlo ve Francii k demografickému růstu. Země byla v tomto období ve fázi rekonstrukcí, měla stabilní ekonomický růst a v šedesátých letech také příliv pracovních imigrantů. V roce 1973 vedla energetická krize ke změnám imigračních zákonů a dnes má Francie rostoucí populaci především díky přirozenému růstu. V roce 1977 bylo zavedeno očkování proti chřipce pro starší občany a od roku 1985 je dokonce pro rizikové skupiny zdarma, což dopomohlo ke stabilizaci rozdílů v úmrtnosti ve vyšších věcích. Dalším počinem země vedoucím ke snížení úmrtnosti jsou kampaně zaměřené na bezpečnost jízdy na silnicích, což ovlivnilo především věkovou skupinu 20–40 a vedlo ke snížení počtu úmrtí způsobených dopravními nehodami mezi lety 2001 a 2005 o 36 %. Stejně jako v ostatních zemích, i zde došlo ke snížení počtu kuřáků, ačkoli Francie byla poslední z uvedených zemí, která zavedla zákaz kouření na pracovištích (2007).

### 1.4 Německo

Německo má velmi extrémní sociální historii. První útrapy přišly s první světovou válkou v podobě nedostatku jídla a přidělovým systémem na chleba a mouku, později ještě mléko a maso. Další rána přišla s krachem na newyorské burze, který se projevil rekordní nezaměstnaností. Druhá světová válka pak přinesla vynucené migrace a další hladovění. Také historie týkající se spotřeby cigaret je jedinečná. Země byla jednou z prvních mající kampaně proti kouření před druhou světovou válkou i během ní. Jelikož ale šlo o nacistická opatření, poválečné Německo naopak zaostávalo za ostatními zeměmi v zákazech kouření v uzavřených veřejných prostorech. Dokonce až do roku 2007 bylo v Německu neslýchané mít v restauracích nekuřáckou sekci, a každý, včetně dětí, si mohl koupit cigarety v prodejních automatech. Pokud by někoho tato problematika zajímala více, může se podívat do [11].

### 1.5 Japonsko

Japonsko je v současnosti zemí s nejdelší očekávanou dobou života na světě, a to jak pro muže, tak ženy. Významným příspěvkem ke snížení úmrtnosti byla redukce infekčních onemocnění v letech 1955–1960 (například díky antibiotikům), které do té doby byly hlavní příčinou úmrtí nemluvňat a dětí. Největší příspěvek k rostoucí dlouhověkosti se pak posunul z mladších věkových skupin do těch starších. K vývoji úmrtnosti v Japonsku během posledních let 20. století velmi přispěl velký pokles onemocnění mozkových cév, který tradičně býval výrazně vyšší než v evropských zemích, vyjma Itálie. Za tento trend může hlavně změna v japonském stravování, kde se ze slaných jídel přešlo k méně slaným pokrmům bohatším na proteiny. Za to může například zavedení ledniček do japonských domácností. Očekávaná doba života Japonců dokonce roste rychleji než u ostatních

zemí i navzdory tomu, že japonští muži mají zatím nejvyšší zastoupení kuřáků ze zemí zde zmíněných a míra úmrtnosti na rakovinu plic stále roste, zatímco ve většině ostatních zemích klesá.

## 1.6 Švédsko

Švédsko bylo jen velmi málo ovlivněno válkou, jelikož když vypukla první světová válka, Švédsko se prohlásilo za neutrální a zůstalo jím dokonce i během druhé světové války. Po válce byla země netknutá, s nedotčeným průmyslem a produkovala zboží chtěné v mnoha zdevastovaných zemích Evropy. Tyto okolnosti vedly k ekonomickému rozmachu a rychlému rozvoji sociálního systému. Co se týče kouření, Švédsko tvoří výjimku mezi ostatními zeměmi zde zmíněnými. Zatímco u žen je spotřeba tabákových výrobků víceméně v souladu s ostatními evropskými zeměmi, muži ji většinou mají menší. Důvodem je používání tzv. vlhkého šňupacího tabáku, který je v ostatních zemích Evropské unie zakazán, ačkoli různé studie ukazují, že jeho používání nevede k významnému zvyšování úmrtnostních měr. Podrobnější studie na téma spotřeby tabákových výrobků ve Švédsku je v [14] a [17].

## 1.7 USA

Společně s ostatními zmíněnými zeměmi má USA v posledním století výrazně klesající míry úmrtnosti v kojeneckém věku a ve věcích pod 65 let. Dnes nastávají tři čtvrtiny všech úmrtí po 65. roce života, kdežto v roce 1900 to bylo o něco méně než jedna pětina. USA je země s nejvyšší mírou obezity mezi všemi zeměmi OECD (Organization for Economic Cooperation and Development) a její růst pravděpodobně povede k nástupu chronických onemocnění jako jsou cukrovka nebo astma. Mezi zmíněnými zeměmi je USA výjimečné ve výši útrat za zdravotnictví (poměrově k HDP), která je nejvyšší. Zajímavé ovšem je, že i přes vysoké dotování zdravotnictví jsou očekávané doby života nižší než u ostatních zemí vynakládajících do tohoto odvětví méně. O kouření pak můžeme říci, že spotřeba tabáku u mužů se v posledních desetiletích významně snížila, a to až do té míry, že se spotřeba u žen a mužů téměř vyrovnala. Spolu s Kanadou a Švédskem má USA nejmenší poměr kuřáků mezi dospělými mezi zeměmi OECD.

## 2. Úvod do metod projekce měř úmrtnosti

Pro úlohu projekce měř úmrtnosti se dá vybírat z velkého množství různých metod. Nedá se obecně říci, že by některé byly lepší, než-li jiné. Liší se v přístupu a na různá data nemusí být vhodná vždy stejná metoda. Obecně se však nedá určit, která z metod by byla zrovna pro naše data nejvhodnější, ani kdy již zvolená metoda nejlépe funguje.

Metody projekce měř úmrtnosti se mohou dělit do různých skupin. V různých článcích se však toto dělení velmi liší a záleží tak na každém, jaké dělení zrovna vyhovuje jeho záměrům. Já zde ukážu některé způsoby, jakými lze metody dělit do jednotlivých skupin, používat však budu také pouze jeden z nich.

Jednou z nejjednodušších možností, jak metody dělit, je na demografické a nedemografické modely. Rozdíl mezi těmito dvěma skupinami je v povaze proměnných. Zatímco v demografických modelech se používají demografické proměnné, v nedemografických máme proměnné vnější. Příkladem demografického modelu jsou extrapoláčnı a interpolačnı metody, mezi nedemografické modely patří například biomedicınské nebo epidemiologické.

Dalším způsobem je dělení metod na objektivnı a subjektivnı, přičemž objektivnı metody mohou být dále děleny například na procesové, výkladové, extrapoláčnı a metody očekávání. Toto dělení budu dále používat v této práci. Obecně lze o objektivnıch metodách říci, že jejich proměnné, zdroje dat a projekčnı techniky jsou přesně určené a následná interpretace je tedy jednodušší, než v případě subjektivnıch metod, u kterých není nic z tohoto určeno přesně. I u objektivnıch metod je však prostor pro určitou subjektivitu. Vždy je potřeba nějakého soudu týkajícího se zdrojů dat nebo volby projekčnı metody. Výběr metody dokonce nemusí být nijak striktnı. Můžeme použít i nějakou kombinaci různých metod, například pro různé věkově skupiny nebo časově úseky. Modely můžeme také dělit podle počtu faktorů na tzv. bezfaktorové, jednofaktorové, dvoufaktorové a třífaktorové. Tedy modely mohou zahrnovat až tři faktory, kterými jsou věk, kohorta a období. Projekce se také dají vytvářet buď pro nějaký celek, nebo odděleně pro částı, jako v případě, kterým se zabývá tato práce, kdy projektujeme úmrtnost zvlášť pro jednotlivě příčiny úmrtı.

Než přistoupım k popisu jednotlivých metod, zmınım se zde ještě stručně o historii projekcí, převzaté převážně z [12] a zčástı z [17].

### 2.1 Historie

K největší změně na poli demografických prognóz došlo zhruba před pětadvaceti lety, kdy se demografové začali zajımat o populačnı projekce, které byly do této doby demografy nepoužívané a nedocenené. Demografové začali pro své cíle využívat metody používané v jiných disciplınách a vyvíjet nově přístupy.

V roce 1980 se pro populačnı projekce používala tzv. kohort-komponentnı metoda. Byla představena Cannanem v roce 1895 v [6] a později používána Bowleym (1924, více v [4]) a ještě později znovu objevena Whelptonem (1928, více v [20]). Metoda je založena na rozdělení populace do skupin podle věku a pohlavı a ná-

slednému počítání porodnosti, úmrtnosti a migrace pro každou takovou skupinu zvlášť. K projekci se potom dá využít různých technik pro každou skupinu zvlášť. Nejobtížnější částí je zde právě ono rozdělení populace do skupin.

Populační prognózy jsou velmi významné a mají široké využití. Používají se například v sociální, ekonomické, ekologické nebo i politické sféře pro plánování a mají také další uplatnění týkající se například porodů, školství (pro školní zápisy) nebo domácností. To vede také k širokým možnostem, jak s populačními projekcemi pracovat. Dají se vytvářet pro různé časové horizonty, které mohou být kratší než rok, ale také dlouhé jako délka života. Lze je vytvářet také pro celosvětovou populaci nebo populaci pouze na nějakém konkrétním území, oblasti, zemi nebo jenom městě. Pro větší detail se populace ještě klasicky dělí podle pohlaví. Základními složkami ovlivňujícími velikost populace jsou potom porodnost, úmrtnost a migrace. Tyto složky se mohou projektovat odděleně každá zvlášť a následně je lze vhodně kombinovat.

Ve dvacátém století byla považována za nejdůležitější složku právě porodnost. Zjistilo se však, že je tento úkol dost složitý, jelikož po válce nikdo nepočítal s „baby boomem“ ani s následným „baby bustem“. Stejně tak nikdo nepočítal s rychlým poklesem úmrtnosti po válce a s následným růstem střední délky života. Složka migrace byla zase považována za nedůležitou. Toto všechno nějakým způsobem ovlivňuje strukturu obyvatelstva a po neočekávaných událostech v porodnosti získaly populační projekce na důležitosti.

V současné době jsou populační projekce důležité z hlediska stárnutí obyvatel. Dochází k nerovnoměrnostem ve věkové struktuře, zapříčiněné menší porodností a prodlužováním očekávané doby života. Jak už jsme si řekli hned v úvodu, toto vede k problémům v sociálním zabezpečení a ve zdravotnictví, což činí populační projekce stále více žádoucími. Některé země se, pomocí migračních zákonů, snažily stárnutí obyvatel zabránit, nebo alespoň snížit. Nechaly do svých zemí legálně proudit imigranty, což ovšem vedlo k opačnému efektu, kdy populace rostla více díky migraci, než díky porodnosti původních obyvatel. Toto je případ například Kanady, jak jsem zmínila už v předchozí kapitole.

## 3. Extrapolační metody

Extrapolační metody fungují tak, že se sledují trendy úmrtnosti v minulosti a poté se projektují do budoucnosti. Jinými slovy se prokládají historická data pomocí matematických modelů, pomocí kterých se potom i projektují budoucí hodnoty. Předpoklad je zde v podstatě jediný, a to ten, že budoucnost bude v nějakém smyslu následovat minulost. To je u těchto metod jak záporem, tak i kladem. Ve většině případů je tento předpoklad rozumný, ale nemusí to platit úplně vždy. V průběhu let dochází k lékařským pokrokům, vzniku nových nemocí nebo k ekonomickým změnám, které mohou vést ke změnám v trendu úmrtnosti. To potom může mít za následek, že minulý trend nemusí být úplně nejlepším vodítkem pro budoucnost.

Metod založených na extrapolaci je velmi mnoho a díky malým požadavkům na data a nízké nákladnosti jsou také velmi často používané. Všechny takové metody mají společnou jednu věc, kterou je subjektivní rozhodnutí o době, po kterou je nutné sledovat historický trend. Jednoduché metody jsou spolehlivé do té míry, dokud budou mít podmínky, které vedly ke změnám měř úmrtnosti v minulosti, stejný vliv i v budoucnosti.

V této kapitole budu používat rozdělení extrapolačních metod, kterým jsem se inspirovala v práci [16], dělím je tedy na tři podskupiny: jednoduché extrapolační metody, podílové extrapolační metody a složité extrapolační metody. Jednoduché a podílové extrapolační metody zde ukážu pouze na populačních projekcích, protože je v této práci nebudu dále používat pro projekce úmrtnosti. Připomenu ještě, že na téma populačních projekcí jsem v minulé kapitole uvedla, že populaci lze dělit na tři složky: porodnost, úmrtnost a migraci, které se potom mohou dále projektovat zvlášť.

Ještě než přistoupím k popisu jednotlivých metod, uvedu zde pár pojmů, které využiji pro budoucí výklad:

- Základní rok je rok, ze kterého se berou nejstarší data použitá pro projekci, tedy rok, ve kterém se začínají sledovat historická data.
- Zahajovací rok je rok, ze kterého jsou k dispozici nejnovější data použitá pro projekci, tedy poslední rok, pro který jsou hodnoty ještě známé.
- Cílový rok je rok, pro který se projektuje.
- Základní období je počet let mezi základním a zahajovacím rokem, tedy počet let, pro které jsou hodnoty známé.
- Horizont projekce je počet let mezi zahajovacím a cílovým rokem, tedy počet let, pro které se projekce vytváří.
- Interval projekce je časový přírůstek, pro který se projekce vytváří (například jeden rok nebo pět let).

### 3.1 Jednoduché extrapolační metody

U jednoduchých extrapolačních metod se používají tři základní přístupy: lineární změna, geometrická změna a exponenciální změna.

### 3.1.1 Lineární změna

U tohoto přístupu se předpokládá, že se populace bude v daném období v budoucnosti měnit stejnou měrou jako tomu bylo v základním období. Následující vztah vyjadřuje průměrnou absolutní změnu během základního období:

$$\Delta = \frac{P_1 - P_b}{y}.$$

$\Delta$  je průměrná absolutní změna,  $P_1$  je velikost populace v zahajovacím roce,  $P_b$  velikost populace v základním roce a  $y$  je počet let základního období. Do cílového roku  $t$  se dále projektuje pomocí vztahu

$$P_t = P_1 + [z \cdot \Delta],$$

kde  $z$  je počet let horizontu projekce.

Prakticky se tento postup může použít například pro Prahu a pro období mezi lety 2004 a 2012 (hodnoty z Českého statistického úřadu):

$$\Delta = \frac{P_1 - P_b}{y} = \frac{1\,241\,664 - 1\,165\,581}{8} = 9\,510,375,$$

$$P_{2020} = 1\,241\,664 + [8 \cdot 9\,510,375] = 1\,317\,747.$$

### 3.1.2 Geometrická změna

Při tomto přístupu se namísto předpokladu stejné míry změny předpokládá stejná procentuální míra změny. Průměrná geometrická míra změny populace a následná projekce se počítají pomocí vztahů

$$r = \left(\frac{P_1}{P_b}\right)^{1/y} - 1,$$

$$P_t = P_1 (1 + r)^z.$$

Konkrétně pro Prahu a stejné období jako v předchozím příkladě jsou výsledky následující:

$$r = \left(\frac{1\,241\,664}{1\,165\,581}\right)^{1/8} - 1 \approx 7,935 \cdot 10^{-3},$$

$$P_{2020} = 1\,241\,664 \left(1 + 7,935 \cdot 10^{-3}\right)^8 = 1\,322\,713,298.$$

### 3.1.3 Exponenciální změna

Tento přístup úzce souvisí s přístupem geometrické změny, zde jde však o spojitou variantu. Vztahy vyjadřující průměrnou exponenciální míru změny populace během základního období a projekci jsou následující:

$$r = \frac{\ln \frac{P_1}{P_b}}{y},$$

$$P_t = P_1 e^{rz}.$$



## 3.2 Podílové extrapolační metody

Tyto metody se dají využít například v případě, kdy jsou již známé projekce pro nějaký větší celek a cílem je získat projekce také pro menší celky, ze kterých se tento větší celek skládá. Předpokládá se přitom, že dílčí části jsou disjunktní a dohromady tvoří daný celek. Pro podílové extrapolační metody se opět využívá tři přístupů: tzv. constant-share, shift-share a share-of-growth. Názvy ponechám v angličtině, protože takto lépe vystihují podstatu samotných přístupů, než by tomu bylo v případném českém překladu.

### 3.2.1 Constant-share

V tomto přístupu je podíl velikosti populace dílčí oblasti a celkové oblasti, pozorovaný během základního období, konstantní. Projekci do cílového roku  $t$  pro takovou dílčí část pak můžu snadno vypočítat jako

$$P_{it} = \frac{P_{il}}{P_{jl}} P_{jt},$$

kde  $P_{il}$ , respektive  $P_{jl}$ , je velikost populace menší, respektive celkové, oblasti v zahajovacím roce  $l$ ,  $P_{jt}$  je projekce do cílového roku  $t$  pro celkovou oblast.

### 3.2.2 Shift-share

U předchozího přístupu se předpokládá, že podíl velikosti populace dílčí části a celkové oblasti je konstantní. Tento přístup už počítá se změnami tohoto podílu. Pro vyjádření projekce populace dílčí oblasti můžu použít například výraz

$$P_{it} = P_{jt} \left[ \frac{P_{il}}{P_{jl}} + \frac{z}{y} \left( \frac{P_{il}}{P_{jl}} - \frac{P_{ib}}{P_{jb}} \right) \right],$$

kde  $z$  je počet let horizontu projekce,  $y$  počet let základního období a  $b$ , respektive  $l$ , je základní rok, respektive zahajovací rok.

### 3.2.3 Share-of-growth

Poslední přístup podílových extrapolacních metod se liší od předchozího přístupu v tom, že bere v úvahu spíše podíl změny populace, nežli velikosti populace. Zde se dá projekce dílčí oblasti vyjádřit následovně:

$$P_{it} = P_{il} + \left[ \frac{P_{il} - P_{ib}}{P_{jl} - P_{jb}} (P_{jt} - P_{jl}) \right].$$

## 3.3 Složitě extrapolační metody

Složitě extrapolacní metody mají oproti jednoduchým a podílovým extrapolacním metodám tu výhodu, že se lépe vyrovnávají s nelineární změnou populace. Odhady parametrů modelů jsou zde získávány pomocí statistických algoritmů, což dává větší základ pro měření nejistot předpovědí. I přes tyto výhody se ovšem

nemusí pomocí složitých extrapoláčních metod dojít k lepším výsledkům, než by tomu bylo při použití jednoduchých nebo podílových extrapoláčních metod.

Prvním krokem při tvorbě projekcí pomocí těchto metod je sběr historických dat pro různé datумы základního období. Když jsou data k dispozici, musejí se odhadnout parametry modelu zvoleného pro projekci. Nakonec se musí vytvořit samotná projekce. U složitých extrapoláčních metod se pro tvorbu předpovědí často využívá časových řad. Asi nejčastěji používanou metodou pro projekce je jednorozměrný ARIMA (Autoregressive Integrated Moving Average) model. V případě modelu s více parametry je ale vhodné brát v potaz také závislosti mezi jednotlivými parametry, což umožňuje vícerozměrný ARIMA model. Délka datových řad by měla být alespoň tak dlouhá, jako je dlouhý předpovídaný horizont. Ideální délka datových řad je však dvakrát až třikrát delší, než je délka předpovídaného horizontu.

Místo časových řad lze použít i jiné přístupy, například GLM (Generalized Linear Model). V tomto případě se pak u jednofaktorových modelů získá obyčejná regrese na čase, u dvoufaktorových modelů bivarianční regresní model.

V úvodu této kapitoly jsem uvedla, že extrapoláční metody lze dělit na tři skupiny: jednoduché, podílové a složité extrapoláční metody. Pro účely budoucího textu nyní zmíním ještě jedno možné dělení extrapoláčních metod, kterým jsem se inspirovala v literaturách [2] a [3]. Extrapoláční metody tedy rozdělím do skupin podle počtu faktorů, které, jak jsem již zmínila, mohou být až tři. Předchozí dvě skupiny extrapoláčních metod nemají faktor žádný a jsou tedy modelovány a projektovány přímo.

### 3.3.1 Jednofaktorové modely

Mezi tyto modely patří například Brassův relační model nebo parametrické modely. Časovou řadu zde tvoří vyrovnávané parametry modelu, kterých může být větší množství. Může se zde tedy objevovat dříve zmíněný problém se vzájemnou závislostí mezi jednotlivými parametry, což činí takové metody složitějšími.

#### Brassova relační metoda

William Brass přišel s nápadem nemodelovat vlastní průběhy intenzit úmrtnosti, ale modelovat spíše změny úmrtnosti, ke kterým v čase dochází. Zde nebude Brassovu metodu více popisovat, v případě zájmu si o ní čtenář může přečíst v [5].

#### Parametrické metody

Parametrické metody jsou založeny na prokládání historických dat parametrickou křivkou a následné projekci jejích parametrů. Nevýhodou těchto metod může být, že tato křivka nemusí budoucí úmrtnost popisovat dobře, protože nebude dostatečně flexibilní. Na druhou stranu, pokud bude křivka budoucnost popisovat dobře, tak zase hrozí, že bude mít příliš mnoho parametrů. Výhodou těchto metod je, že mají ve většině případů hladkou projekci měř. Velmi známým příkladem těchto metod je model Heligman-Pollarda, který má osm parametrů. Je popsán rovnicí

$$\frac{q_x}{p_x} = A^{(x+B)^C} + D \exp \left[ -E (\ln x - \ln F)^2 \right] + GH^x,$$

kde  $p_x = 1 - q_x$  a velká písmena od  $A$  do  $H$  jsou parametry modelu. Pravá strana rovnice je interpretována pomocí rozdělení na tři části. První z nich popisuje chování úmrtnosti v dětství, druhá v dospívání (zachycuje „nehodový hrb“ na křivce popisující úmrtnost okolo dvacátého roku života) a poslední ve stáří. Každá z těchto částí přitom nabývá kladných hodnot pouze pro příslušné věky.

Jako další parametrický model zmíním tzv. multiexponenciální model. Původně byl navržen pro modelování migrací a má devět parametrů. Stejně jako Heligman-Pollardův model, může být i tento rozdělen na několik částí, tentokrát na čtyři. První z nich je konstantní, druhá zachycuje exponenciální pokles úmrtnosti v dětství, třetí představuje dvojitý exponenciální nehodový hrbol a poslední částí je gompertzova úmrtnost ve stáří. Model se popisuje následující rovnicí:

$$m_x = a_0 + a_1 \exp(-\alpha_1 x) + a_2 \exp[-\alpha_2(x - \mu) - \exp[\lambda(x - \mu)]] + a_3 \exp(-\alpha_3 x),$$

kde  $m_x$  je centrální míra úmrtnosti ve věku  $x$ ,  $a_1$ ,  $a_2$  a  $a_3$  jsou úrovní parametrů a  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ ,  $\mu$  a  $\lambda$  definují úmrtnostní profil.

Pro zajímavost už jen zmíním několik dalších parametrických modelů. Například model A. Kostakiho má devět parametrů a je vhodnější pro nízké dospělé věky než model Heligman-Pollarda. Model navržený J. F. Carriérem má osm parametrů a je vhodnější na americká data a existuje i jeho varianta s jedenácti parametry.

### 3.3.2 Dvufaktorové modely

V těchto modelech se většinou za faktory bere věk a perioda, popřípadě věk a kohorta. V současnosti jsou tyto modely asi nejpoužívanější. Příkladem je Lee-Carterův model, používaný pro projekce úmrtností. Podrobněji tento model popíšu v dalších odstavcích.

#### Lee-Carterova metoda

Model Leea a Cartera je bilineárním modelem s věkově specifickými a časově specifickými proměnnými. Je definován vztahem

$$\ln(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t},$$

kde

- $m_{x,t}$  je centrální míra úmrtnosti pro věk  $x$  a kalendářní rok  $t$ ,
- $a_x$  je průměrný věkově specifický vzor úmrtnosti,
- $b_x$  ukazuje, které míry klesají rychle a které pomalu v závislosti na změnách  $k$ ,
- $k_t$  je ukazatel úrovně úmrtnosti v kalendářním roce  $t$ ,

- $\epsilon_{x,t}$  je náhodná chyba ve věku  $x$  a kalendářním roce  $t$ , vyjadřuje především věkově specifické historické výkyvy, které nebyly zachyceny modelem, má normální rozdělení s nulovou střední hodnotou a rozptylem  $\sigma_\epsilon^2$ .

Parametry  $b_x$  jsou většinou kladné, ale nemusí tomu tak být vždy. V případě, že úmrtnost v daných věcích roste, vycházejí  $b_x$  záporná. Pokud je  $k$  v čase lineární, znamená to, že se úmrtnost v každém věku mění stálou exponenciální měrou. Pokud jde  $k$  do záporného nekonečna, všechny věkově specifické míry jdou k nule.

Řešení rovnice není jednoznačné, proto jsou přidány dodatečné podmínky, které musejí parametry splňovat:

$$\sum_x b_x = 1, \sum_t k_t = 0.$$

Minimalizací výrazu

$$\mathcal{O}_{LS}(a, b, k) = \sum_{x=x_1}^{x_m} \sum_{t=t_1}^{t_n} (\ln \widehat{m}_{x,t} - a_x - b_x k_t)^2,$$

se získá odhad pro  $a_x$ :

$$\widehat{a}_x = \frac{1}{T} \sum_{t=1}^T \ln \widehat{m}_{x,t}. \quad (3.1)$$

V dalším kroku se musí odhadnout zbylé parametry  $b_x$  a  $k_t$ . Toho lze dosáhnout více způsoby.

První možností, která se může k odhadům parametrů použít, je singulární dekompozice matic (SVD). Při použití tohoto způsobu se vytvoří matice

$$Z_{x,t} = \ln(m_{x,t} - \widehat{a}_x) \quad (3.2)$$

a pomocí SVD se získá rozklad ve tvaru

$$SVD(Z_{x,t}) = \sum_{i=1}^r \rho_i U_{x,i} V_{i,t},$$

kde  $r$  je hodnota matice  $Z$ ,  $\rho_i$  jsou singulární hodnoty matice  $Z$ , pro něž platí uspořádání  $\rho_1 \geq \dots \geq \rho_r$ ,  $U_{x,i}$  a  $V_{i,t}$  jsou levé a pravé singulární vektory. Lee a Carter ve své metodě používají aproximaci pomocí SVD pouze prvního řádu, tedy  $\widehat{Z}_{x,t} = \rho_1 U_{x,1} V_{1,t} = b_x \cdot k_t$  a

$$\widehat{b}_x = U_{x,1}, \quad (3.3)$$

$$\widehat{k}_t = \rho_1 \cdot V_{1,t}. \quad (3.4)$$

Tímto však vzniká určitá chyba. Označím si  $\epsilon_t^{(p)}$  jako chybu, které se dopustím použitím  $p$  komponent v SVD, tedy

$$\epsilon_t^{(p)} = Z_t - (b_1 k_{1t} + b_2 k_{2t} + \dots + b_p k_{pt}).$$

Chybu aproximace potom získám jako podíl

$$\frac{\sum_t \|\epsilon_t^{(p)}\|^2}{\sum_t \|Z_t\|^2} = \frac{\sum_{i=p+1}^r \rho_i^2}{\sum_{i=1}^r \rho_i^2},$$

kde  $r$  je hodnost matice  $Z$ . Spíše než chyba aproximace by mohlo být zajímavější, jak moc je taková aproximace správná, tedy

$$\Delta E_p = 1 - \frac{\sum_{i=p+1}^r \rho_i^2}{\sum_{i=1}^r \rho_i^2} = \frac{\sum_{i=1}^p \rho_i^2}{\sum_{i=1}^r \rho_i^2}. \quad (3.5)$$

Získané parametry  $k_t$  však nejsou v souladu se skutečnými počty úmrtí, protože v Lee-Carterově metodě se modelují logaritmy měr úmrtnosti a nikoli přímo míry úmrtnosti samotné. Proto se doporučuje použít ještě druhý krok – rekalibraci, pro odhady parametrů  $k_t$ . Tyto odhady se získají řešením rovnice

$$\sum_x d_{x,t} = \sum_x E_{x,t} \exp(\hat{a}_x + \hat{b}_x \tilde{k}_t) \forall t, \quad (3.6)$$

kde  $d_{x,t}$  je počet úmrtí ve věku  $x$  v roce  $t$  a  $E_{x,t}$  je počet žijících ve věku  $x$  a roce  $t$ , tedy expozice riziku. Tato rovnice se nedá řešit analyticky, ale pouze numericky, pomocí některé ze známých iteračních metod.

Dalším krokem je předpověď modelu do budoucna. Předpokládá se, že parametry  $b_x$  a  $a_x$  jsou v čase konstantní, tudíž stačí předpovídat pouze index úmrtnosti  $k_t$ . K tomu se používá vhodného modelu ARIMA. V [9] došli Lee a Carter k tomu, že takovým vhodným modelem je v naprosté většině případů model náhodné procházky s driftem. Tento model je dán předpisem

$$k_t = k_{t-1} + \theta + \psi_t, \quad (3.7)$$

kde  $\theta$  je tzv. parametr driftu a  $\psi_t$  jsou vzájemně nezávislé náhodné veličiny normálního rozdělení s nulovou střední hodnotou a rozptylem  $\sigma_{rw}^2$ . Metodou maximální věrohodnosti se získá odhad pro driftovou složku i rozptyl chybové složky:

$$\hat{\theta} = \frac{k_T - k_1}{T - 1}, \quad (3.8)$$

$$\sigma_{rw}^2 = \frac{1}{T - 1} \sum_{t=1}^{T-1} (\hat{k}_{t+1} - \hat{k}_t - \hat{\theta})^2. \quad (3.9)$$

Pokud bych v předpisu pro náhodnou procházku s driftem chtěla posun o dva kroky místo jenom o jeden, dostávám předpis ve tvaru

$$\hat{k}_t = \hat{k}_{t-1} + \hat{\theta} + \psi_t = (\hat{k}_{t-2} + \hat{\theta} + \psi_{t-1}) + \hat{\theta} + \psi_t = \hat{k}_{t-2} + 2\hat{\theta} + (\psi_{t-1} + \psi_t).$$

Předpis pro  $k$  v čase  $T + l$  pro  $l > 0$  je tedy následující:

$$\hat{k}_{T+l} = k_T + l\hat{\theta} + \sum_{i=1}^l \psi_{T+i-1} = k_T + l\hat{\theta} + \sqrt{l}\psi_t. \quad (3.10)$$

Druhá rovnost vychází z předpokladu, že  $\psi_t$  jsou nezávislé náhodné veličiny se stejným rozptylem.

Odhad na základě tohoto modelu označím  $\mu_{T+l}$  a jeho tvar je

$$\mu_{T+l} = E \left[ \widehat{k}_{T+l} | k_1, \dots, k_T \right] = k_T + l\widehat{\theta}. \quad (3.11)$$

Projekce centrálních měr úmrtnosti do budoucnosti se pak už jednoduše vypočítá dosazením těchto odhadů do modelu, tedy

$$\widehat{m}_{x,T+l} = \widehat{a}_x + \widehat{b}_x \left( \widehat{k}_T + l\widehat{\theta} \right). \quad (3.12)$$

Jak jsem již zmínila výše, pro odhad parametrů lze použít i jiných postupů než SVD. Parametry lze získat také řešením soustavy rovnic, kterou jde získat parciálními derivacemi výrazu

$$\widetilde{O}_{LS}(b, k) = \sum_x \sum_t (\ln(m_{x,t}) - a_x - b_x k_t)^2$$

podle jednotlivých parametrů a jsou pokládány rovny nule:

$$0 = \sum_t (\ln(m_{x,t}) - a_x - b_x k_t),$$

$$0 = \sum_x b_x (\ln(m_{x,t}) - a_x - b_x k_t),$$

$$0 = \sum_t k_t (\ln(m_{x,t}) - a_x - b_x k_t).$$

Newton-Raphsonova metoda řeší tyto rovnice iteračním procesem založeným na rekurzivní formuli

$$\xi^{h+1} = \xi^h - \frac{f(\xi^h)}{f'(\xi^h)},$$

čímž se dojde k odhadům

$$\begin{aligned} \widehat{a}_x^{(h+1)} &= \widehat{a}_x^{(h)} + \frac{\sum_t (\ln(\widehat{m}_{x,t}) - \widehat{a}_x^{(h)} - \widehat{b}_x^{(h)} \widehat{k}_t^{(h)})}{T}, \\ \widehat{k}_t^{(h+1)} &= \widehat{k}_t^{(h)} + \frac{\sum_x \widehat{b}_x^{(h)} (\ln(\widehat{m}_{x,t}) - \widehat{a}_x^{(h+1)} - \widehat{b}_x^{(h)} \widehat{k}_t^{(h)})}{\sum_x (\widehat{b}_x^{(h)})^2}, \\ \widehat{b}_x^{(h+1)} &= \widehat{b}_x^{(h)} + \frac{\sum_t \widehat{k}_t^{(h+1)} (\ln(\widehat{m}_{x,t}) - \widehat{a}_x^{(h+1)} - \widehat{b}_x^{(h)} \widehat{k}_t^{(h+1)})}{\sum_t (\widehat{k}_t^{(h+1)})^2}. \end{aligned}$$

Dalším postupem je metoda vážených nejmenších čtverců, navržená J. R. Wilmothem. Zde se minimalizuje výraz

$$\sum_x \sum_t w_{x,t} (\ln(m_{x,t}) - a_x - b_x k_t)^2.$$

Za váhy  $w_{t,x}$  se mohou brát počty úmrtí  $d_{x,t}$  ve věku  $x$  a kalendářním roce  $t$ . Při této volbě nejde o váhy v pravém slova smyslu, jelikož se nesečtou na jedničku. Toho lze dosáhnout poměrně snadno volbou vah  $w_{t,x} = d_{x,t} / \sum_x \sum_t d_{x,t}$ .

Poslední možností, kterou zde zmíním, je použití metody maximální věrohodnosti. V tomto případě se předpokládá znalost počtů úmrtí  $D_{x,t}$  a expozičních rizik  $E_{x,t}$ , a že mají  $D_{x,t}$  Poissonovo rozdělení s parametrem  $E_{x,t} m_{x,t} =_{x,t} \exp(a_x + b_x k_t)$ . Odhady parametrů se pak získají maximalizací logaritmicke věrohodnostní funkce

$$\sum_x \sum_t [D_{x,t} \ln(E_{x,t} m_{x,t}) - E_{x,t} m_{x,t} - \ln(D_{x,t}!)]$$

a použitím Newton-Raphsonovy metody.

Posledním krokem je určení chybové složky metody. Už jsem ukázala, že předpověď logaritmu věkově specifických měr úmrtnosti o  $l$  let dopředu ze základního roku  $T$  jdou vypočítat na základě vztahu

$$\text{forecast} [\ln(m_{x,T+l})] = \hat{a}_x + \hat{k}_{T+l} \hat{b}_x, \quad (3.13)$$

stříškami jsou v případě  $a_x$  a  $b_x$  označeny odhadnuté hodnoty, v případě  $k_t$  značí stříška předpovězenou hodnotu. Každý parametr metody je však zatížen chybou, a tak je skutečná hodnota předpovědi následující:

$$\ln(m_{x,T+l}) = (\hat{a}_x + \alpha_x) + (\hat{k}_{T+l} + u_{T+l}) (\hat{b}_x + \beta_x) + \epsilon_{x,T+l}, \quad (3.14)$$

kde řecká písmena představují chybové složky jednotlivých parametrů, tedy  $\alpha_x$  a  $\beta_x$  jsou chyby vzniklé při odhadování  $a_x$  a  $b_x$ ,  $u_{x,T+l}$  je chybová složka vzniklá při předpovědi  $k$  o  $l$  let dopředu se základním rokem  $T$  a  $\epsilon_{x,T+l}$  je chyba modelu. Celková chyba se získá jako rozdíl rovnic 3.13 a 3.14 je tedy

$$E_{x,T+l} = \alpha_x + \epsilon_{x,T+l} + (\hat{b}_x + \beta_x) u_{T+l} + \beta_x \hat{k}_{T+l}$$

a její rozptyl je za předpokladu nezávislosti jednotlivých chyb

$$\sigma_{E_{x,T+l}}^2 = \sigma_{\epsilon_{x,T+l}}^2 + \sigma_{\alpha_x}^2 + \hat{b}_x^2 \sigma_{u_{T+l}}^2 + \sigma_{\beta_x}^2 (\hat{k}_{T+l}^2 + \sigma_{u_{T+l}}^2).$$

V článku [9] autoři popisují, že nejsou jasné korelace mezi zdroji jednotlivých chybových složek, čímž zdůvodňují předpoklad nezávislosti. Dále se v [9] tvrdí, že chybová složka vzniklá při předpovědi  $k_t$  do budoucna představuje kolem 95 % celkové chyby modelu, a to i v nejhorsím případě (v případě věkové skupiny 20–24 let). Na základě tohoto poznatku se i já v této práci omezím pouze na tuto část celkové chyby, tedy na určení hodnoty

$$\hat{b}_x^2 \sigma_{u_{T+l}}^2. \quad (3.15)$$

Odhad pro  $\sigma_{u_{T+l}}^2$  se získá ze vztahů 3.8 a 3.10:

$$\sigma_{u_{T+l}}^2 = l^2 \sigma_{\theta}^2 + l \sigma_{rw}^2,$$

kde  $\sigma_{rw}^2$  je vyjádřena již dříve (3.9) a  $\sigma_{\theta}^2$  se za předpokladu nezávislosti a normality chyb  $\psi_t$  určí ze vztahu 3.10:

$$\sigma_{\theta}^2 = \frac{\sigma_{rw}^2}{T-1}. \quad (3.16)$$

Nakonec je ještě možné určit interval spolehlivosti na hladině spolehlivosti 95 %. Ten je podle [1] za předpokladu normality celkové chyby  $E_{x,T+l}$  dán vztahem

$$\ln(\hat{m}_{x,T+l}) \pm \phi^{-1}(0,975) \hat{b}_x \sigma_{u_{T+l}}, \quad (3.17)$$

kde  $\phi$  je distribuční funkce rozdělení  $N(0,1)$ .

## 4. Jiné metody projekce úmrtnosti

### 4.1 Procesové metody

Tyto metody se většinou nepoužívají k samotným projekcím úmrtnosti, ale mohou být využity jako doplněk k extrapolacním metodám. Úmrtností se zabývají z bio-medicínského pohledu. Předpokládá se, že úmrtí je způsobeno nějakým nahromaděním defektů, které vedou k úmrtí a lze je matematicky popsat. K tomu jsou potřeba znalosti procesů vedoucích ke smrti a možnost jejich matematického modelování. Je možné, že se v budoucnu díky svým výhodám budou používat častěji, než je tomu dnes.

### 4.2 Výkladové metody

Výkladové metody se opět moc nepoužívají k oficiálním projekcím úmrtnosti, jelikož nejsou dostatečně známy vazby, které se v nich používají. Pracuje se zde s kauzálním nebo prognostickým přístupem a využívá se například ekonometrických technik založených na ekonomických proměnných nebo ekologických faktorech. Pokud jsou navíc výkladové proměnné předpověditelné stejně obtížně (nebo i více) jako závislé proměnné, vůbec se nevyplatí tyto metody zahrnovat do modelu.

V případě projekce úmrtnosti do budoucnosti se vytvářejí podtřídy, například podle příjmu nebo příslušnosti k socioekonomické třídě, což může být velice obtížné. Potřebné informace většinou nejsou k dispozici z úmrtních listů a je tedy nutné získat je z jiných zdrojů, například z medicínských studií.

Mnoho modelů používaných ve výkladových předpovědích je založeno na regresi nebo GLM.

### 4.3 Metody očekávání

Metody založené na očekávání používají data na základě očekávání jednotlivců nebo skupin odborníků ohledně chování těchto dat. Tato očekávání jsou založena na historických datech a průzkumech nebo podložených úsudcích odborníků v procesu predikce. U předpovědí úmrtnosti se tyto metody často používají právě ve formě expertního odhadu. Je dána předpověď nebo scénář a často máme také k dispozici dolní a horní odhady.

Nejběžnější použití těchto metod je pro porodnost, kde jsou očekávání porodů založena na zjištěných datech a následně použita pro tvorbu predikce. Problém těchto metod je, že zaostávají za aktuálním trendem. Výhodou je naopak možnost jejich použití v případě nedostatku historických dat.

Zastánci tohoto přístupu tvrdí, že velkou výhodou těchto metod v případě očekávání skupiny odborníků je možné zahrnutí budoucích strukturálních změn nebo neočekávaných událostí do modelu. Začlenit se dají také demografické, epidemiologické a další relevantní vědomosti, a to alespoň v kvalitativním smyslu.



Nevýhodou je potom možné zkreslení a subjektivita, čemuž se dá částečně předejít zprůměrováním úsudků více odborníků.

Jako zástupce těchto metod zmíním tzv. cílení očekávané délky života. Cílení je také používané pro redukční faktory věkově-specifické úmrtnosti. The Continuous Mortality Investigation Bureau (CMIB) navrhla poměr žijících v daném věku  $x$ , kteří zemřou před dožitím daného věku  $x + 1$ ,  $q_x$ , od roku 0 násobením  $t$ -letým redukčním faktorem  $RF(x, t)$ , kde

$$RF_x(t) = \alpha_x + [1 - \alpha_x] \cdot [1 - f_x(n)]^{\frac{t}{n}},$$

tedy

$$q_x(t) = g_x(t') \left[ \alpha_x + [1 - \alpha_x] \cdot [1 - f_x(n)]^{\frac{t-t'}{n}} \right],$$

$\alpha_x$  a  $f_x(n)$  jsou konečné redukční faktory. Předpokládá se, že poměr celkového poklesu  $(1 - \alpha_x)$  nastane za  $n$  let. Přístup ztělesňuje exponenciální pokles úmrtnosti v průběhu času na asymptotickou hodnotu  $\alpha_x$  a využívá expertního odhadu ke stanovení cílů  $\alpha_x$  a  $f_x(n)$ .

# 5. Rozdělení podle příčin úmrtí

## 5.1 Vývoj hlavních příčin úmrtí ve 20. století

Ve dvacátém století došlo k velkým pokrokům v dlouhověkosti v mnoha vyspělých zemích, jak už jsem ostatně uvedla hned v úvodu práce. Na začátku století došlo k významnému poklesu nemocí dýchací soustavy, později měl hlavní vliv na klesající úmrtnost pokles počtu nemocí oběhové soustavy. Následující poznatky vychází podle [13] z dat European Mortality Database, která poskytuje data podle 67 příčin úmrtí, národnosti, pohlaví a věkových skupin. Pro většinu zemí jsou zde data dostupná pro roky 1980–2007. The European Public Health Alliance v roce 2006 zkoumala vedoucí příčiny úmrtí v EU a dospěla k závěru, že 41 % všech úmrtí byla způsobena chorobami oběhové soustavy a 25 % rakovinami. Nyní se zaměřím na jednotlivé věkové skupiny a dále na vývoj podle příčin úmrtí ve vybraných zemích a na rizikové faktory podle příčin úmrtí, stejně jak je popisováno v [13]:

- Věky 20–44: sebevraždy představují 12 % úmrtí, přičemž u mužů je až čtyřikrát vyšší pravděpodobnost spáchání sebevraždy, než u žen.
- Věky 45–64: rakoviny představují 41 % úmrtí, nejčastější jsou u mužů rakoviny plic a jícnu, u žen rakoviny prsu.
- Věky 65–84: 42 % úmrtí nastává kvůli chorobám oběhové soustavy.

### 5.1.1 Vývoj v Anglii a Walesu

Vývoj zde nebyl pravidelný, liší se různě jak podle věkových skupin a pohlaví, tak podle času. Faktory ovlivňující změnu, pozitivní i negativní, zahrnují vývoj ve zdravotnictví, změny v chování obyvatel, práci, sociální politiku, bohatství a války. V posledních čtyřiceti letech bylo v UK hlavním tahounem velkého poklesu úmrtnosti snížení úmrtí na následky chorob oběhové soustavy, a to okolo 60 % u mužů i žen. Úmrtnost vlivem rakovin postupně rostla u obou pohlaví až do začátku osmdesátých let, kdy nabyla vrcholu a od devadesátých let klesá. U mužů také došlo k významnému zlepšení v úmrtnosti na choroby dýchací soustavy. V období od roku 1979 do roku 2003 byl vývoj v jednotlivých věkových skupinách následující:

- Novorozenecké míry úmrtnosti se snížily téměř o 60 %, zde nejsou známy příčiny úmrtí.
- Věky 1–14: úmrtnost se snížila u chlapců o 57 %, u dívek o 50 %. Téměř polovina těchto úmrtí byla způsobena zraněními a otravami.
- Věky 15–44: otravy a zranění (zatím nejvyšší zastoupení příčin úmrtí) u mužů klesají jen pozvolně, u žen klesá počet úmrtí způsobených rakovinami (největší zabiják) o 45 %.

- Věky 45–64: v této skupině byly na začátku období největším zabijákem mužů nemoci oběhové soustavy. Později tento počet klesl asi o dvě třetiny a největším zabijákem se tak staly rakoviny, i když jejich počet klesl o 34 %. Pro ženy čítají téměř polovinu všech příčin úmrtí rakoviny, třetinu nemoci oběhové soustavy. Obě příčiny v tomto období klesly o 29 % pro rakoviny a o 65 % pro nemoci oběhové soustavy.
- Věky 65–84: věkově standardizované míry úmrtnosti klesly o 40 % u mužů a o třetinu u žen. Pro obě pohlaví byly hlavní příčinou úmrtí nemoci oběhové soustavy. Druhým největším zabijákem byly rakoviny, jejichž počet se v průběhu období snížil u mužů, ale u žen nikoliv.
- Věky nad 85: největšími zabijáky jsou zde nemoci oběhové a dýchací soustavy. Obě příčiny během období výrazně poklesly.

### 5.1.2 Vývoj v Japonsku

Jak již bylo řečeno v úvodu, Japonsko je země s nejvyšší očekávanou dobou života. Je proto zajímavé sledovat trendy v úmrtnosti podle příčin úmrtí. Období 1950–2000 vykazuje významný pokles ve věkově-standardizovaných mírách úmrtnosti, zejména v letech 1970–1990, kdy došlo k ohromnému poklesu cerebrovaskulárních chorob. Dalšími významnými ukazateli jsou také pokles srdečních chorob a na začátku období snížení výskytu tuberkulózy.

### 5.1.3 Vývoj v USA

S výjimkou malého růstu v roce 1993 zde úmrtnost vlivem srdečních onemocnění nepřetržitě klesá od roku 1980. Věkově upravené míry úmrtnosti zapříčiněné rakovinami, druhou nejčastější příčinou úmrtí, vykazují postupný, ale konzistentní sestupný trend od roku 1993. Míry pro cévní příhodu mozkovou obecně klesají od roku 1958 s výjimkou let 1992–1995. Vypozorovalo se, že úmrtnostní trend pro Alzheimerovu chorobu prudce rostl, ale část tohoto zvyšování byla zapříčiněna vývojem diagnóz, povědomím lékařské obce a také změnou v kódování z ICD-9 na ICD-10.

## 5.2 Rizikové faktory podle příčin úmrtí

### 5.2.1 Ischemická choroba srdeční a cévní příhoda mozková

U cévní příhody mozkové je nejvýznamnějším rizikovým faktorem suboptimální systolický krevní tlak, nižší než 115 mmHg. Méně významnými faktory jsou potom vysoký cholesterol a příjem malého množství ovoce a zeleniny. U ischemických chorob srdečních je největším faktorem vysoký cholesterol, těsně za ním suboptimální systolický krevní tlak. Dalšími rizikovými faktory v uvedeném pořadí podle významnosti jsou příjem malého množství ovoce a zeleniny a málo pohybu. S. Capewell provedl analýzu a dospěl k závěru, že ve Skotsku, Anglii a Walesu, Irsku, Novém Zélandu, Finsku a USA můžeme přisoudit 45–75 % poklesu úmrtnosti právě změně v rizikových faktorech, zbývajících 25–55 % změně ve zdravotnictví.

## 5.2.2 Rakovina

Klíčovými faktory, kterým se dá vyhnout, jsou podle WHO (World Health Organization) tyto:

- Užívání tabákových výrobků – to je zodpovědné až za 1,8 milionu výskytů rakovin za rok.
- Nadváha, obezita a nedostatek pohybu dohromady způsobí až 274 000 úmrtí na rakovinu ročně.
- Užívání škodlivého množství alkoholu může až za 351 000 rakovin ročně.
- Sexuálně přenosný lidský papillomavirus, který způsobuje až 235 000 rakovin za rok.
- Karcinogeny, které mohou nejméně za 152 000 rakovin ročně.
- Citlivost na určitý druh rakoviny způsobená genetickými predispozicemi.

## 5.2.3 Nemoci dýchací soustavy: chronická obstrukční plicní nemoc (COPD)

Hlavní příčinou COPD je užívání tabákových výrobků. V dnešní době jsou postiženy ženy i muži téměř stejnou měrou, a to především zvýšením počtu ženských kuřáček v bohatých zemích. Dalšími významnými rizikovými faktory jsou vystavení prachu a chemikáliím, příslušnost k nižší socio-ekonomické třídě nebo genetické predispozice. Tato choroba není léčitelná, nicméně léčba může postup nemoci velmi zpomalit. Lidé trpící COPD často umírají na jinou chorobu. Podle studií provedených v Anglii a Walesu je přibližně 60 % úmrtí přisouzeno COPD, 25 % nemocím oběhové soustavy a 5 % rakovinám.

## 5.2.4 Diabetes

Diabetes prvního i druhého typu jsou zabijáky samy o sobě. Předpokládá se však, že jiné nemoci způsobené diabetem významně snižují skutečný počet úmrtí přisouzených této chorobě. Diabetes je také rizikovým faktorem pro kardiovaskulární choroby i celkovou úmrtnost. Pánové Franco, Oscar, Steverberg, Ewout, Hu, Frank, Mackenbach, Johan, Nusselder a Wilma v roce 2007 počítali souvislost mezi očekávanou dobou života a onemocněním na diabetes. K dispozici měli skupinu diabetiků starších padesáti let a podobnou skupinu netrpící diabetem. Přišli na to, že jedinci ve skupině diabetiků žili průměrně o 7,5 až 8,2 let kratší dobu než jedinci z druhé skupiny. WHO předpokládá, že se v letech 2005 až 2030 počet diabetiků zdvojnásobí a předkládá jednoduchá životní opatření, která se v prevenci nebo alespoň oddálení nástupu diabetu druhého typu ukázala jako efektivní:

- Dosažení a udržení zdravé tělesné váhy.
- Být fyzicky aktivní, doporučeno je nejméně 30 minut pravidelné aktivity střední intenzity po většinu dní.

- Zdravě jíst, doporučuje se příjem tří až pěti porcí ovoce a zeleniny denně, omezení příjmu cukru a nasycených tuků.
- Vyhýbat se tabákovým výrobkům, kouření zvyšuje i riziko kardiovaskulárních onemocnění.

### 5.3 Projekce podle příčin úmrtí

Všechny metody, které jsem dosud zmínila, je možné pro projekce úmrtnosti použít buď pro souhrnná data nebo oddělená, například podle příčin úmrtí. V tomto případě musíme úmrtí rozdělit podle příčin do několika hlavních skupin, projektovat zvlášť míry úmrtnosti pro každou skupinu a výsledné míry nějakým způsobem opět zkombinovat k vytvoření souhrnné míry úmrtnosti.

Projekce podle příčin úmrtí mají například tu výhodu, že poskytují dobrý náhled na způsob, jakým se úmrtnost v čase mění, nemusí to však být pravda úplně vždy. Jedním z největších problémů je naopak fakt, že úmrtí z určitých příčin většinou nejsou nezávislá. Příkladem může být jev, který se stal na začátku 20. století, kdy se sice snížil počet infekčních chorob, ale místo toho se zvýšil počet úmrtí na rakovinu a srdeční onemocnění. Dalším problémem je určení konečné příčiny úmrtí. Smrt může nastat v důsledku kombinace několika různých chorob. Právě ta, která je zaznamenána jako ta hlavní, nemusí být zdaleka jediná a dokonce ani ne ta skutečná. U starých lidí je určení konečné příčiny úmrtí zvláště obtížné. Další omezení vychází z dostupných časových řad pro data rozdělených podle příčin. Ta jsou většinou pro použití ARIMA modelu pro předpovědi příliš krátká. Parametry, které se v modelu musí predikovat, jsou v případě projekcí pro jednotlivé příčiny většinou méně stabilní, než pro celkovou úmrtnost.

# 6. Projekce úmrtnosti podle příčin úmrtí pro českou populaci

## 6.1 Data

K projekci úmrtnosti pro českou populaci budu používat Lee-Carterovu metodu, kterou jsem podrobně popsala ve třetí kapitole. Pro vytvoření projekce úmrtnosti podle příčin úmrtí je nejprve potřeba sestavit úmrtnostní tabulky podle jednotlivých příčin úmrtí, pohlaví a věku pro období 1950–2010. Počty úmrtí podle příčin, pohlaví a věkových skupin jsou dostupné na webových stránkách ČSÚ (Český statistický úřad), pro účely práce však vyžadují další úpravu.

### 6.1.1 Skupiny příčin úmrtí

Kvůli cílům diplomové práce jsem podrobné příčiny úmrtí rozdělila do šesti základních skupin – rakoviny (rak), nemoci oběhové soustavy (oběh), nemoci dýchací soustavy (dých), nemoci trávicí soustavy (tráv), vnější příčiny úmrtí (vněj) a ostatní příčiny úmrtí (ost). Do skupiny ostatních příčin patří například nemoci nervové soustavy nebo nemoci močové a pohlavní soustavy, u žen do ostatních příčin patří také těhotenství, porod a šestinedělí. V tabulce 6.1 jsem pro představu uvedla některé zástupce od každé skupiny nemocí.

Podrobné příčiny úmrtí jsou značeny podle Mezinárodní klasifikace nemocí a souvisejících zdravotních problémů (v České republice se používá zkratka MKN, v originále zkratka ICD – International Classification of Diseases and Related Health Problems). Aktuální verze této klasifikace se během období let 1950–2010 několikrát změnila. Tabulka 6.2 znázorňuje, ve kterých letech se používala jaká verze mezinárodní klasifikace nemocí (MKN) a tomu odpovídající označení ICD.

I když se z důvodu změn v klasifikaci mění označení jednotlivých skupin, jsou do nich přidávány nové nemoci nebo naopak některé odebírány, na zvolené skupiny to ve většině případů nemělo žádný vliv, jelikož změny probíhaly pouze v rámci jednotlivých skupin. Rakoviny například po celé období 1950–2010 spadají do klasifikační skupiny II – Novotvary. Naopak u skupiny nemocí oběhové soustavy došlo během tohoto období hned k několika změnám. V roce 1968 došlo k přejmenování skupiny z Nemoci ústrojí cévního na Nemoci ústrojí oběhového a v roce 1979 k dalšímu přejmenování na Nemoci oběhové soustavy. Tento název se používá i v současnosti. Další změna nastala v roce 1994, kdy se změnilo číslo klasifikační skupiny ze VII na IX. V tomto roce ostatně došlo ke změně označení v klasifikaci u více skupin. U nemocí trávicí soustavy se přešlo z čísla IX na XI, u nemocí dýchací soustavy z VIII na X a u vnějších příčin úmrtí z XVII na XX. Ve skupině ostatních příčin úmrtí jsou pak zařazeny zbylé klasifikační skupiny od I do XVI, od roku 1994 až do XVIII.

K výraznějším změnám v klasifikaci, které mají vliv i na mnou zvolené skupiny, došlo pouze v roce 1968. V důsledku vývoje medicíny v oblasti nervových a oběhových nemocí došlo právě v tomto roce k přesunu některých nemocí ze skupiny nemocí nervového ústrojí (v mém případě zařazených do skupiny ostatních chorob) do skupiny nemocí oběhové soustavy. Konkrétně šlo o následující nemoci:

<i>Skupina</i>	<i>Příklady</i>
Rakoviny	Zhoubný novotvar tlustého střeva Zhoubný novotvar prsu (u žen) Zhoubný novotvar průdušky – bronchu a plíce
Nemoci oběhové soustavy	Chronická ischemická nemoc srdeční Mozkový infarkt (mozková mrtvice) Ateroskleróza
Nemoci trávicí soustavy	Alkoholické onemocnění jater Fibróza a cirhóza jater Vaskulární onemocnění střeva
Nemoci dýchací soustavy	Pneumonie Bakteriální zánět plic Jiná chronická obstruktivní plicní nemoc
Vnější příčiny úmrtí	Pády Dopravní nehody Úmyslné sebepoškození
Ostatní příčiny úmrtí	Diabetes Chronické selhání ledvin Alzheimerova nemoc

Tabulka 6.1: Příklady nemocí pro jednotlivé skupiny příčin úmrtí.

mozkové krvácení, subarachnoidální krvácení<sup>1</sup>, mozkový vmetek a záhať<sup>2</sup>, jiné a špatně určitelné cévní poruchy centrálního nervstva, phlebitis a thrombophlebitis<sup>3</sup> a spasmus mozkových tepen<sup>4</sup>. K dosažení konzistence dat jsem až do roku 1968 včetně přesunula tyto nemoci ze skupiny nervových onemocnění do skupiny nemocí oběhové soustavy. V případě, že by se tato úprava dat neudělala, došlo by v roce 1968 k výraznému poklesu nemocí nervové soustavy na úkor velkého růstu ve skupině oběhových chorob.

### 6.1.2 Věkové skupiny

Data získaná od ČSÚ, která jsou dělena podle příčin úmrtí, jsou dále dělena pouze podle věkových skupin. Pro věky po jednom roce jsou data dělena pouze v letech 1950–1993 a to pouze pro věky 0–4. V letech 1950–1959 jsou dále data dělena do desetiletých věkových intervalů a poslední věkovou skupinou jsou věky nad 85 včetně. V letech 1960–1993 jsou data dělena obdobně, ale do pětiletých věkových intervalů. Od roku 1994 jsou data dostupná pro věk 0 a dále pro pětileté

<sup>1</sup>O subarachnoidálním krvácením se mluví, jestliže se krev vylije do prostoru mezi pavučnicí a tuhou plenu, což jsou obaly, ve kterých je uložen mozek. Tento prostor se nazývá subarachnoidální prostor.

<sup>2</sup>Mozkový vmetek je jinými slovy embol, neboli cizí těleso v cévě, které může vést k mozkové mrtvici. Mozková záhať je pak přímo výraz pro mozkovou mrtvici.

<sup>3</sup>Phlebitis znamená zánět žil, thrombophlebitis je zánět žil doprovázený srážením krve.

<sup>4</sup>Spasmus znamená křeč nebo křečový stah.

<i>Období</i>	<i>MKN</i>	<i>ICD</i>
1950–1957	MKN 1948	ICD-6
1958–1967	MKN 1955	ICD-7
1968–1978	MKN 1966	ICD-8
1979–1993	MKN 1976	ICD-9
1994–2010	MKN 1993	ICD-10

Tabulka 6.2: Verze MKN v jednotlivých obdobích v rozmezí let 1950–2010.

věkové intervaly končící skupinou pro věky nad 95 včetně. Je však potřeba sestavit úmrtnostní tabulky pro věky po jednom roce. Toho docílím použitím vztahu

$$q_{x,t}^{(P)} = q_{x,t}^{(C)} \frac{D_{S,t}^{(P)}}{D_{S,t}^{(C)}},$$

zvlášť pro muže a ženy, kde  $x = 0, \dots, 103$ ,  $t = 1950, \dots, 2010$ ,  $P = 1, \dots, 6$ ,

- $q_{x,t}^{(P)}$  je pravděpodobnost úmrtí z příčiny  $P$  ve věku  $x$  v roce  $t$ ,
- $q_{x,t}^{(C)}$  je celková pravděpodobnost úmrtí ve věku  $x$  v roce  $t$ , získaná z celkových úmrtnostních tabulek ČSÚ (tyto hodnoty jsou známé i pro věky po jednom roce),
- $D_{S,t}^{(P)}$  je počet úmrtí z příčiny  $P$  ve věkové skupině  $S$  v roce  $t$ ,
- $D_{S,t}^{(C)}$  je celkový počet úmrtí ve věkové skupině  $S$  a roce  $t$  (opět převzato z dat ČSÚ, oproti součtu  $D_{S,t}^{(P)}$  přes jednotlivé příčiny se tyto hodnoty liší výjimečně a minimálně, případné rozdíly jsou způsobené nenulovými počty úmrtí z neurčené příčiny).

Abych mohla celkové pravděpodobnosti úmrtí rozdělit na pravděpodobnosti úmrtí podle jednotlivých příčin takto jednoduše, je potřeba předpokládat nezávislost jednotlivých příčin úmrtí. Tento předpoklad budu využívat i při pozdější agregaci, kde vysvětlím, proč je možné ho používat. Takovýmto dělením dat pro věkové skupiny na jednotlivé věky však dochází v průběhu přes věky ke skokům v okrajových věcích skupin. Skoky by bylo možné odstranit vhodnou vyhlazovací metodou, to by však ovlivnilo budoucí projekci. Z tohoto důvodu jsem data nechala nevyhlazená.

### 6.1.3 Nulové hodnoty

Počty úmrtí jsou pro některé příčiny v některých letech a věcích nulové. To vede i k nulovým pravděpodobnostem úmrtí a centrálním měrám úmrtnosti, které budu později logaritmovat. Nulové hodnoty pravděpodobností úmrtí jsem v případě nenulových hodnot v okolních kalendářních letech nahradila hodnotami získanými lineární interpolací. V případě nulových hodnot na konci našeho kalendářního období jsem nahrazovala hodnotou  $10^{-7}$ , protože při použití lineární extrapolace vycházely záporné hodnoty.



## 6.2 Aplikace Lee-Carterovy metody

Než přistoupím k samotným výpočtům, krátce zopakuji, jak je Lee-Carterův model definován. Model je dán vztahem

$$\ln(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t},$$

kde  $m_{x,t}$  jsou centrální míry úmrtnosti,  $a_x$  a  $b_x$  jsou věkově specifické parametry modelu, které je třeba odhadnout,  $k_t$  je index úmrtnosti, který je taktéž potřeba odhadnout a  $\epsilon_{x,t}$  je chybová složka modelu. Podrobnější popis metody obsahuje třetí kapitola, na kterou se budu i nadále odkazovat při jednotlivých výpočtech.

Nejprve si vypočítám centrální míry úmrtnosti podle vztahu

$$q_{x,t}^{(P)} = 1 - \exp^{-m_{x,t}^{(P)}}$$

používaného ČSÚ. Následně spočtu logaritmy centrálních měr úmrtnosti  $\ln(m_{x,t}^{(P)})$  pro ženy i muže,  $x = 0, \dots, 103$ ,  $t = 1950, \dots, 2010$  a  $P = 1, \dots, 6$ .

Poté je třeba odhadnout parametry modelu  $a_x^{(P)}$ ,  $b_x^{(P)}$  a  $k_t^{(P)}$ . Jelikož je každou operaci nutno provést pro muže i ženy a každou příčinu úmrtí zvlášť, nebudu nadále psát index (P) ani výslovně psát informaci o pohlaví, kromě případů, ve kterých by mohlo dojít k nejasnostem.

### 6.2.1 Věkově-specifické parametry

Parametr  $a_x$  lze jednoduše vypočítat podle vztahu 3.1. Tento parametr představuje průměrný věkový profil úmrtnosti.

Parametr  $b_x$  získám singulárním rozkladem matice 3.2 a použitím vztahu 3.3. Aby řešení bylo jednoznačné, musí se ještě splnit podmínka jednotkového součtu  $b_x$  přes věky  $x = 0, \dots, 103$ . Toho docílím normalizací ve tvaru

$$\hat{b}_x = \frac{U_{x,1}}{\sum_{i=0}^{103} U_{i,1}},$$

kde  $U_{x,1}$  je levý singulární vektor získaný singulárním rozkladem.

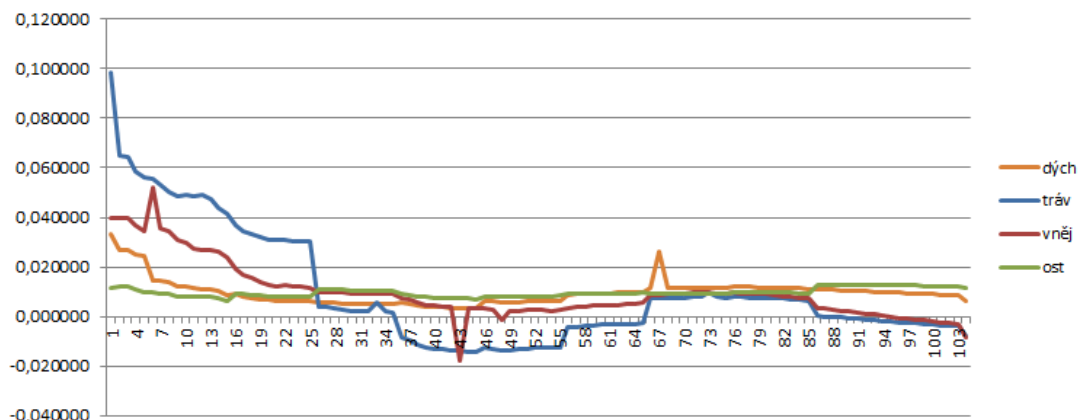
Použitím rozkladu prvního řádu však vzniká chyba, jak je popsáno ve třetí kapitole. Jak moc je tato aproximace správná vypočítám podle vztahu 3.5. Výsledné hodnoty jsem uvedla v tabulce 6.3.

$\Delta E_1$	<i>Rak.</i>	<i>Oběh.</i>	<i>Tráv.</i>	<i>Dých.</i>	<i>Vněj.</i>	<i>Ost.</i>
Muži	43,2 %	71,8 %	72,2 %	49,1 %	45,5 %	85,0 %
Ženy	20,7 %	58,9 %	47,9 %	74,8 %	62,3 %	85,3 %

Tabulka 6.3: Správnost aproximace prvním řádem SVD pro muže, ženy a jednotlivé příčiny úmrtí.

Jak ukazuje graf 6.1, pro českou populaci vychází parametry  $b_x$  pro některé věky záporné. To značí skutečnost, že úmrtnost v těchto věcích roste, narozdíl od okolních kladných hodnot, kde úmrtnost klesá. Dále graf pro  $b_x$  poukazuje na to, že mužská úmrtnost z ostatních příčin se v žádném věku výrazně nesnižuje. Naproti tomu úmrtnost zapříčiněná trávicími chorobami se přes věky viditelně

mění. Tvar této křivky vypovídá o tom, že v nízkých věcích se úmrtnost výrazně snížila, ve věcích 35–64 už však úmrtnost mírně roste a do věku 86 opět mírně klesá. Na obrázku je také patrný výskyt „hrbolů“ pro některé příčiny a věky. To jsou chyby způsobené použitím pouze prvního řádu SVD.



Obrázek 6.1: Parametr  $b_x$  přes věky  $x$  pro muže podle vybraných příčin úmrtí.

## 6.2.2 Odhad indexu úmrtnosti

Singulárním rozkladem, pomocí kterého jsem získala odhady parametru  $b_x$ , získám i odhady indexu úmrtnosti  $k_t$ . Pro výpočet použiji vztah 3.4. Jak jsem již uvedla ve třetí kapitole, tento odhad neodpovídá skutečnému počtu úmrtí a je vhodné index úmrtnosti rekalibrovat. K tomu použiji vztah 3.6, taktéž popsaný ve třetí kapitole. Rovnice jsem řešila pomocí zabudovaných funkcí v programu Wolfram Mathematica. V případě přibližně poloviny skupin nebyl s řešením žádný problém – konkrétně šlo o úmrtí na nemoci oběhové soustavy, dýchací soustavy a úmrtí z vnějších příčin u mužů a úmrtí na nemoci dýchací soustavy, trávicí soustavy, úmrtí z vnějších a ostatních příčin u žen. V ostatních případech (tedy u úmrtí na rakoviny, na nemoci trávicí soustavy a úmrtí z ostatních příčin u mužů a u úmrtí na rakoviny a nemoci oběhové soustavy u žen) nebylo v některých kalendářních letech nalezeno reálné řešení zmíněné rovnice, z toho u ostatních příčin u mužů a u nemocí oběhové soustavy u žen nebylo nalezeno reálné řešení pouze v jednom kalendářním roce. Podrobnější informace jsou uvedeny v tabulkách 6.4 a 6.5.

Ty hodnoty  $k_t$ , které v problémových kalendářních letech  $t$  vyšly pouze komplexní, je nutné nahradit reálnými hodnotami. Vyzkoušela jsem dva postupy, které připadají v úvahu:

1. Nahrazení komplexních hodnot  $k_t$  hodnotami minimalizující výraz

$$\left\| \sum_x d_{x,t} - \sum_x E_{x,t} \exp(\hat{a}_x + \hat{b}_x \tilde{k}_t) \right\|. \quad (6.1)$$

2. Náhrada komplexních hodnot  $k_t$  těmi, které jsme získali v prvním kroku odhadu metodou SVD.

<i>Příčina úmrtí – muži</i>	<i>Kalendářní rok</i>
Rakoviny	1950–1960, 2005–2010
Nemoci trávicí soustavy	1955, 1956, 1958–1963, 1968, 1993, 1996–2010
Ostatní příčiny	1996

Tabulka 6.4: Kalendářní roky výskytů nereálných řešení rovnice 3.6 u mužů.

<i>Příčina úmrtí – ženy</i>	<i>Kalendářní rok</i>
Rakoviny	2001, 2004–2010
Nemoci oběhové soustavy	2007

Tabulka 6.5: Kalendářní roky výskytů nereálných řešení rovnice 3.6 u žen.

Ze zmíněných dvou možností náhrady komplexních  $k_t$  jsem se rozhodla pro první z nich, s výjimkou skupiny rakovin u mužů. Proč jsem v tomto případě volila jiný přístup ještě vysvětlím v následující podkapitole. V případě, že již mám reálné hodnoty indexů úmrtnosti pro všechny kalendářní roky, musím je ještě upravit tak, aby splňovaly podmínku nulového součtu přes kalendářní roky  $t$ . Toho docílím odečtením průměru hodnot  $k_t$  od každého z  $k_t$ ,  $t = 1950, \dots, 2010$ . Konečný výsledek získám jako

$$\hat{k}'_t = \hat{k}_t - \bar{\bar{k}},$$

kde

$$\bar{\bar{k}} = \frac{1}{T} \sum_t \hat{k}_t.$$

Kvůli normalizaci indexů úmrtnosti nyní musím upravit i hodnoty parametrů  $a_x$ . Po dosazení nových hodnot indexů úmrtnosti do Lee-Carterova modelu dostávám nové hodnoty i pro parametr  $a_x$ :

$$\ln(\widehat{m}_{x,t}) = \hat{a}_x + \hat{b}_x \hat{k}_t + \epsilon_{x,t} = \hat{a}_x + \hat{b}_x (\hat{k}'_t + \bar{\bar{k}}) + \epsilon_{x,t} = \hat{a}'_x + \hat{b}_x \hat{k}'_t + \epsilon_{x,t},$$

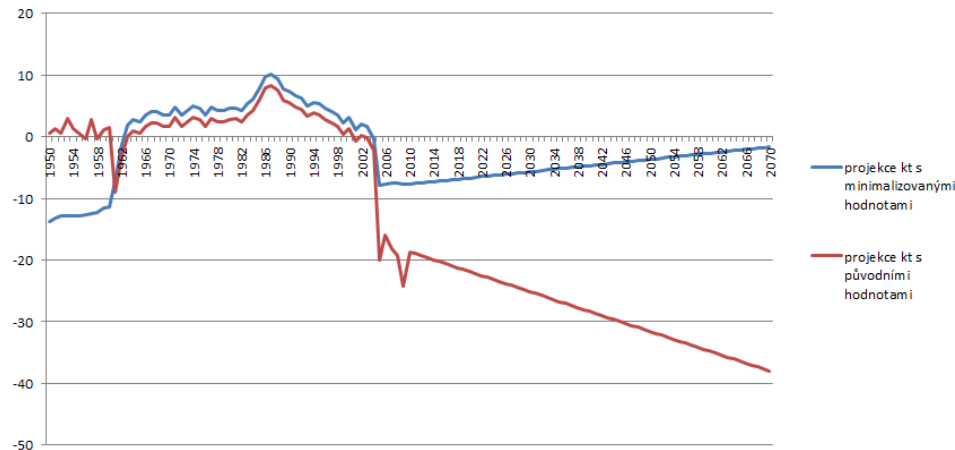
kde

$$\hat{a}'_x = \hat{a}_x + \hat{b}_x \bar{\bar{k}}.$$

### 6.2.3 Projekce indexů úmrtnosti

Výše jsem řešila problém výskytu nereálných indexů úmrtnosti a jejich možnému nahrazení dvěma různými způsoby. K tomu, abych mohla lépe rozhodnout, který z přístupů vybrat pro samotnou projekci, a které indexy úmrtnosti nakonec použít, sestrojila jsem projekci indexů úmrtnosti pro oba případy. U naprosté většiny skupin příčin úmrtí neměl výběr z těchto dvou možností na projekci indexů úmrtnosti výrazný vliv, a proto jsem volila spíše náhradu nereálných indexů úmrtnosti hodnotami minimalizujícími 6.1 (první možnost). Výjimkou však byla skupina mužů a rakovin, u které jsem nakonec zvolila náhradu nereálných indexů

úmrtnosti hodnotami indexů získaných metodou SVD (druhá možnost). Obrázek 6.2 ukazuje, že u této skupiny je rozdíl ve volbě z možných náhrad za indexy úmrtnosti značný a vede dokonce k diametrálně odlišné projekci. Zatímco při použití první z možností vychází projekční křivka rostoucí, při volbě druhé možnosti je křivka klesající. U žen je v obou případech projekční křivka klesající, a jelikož není důvod, kvůli kterému by tyto projekce měly být výrazně jiné u mužů a u žen, volila jsem u mužů – rakovin druhou možnost náhrady.



Obrázek 6.2: Projekce indexů úmrtnosti  $k_t$  pro muže a rakoviny s náhradou ne-reálných indexů oběma zmíněnými možnostmi.

Nyní, když jsem vybrala, se kterými indexy úmrtnosti je pro česká data vhodnější pracovat, ukážu, jak jsem ke konkrétním projekcím došla. Podle termínů, které jsem zadefinovala ve třetí kapitole beru

- základní rok = 1950,
- zahajovací rok = 2010,
- cílový rok = 2070,
- základní období = 61 let (1950–2010),
- horizont projekce = 60 let (2011–2070) a
- interval projekce = 1 rok.

K samotné projekci využiji časových řad, konkrétně modelu  $ARIMA(0, 1, 0)$  – náhodné procházky s driftem (viz 3.7). V tomto modelu je nejprve nutné odhadnout drift vztahem 3.8. Výsledné hodnoty pro jednotlivé skupiny jsou vypsány v tabulce 6.6. Projekci indexů úmrtnosti na 60 let dopředu poté určím za pomoci vztahu 3.11.

#### 6.2.4 Projekce centrálních měř úmrtnosti

Konečně mohu přistoupit k výpočtu samotných projekcí centrálních měř úmrtnosti. Ty vypočítám na základě vztahu 3.12, stačí znát hodnotu indexu úmrtnosti

	<i>Muži</i>	<i>Ženy</i>
Rakoviny	-0,3222	-0,4446
Nemoci oběhové soustavy	-3,1792	-4,1417
Nemoci dýchací soustavy	-1,5862	-2,3283
Nemoci trávicí soustavy	-0,7404	-2,1536
Vnější příčiny	-1,1922	-1,0177
Ostatní příčiny	-2,9579	-3,2968

Tabulka 6.6: Výsledné odhady driftů z modelu náhodné procházky pro jednotlivé příčiny úmrtí, ženy a muže.

pro poslední známý rok (2010) a odhad driftového parametru z modelu náhodné procházky s driftem. Projekci opět provádím pro každý věk  $x = 0, \dots, 103$ , obě pohlaví, každou ze šesti skupin příčin úmrtí, a to až do roku 2070 včetně. V uvedeném vztahu beru  $l = 1, \dots, 60$ .

V této fázi výpočtu provedu další úpravu. Získané křivky projekcí centrálních měr úmrtnosti zbavím „hrbolů“, o kterých jsem se zmínila již v odstavci pojednávajícím o odhadu parametru  $b_x$ . Hodnoty  $\widehat{m}_{x,t}$  pro věky  $x$ , ve kterých se hrboly vyskytují, nahradím průměrem hodnot z okolních věků (tedy průměrem hodnot  $\widehat{m}_{x-1,t}$  a  $\widehat{m}_{x+1,t}$ ). Tato úprava je důležitá především z praktického hlediska. Pokud by například z takových dat vycházela nějaká pojišťovna, měla by ve věcích s odlehlými hodnotami měr úmrtnosti stanovit odlišnou sazbu pojistného než pro bezprostředně okolní věky. To by však prakticky nedávalo smysl. V následujících odstavcích se budu věnovat získaným výsledkům z projekcí.

### Vývoj centrálních měr úmrtnosti v průběhu věků

Ve všech kalendářních letech jsem z dat vyzobrazovala, že centrální míry úmrtnosti vykazují rostoucí trend s přibývajícím věkem pro všechny příčiny úmrtí. Za zmínku stojí „nehodový hrb“, který se vyskytuje pouze u vnějších příčin úmrtí a je patrný ve všech letech přibližně mezi věky 15–35 u mužů, méně výrazný je pak u žen přibližně mezi věky 15–25. Centrální míry úmrtnosti z této příčiny pak u obou pohlaví stagnují až do vyšších věků, kde míry opět rostou. Například u dvacetiletých mužů by mělo být v roce 2030 vnějšími příčinami způsobeno 72 % úmrtí, ve věku 50 let by toto zastoupení mělo být 17 % a ve věku 80 už pouze 3 %. Zastoupení se tedy s rostoucím věkem snižuje, ačkoli míry jako takové neklesají s rostoucím věkem po celou dobu. To je způsobeno tím, že ačkoli ve vysokých věcích míry úmrtnosti na vnější příčiny rostou, jejich růst není tak výrazný jako například u nemocí oběhové soustavy a rakovin.

Zajímavou skutečností je také nulová úmrtnost na nemoci oběhové soustavy přibližně mezi věky 5–13 u mužů v letech 2008–2010, která významně ovlivní následné projektované hodnoty centrálních měr úmrtnosti. Ty jsou v celém projekčním období ve zmíněných věcích také nulové. U žen se takový jev vyskytuje pouze pro věk 0.

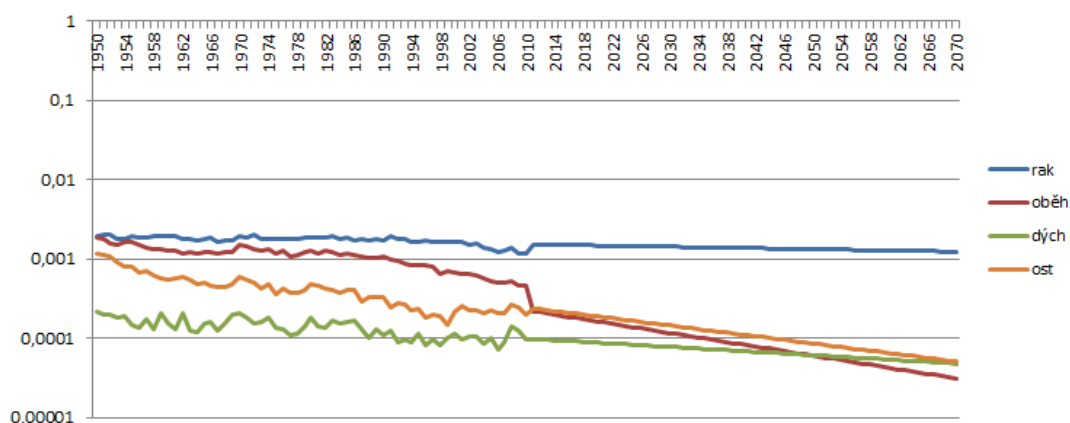
U mužů dále platí, že přibližně v období 1950–2010 ve velmi nízkých věcích převažují úmrtí z ostatních příčin, poté až do středního věku převažují vnější příčiny úmrtí (nehodový hrb) a poté úmrtí na nemoci oběhové soustavy. V pozdějších

letech však přibližně ve věcích 40–70 převažují úmrtí na rakoviny. U žen je vývoj podobný, ale rakoviny převažují nad ostatními příčinami úmrtí už kolem roku 1990 a v trochu větším věkovém intervalu, než-li je tomu u mužů.

U většiny příčin, kalendářních let a věků platí, že je u mužů vyšší úmrtnost než u žen. V roce 1950 jsou však úmrtnosti pro obě pohlaví u všech příčin velmi podobné a rozdíly se vyskytují spíše až v pozdějších kalendářních letech. U rakovin jsou například míry úmrtnosti u žen a u mužů velmi podobné pro téměř všechny kalendářní roky přibližně až do věku 40, poté jsou míry úmrtnosti vyšší pro muže. Míry úmrtnosti vychází pro ženy vyšší než pro muže například pro nemoci oběhové soustavy ve velmi nízkých a pak ve velmi vysokých věcích, podobně je tomu u vnějších příčin úmrtí. U nemocí trávicí soustavy jsou míry pro ženy vyšší ve velmi vysokých věcích.

### Vývoj centrálních měr úmrtnosti v průběhu kalendářních let

U žen centrální míry úmrtnosti pro všechny příčiny úmrtí v průběhu kalendářních let klesají až do velmi vysokého věku. Kolem věku 85 se pak objevuje rostoucí trend pro skupinu rakovin, později také pro nemoci oběhové soustavy a v posledních pozorovaných věcích ještě pro nemoci trávicí soustavy. Nejvýraznější pokles jsem vypočetla především pro skupiny ostatních příčin úmrtí a pro nemoci oběhové soustavy v nižších a středních věcích.



Obrázek 6.3: Vývoj centrálních měr úmrtnosti pro ženy ve věku 50 let pro vybrané příčiny úmrtí.

U mužů je projektovaný vývoj centrálních měr úmrtnosti pro jednotlivé skupiny příčin úmrtí různorodější. Rostoucí trend se vyskytuje už kolem věku 35, a to pro skupinu nemocí trávicí soustavy. Tento trend se mění v klesající až ve vysokých věcích, kdy naopak vykazuje rostoucí trend skupina rakovin. V posledním uvažovaném věku 103 již vykazují rostoucí trend všechny skupiny, kromě nemocí dýchací soustavy a ostatních příčin úmrtí.

Jelikož míry úmrtnosti pro rakoviny začínají ve vysokých věcích růst u mužů dříve než u žen, rostou u mužů projektované míry úmrtnosti trochu rychleji než u žen a lze tedy v budoucnosti očekávat i rostoucí rozdíl mezi úmrtností žen a mužů na rakoviny. Podobně je tomu například i u oběhových nemocí. Existenci takového jevu nasvědčuje i procentuální zastoupení měr úmrtnosti z těchto dvou příčin v celkových měřích úmrtnosti. Například v roce 1950 bylo 12 % všech

úmrtí osmdesátiletých mužů způsobeno rakovinami. Do roku 2070 by tato hodnota měla narůst na 47 %. U stejně starých žen bylo v roce 1950 10 % všech úmrtí způsobeno rakovinami, v roce 2070 už by to však mělo být dokonce 53 %. V základním roce se zastoupení úmrtí na rakoviny pro obě pohlaví příliš neliší (rozdíl je pouze 2 %), v cílovém roce je už rozdíl o něco vyšší (6 %). Důvodem pro takové zvyšování zastoupení u obou pohlaví může být to, že míry úmrtnosti z této příčiny v čase rostou, narozdíl od všech ostatních příčin, které klesají. Zastoupení úmrtnosti na nemoci oběhové soustavy v celkové úmrtnosti šedesátiletých mužů a žen se narozdíl od zastoupení rakovin nezvyšuje, ale zmenšuje, ačkoli rozdíl mezi pohlavími opět roste. V roce 1950 bylo 41 % všech úmrtí šedesátiletých mužů způsobeno nemocemi oběhové soustavy, v roce 2070 by tato hodnota měla být 22 %. U žen jsou tyto hodnoty 42 % pro rok 1950 a už pouze 4 % pro rok 2070.

U obou pohlaví jsou ve velmi nízkých věcích nejvyšší centrální míry úmrtnosti pro skupinu ostatních příčin. Ve věku 0 a v roce 2070 by dokonce zastoupení měr úmrtnosti z ostatních příčin v celkových měrách úmrtnosti mělo být pro muže 97 % a u žen 91 %. U mužů jsou pak kolem 20. roku života výrazně vyšší míry pro skupiny vnějších příčin, u žen tento rozdíl není tak výrazný, přesto pozorovatelný. U skupiny vnějších příčin úmrtí je opět zajímavé procentuální zastoupení. Zatímco v roce 1950 bylo 55 % všech úmrtí u dvacetiletých mužů způsobeno právě vnějšími příčinami, v roce 2070 už by to mělo být 72 %. U žen tyto hodnoty vycházejí na 22 % v roce 1950 a 74 % v roce 2070, je zde tedy ještě větší nárůst v zastoupení. Přesto platí, že křivka centrálních měr úmrtnosti v průběhu kalendářních let pro obě pohlaví v daném věku vykazuje klesající trend. To znamená, že ačkoli by v budoucnu mělo dojít ke zlepšení úmrtnosti z této příčiny, toto zlepšení není tak výrazné jako u jiných skupin příčin úmrtí. U žen od středního do vysokého věku pak převažuje skupina rakovin, ve vysokých věcích je to pak skupina nemocí oběhových soustav. U mužů jsou míry pro rakoviny nejvyšší pouze v projekčním období a pro menší věkový interval, převažující skupinou jsou spíše nemoci oběhových soustav.

### 6.2.5 Chybová složka

Chybovou složku Lee-Carterovy metody pro data české populace vypočítám podle vztahu 3.15 ze třetí kapitoly. Výpočet provedu opět pro obě pohlaví, pro každý věk, pro každý kalendářní rok z projekčního období a pro každou skupinu příčin úmrtí zvlášť. Takto získám velký počet hodnot, proto zde ukážu jen některé z nich. Ze získaných hodnot chybových složek mohu ještě určit konfidenční intervaly pomocí vztahu 3.17.

V tabulkách 6.7, resp. 6.8, jsou uvedeny hodnoty  $\sigma_{\theta}^2$  a  $\sigma_{rw}^2$ , určené podle vztahů 3.16 a 3.9, pro každou příčinu úmrtí a pro muže, resp. pro ženy.

Podle tabulek 6.7 a 6.8 jsou nejvyšší hodnoty  $\sigma_{\theta}^2$  a  $\sigma_{rw}^2$  pro nemoci oběhové soustavy a pro nemoci dýchací soustavy, a to u obou pohlaví. To je způsobené především tím, že křivka indexů úmrtnosti přes kalendářní roky je u těchto příčin více skokovitá než u ostatních příčin. To znamená, že rozdíly  $k_t - k_{t-1}$ , pomocí kterých se určují hodnoty rozptylů, jsou pro některá  $t$  poměrně vysoké.

<i>Muži</i>	$\sigma_\theta^2$	$\sigma_{rw}^2$
Rakoviny	0,170	10,186
Nemoci oběhové soustavy	2,556	153,354
Nemoci dýchací soustavy	1,524	91,464
Nemoci trávicí soustavy	0,776	46,547
Vnější příčiny	0,482	28,923
Ostatní příčiny	0,616	36,931

Tabulka 6.7: Hodnoty  $\sigma_\theta^2$  a  $\sigma_{rw}^2$  pro české muže a jednotlivé příčiny úmrtí.

<i>Ženy</i>	$\sigma_\theta^2$	$\sigma_{rw}^2$
Rakoviny	0,192	11,539
Nemoci oběhové soustavy	1,692	101,502
Nemoci dýchací soustavy	3,171	190,261
Nemoci trávicí soustavy	0,877	52,5992
Vnější příčiny	0,522	31,317
Ostatní příčiny	0,813	48,808

Tabulka 6.8: Hodnoty  $\sigma_\theta^2$  a  $\sigma_{rw}^2$  pro české ženy a jednotlivé příčiny úmrtí.

### 6.2.6 Agregace

Jedním z cílů diplomové práce je porovnání získaných projekcí centrálních měr úmrtnosti podle příčin úmrtí s projekcí celkových centrálních měr úmrtnosti. Abych toto mohla udělat, potřebuji ještě vypočítat projekce celkových měr úmrtnosti a seskupit projekce měr úmrtnosti podle jednotlivých příčin úmrtí opět do jedné celkové míry. Ty nakonec budou děleny pouze podle věku a pohlaví. Všechny výpočty provedu velmi podobně jako v předchozím případě, proto už nebudu jednotlivé kroky tolik popisovat.

#### Projekce celkových měr úmrtnosti

Projekci celkových měr úmrtnosti budu vytvářet stejně jako v předchozím případě pro věky  $0, \dots, 103$ , kalendářní roky 2010–2070 pro ženy a muže zvlášť. Data pro kalendářní roky 1950–2010 jsou k dispozici na webových stránkách ČSÚ.

K projekci opět využívám Lee-Carterovu metodu. Musím určit parametr  $a_x$  a pomocí metody SVD také parametry  $b_x$  a  $k_t$ . Správnosti aproximace pomocí prvního řádu SVD v procentech pro muže ( $\Delta E_m$ ) a ženy ( $\Delta E_z$ ) jsou:

$$\Delta E_m = 52,3772,$$

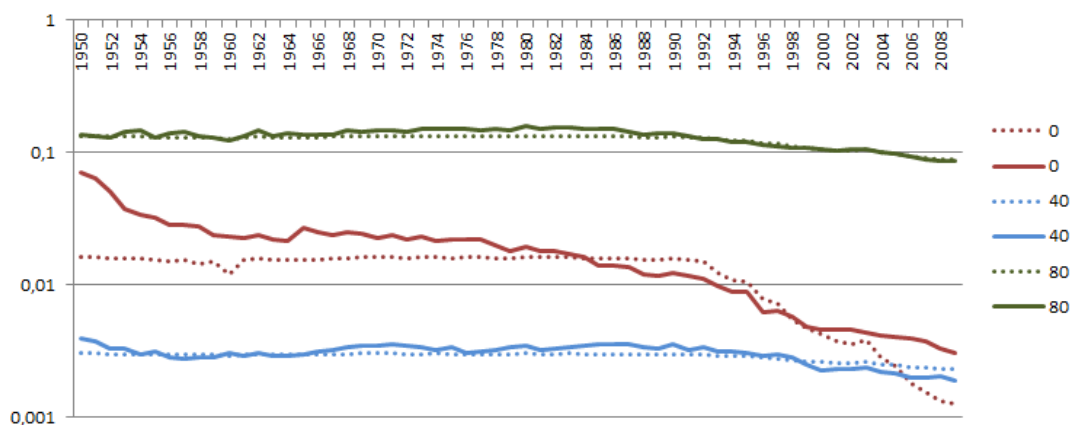
$$\Delta E_z = 57,5077.$$

K určení indexu úmrtnosti, který odpovídá skutečnému počtu úmrtí využiji vztahu 3.6.

Obrázek 6.4 zachycuje, jak je modelování českých dat pomocí Lee-Carterovy metody kvalitní. Porovnám známé hodnoty měr úmrtnosti z let 1950–2010 s měrami úmrtnosti, které vypočítám pomocí právě odhadnutých parametrů Lee-Car-



terova modelu. Na grafu jsou znázorněny hodnoty vypočítané pomocí modelu tečkovanou čarou, známé hodnoty měř plnou čarou. Z grafu lze usoudit, že model není nejpřesnější ve velmi nízkých věcích, ale v pozdějších věcích křivku známých měř aproximuje dobře. To může být způsobeno například tím, že ve velmi nízkých věcích nedochází k velkému počtu úmrtí. Není tedy k dispozici velké množství dat a modelování pomocí Lee-Carterovy metody v těchto věcích není nejlepší.



Obrázek 6.4: Srovnání známých centrálních měř úmrtnosti (plná čára) s hodnotami vypočítanými pomocí Lee-Carterova modelu (tečkovaná čára) ve vybraných věcích pro muže v logaritmickém měřítku.

K projekci indexů úmrtnosti do budoucna opět použiji model ARIMA(0,1,0) s následujícími odhady driftu  $-\theta_m$  pro muže a  $\theta_z$  pro ženy:

$$\theta_m = -3,3931,$$

$$\theta_z = -3,0620.$$

### Samotná agregace

Nyní je zapotřebí získat celkové centrální míry úmrtnosti  $m_{x,t}$  pro  $x = 0, \dots, 103$ ,  $t = 2011, \dots, 2070$  a pro ženy i muže z měř  $m_{x,t}^P$ , které už mám k dispozici pro  $x = 0, \dots, 103$ ,  $t = 2011, \dots, 2070$ .  $P = 1, \dots, 6$  a pro obě pohlaví. Takové celkové centrální míry úmrtnosti pro kalendářní roky 1950,  $\dots$ , 2010 jsou známé od samého počátku. Agregované míry úmrtnosti získám jednoduše jako součet přes jednotlivé příčiny úmrtí, tedy jako

$$m_{x,t} = \sum_{P=1}^6 m_{x,t}^P.$$

Zde jsem opět předpokládala nezávislost jednotlivých příčin úmrtí, kterou jsem zmínila již na začátku této kapitoly. V kapitole 5.3 popisují úskalí projekcí podle příčin úmrtí. Jedním z nich je určení konečné příčiny úmrtí, které je obtížné zejména při úmrtích ve vysokých věcích. Mně však stačí předpokládat nezávislost pouze mnou zvolených skupin příčin úmrtí. Tyto skupiny jsou tak odlišné, že je jen malá pravděpodobnost, že by příčina úmrtí označená jako konečná, patřila do jiné skupiny, než skutečná příčina úmrtí.

Nyní mohu porovnat hodnoty projektovaných měr úmrtnosti z celkových dat s hodnotami projektovaných měr úmrtnosti podle příčin. Obecně u obou pohlaví platí, že míry podle příčin jsou vyšší než celkové míry, přičemž největší rozdíly lze pozorovat v nízkých (do 10 let) a ve velmi vysokých věcích (od 90 let). To může být opět způsobeno tím, že Lee-Carterova metoda v těchto věcích není tak přesná. U žen jsou mezi měrami větší rozdíly než u mužů ve všech věcích. Zatímco u žen se míry úmrtnosti podle příčin k celkovým měrám výrazně nepřibližují v žádném věku, u mužů jsou ve věcích kolem 30–40 obě míry hodně podobné. Z pozorování plyne, že tyto dva způsoby, kterými lze k projekcím měr úmrtnosti dojít, nejsou ekvivalentní a vedou k různým výsledkům pro většinu věků a kalendářních let, přičemž projekce podle příčin úmrtí vede k vyšším hodnotám měr úmrtnosti.

### 6.2.7 Střední doba života

Dalším ukazatelem, který lze pozorovat, je střední délka života. Střední délka života  $e_x$  osoby ve věku  $x$  je střední hodnota zbývající doby života osoby ve věku  $x$   $T_x$ . Platí, že  $T_x = K_x + S_x$ , kde

- $K_x$  je celočíselná zbývající doba života osoby ve věku  $x$ , neboli počet celých let, které ještě prožije osoba ve věku  $x$ ,
- $S_x$  je část roku, po kterou bude osoba ve věku  $x$  naživu v roce úmrtí,  $0 \leq S_x \leq 1$ .

Pak

$$e_x = ET_x = EK_x + ES_x \approx \sum_{k=1}^{\infty} {}_k p_x + \frac{1}{2},$$

kde  ${}_k p_x$  je  $k$ -letá pravděpodobnost přežití osoby ve věku  $x$ , neboli pravděpodobnost toho, že osoba, která je naživu ve věku  $x$ , se dožije věku  $x + k$  a platí

$${}_k p_x = {}_1 p_x \cdot {}_1 p_{x+1} \cdot \dots \cdot {}_1 p_{x+k-1}.$$

K výpočtu dále použiji vztah, který jsem zmínila už na začátku kapitoly:

$$q_{x,t} = 1 - \exp^{-m_{x,t}}$$

a vztah

$$p_x = 1 - q_x,$$

kde  $p_x = {}_1 p_x$ .

Kvůli pozdějšímu srovnání vypočítám střední doby života jak pro celkové míry úmrtnosti, tak pro míry úmrtnosti agregované z měr rozdělených podle příčin úmrtí. Střední doby života by se samozřejmě daly počítat i pro míry úmrtnosti rozdělené podle příčin úmrtí, nicméně by to nemělo velký interpretační smysl.

V obou případech mají ve většině věků vyšší střední dobu života ženy než muži. Kolem věku 90 se střední doby života pro obě pohlaví poměrně vyrovnávají a v ještě pozdějších věcích mají naopak střední dobu života vyšší muži než ženy. Tabulky 6.9 a 6.10 obsahují střední doby života získané z měr úmrtnosti agregovaných z měr rozdělených podle jednotlivých příčin a jejich projekcí.

<i>Muži</i>	1950	1990	2010	2030	2070
$e_0$	62,34	67,59	74,37	75,24	77,34
$e_{30}$	39,59	39,45	45,27	45,94	47,66
$e_{65}$	11,95	11,61	15,26	16,00	17,30

Tabulka 6.9: Střední doby života získané z agregovaných měr úmrtnosti rozdělených podle příčin úmrtí a jejich projekcí pro vybrané kalendářní roky a věky pro muže, zaokrouhlené na dvě desetinná místa.

<i>Ženy</i>	1950	1990	2010	2030	2070
$e_0$	67,02	75,42	80,60	82,44	84,58
$e_{30}$	43,01	46,65	51,14	52,91	54,80
$e_{65}$	13,25	15,24	18,71	20,27	21,79

Tabulka 6.10: Střední doby života získané z agregovaných měr úmrtnosti rozdělených podle příčin úmrtí a jejich projekcí pro vybrané kalendářní roky a věky pro ženy, zaokrouhlené na dvě desetinná místa.

Pro muže se střední doba života ve věku 0 od roku 1950 do roku 2010 zvýšila o 12,03 let, u žen o 13,58 let. Do roku 2070 by pak mělo dojít k dalšímu zlepšení o necelé 3 roky, u žen o necelé 4 roky. Velikost takových zlepšení se s přibývajícím věkem zmenšuje.

Získané hodnoty pro střední doby života vypočítané z celkových měr úmrtnosti a jejich projekcí jsou k nahlédnutí v tabulkách 6.11 a 6.12. Pro období 1950–2010 jsou hodnoty středních dob života shodné jako v předchozích tabulkách, jelikož tyto hodnoty jsou známé. Proto je už nebudu znovu uvádět a místo toho uvádím více hodnot z projekčního období 2011–2070.

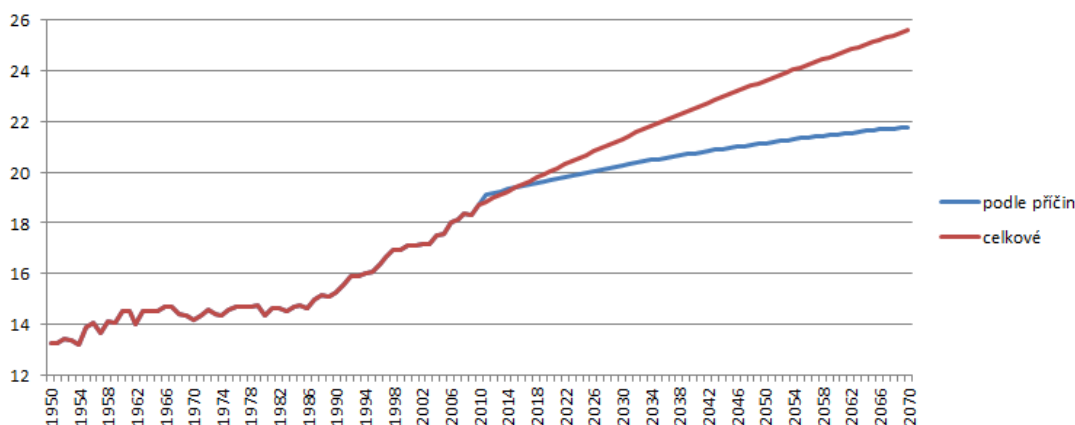
<i>Muži</i>	2011	2030	2050	2070
$e_0$	74,29	76,35	78,47	80,58
$e_{30}$	44,84	46,74	48,74	55,77
$e_{65}$	15,48	17,04	18,74	20,49

Tabulka 6.11: Střední doby života získané z celkových měr úmrtnosti a jejich projekcí pro vybrané kalendářní roky a věky pro muže, zaokrouhlené na dvě desetinná místa.

Ze všech čtyř tabulek je patrné, že hodnoty středních dob života získaných z celkových měr úmrtnosti jsou vyšší než ty, které jsem získala z měr úmrtnosti agregovaných z měr rozdělených podle příčin úmrtí. Zlepšení v období 2010–2070 by v tomto případě mělo být u mužů ve věku 0 o 6,15, což je ještě o 3 roky více než v předchozím případě. U žen je zde zlepšení o 9,14, to je dokonce o 5 let více než v předchozím případě. Zde dokonce dochází ke zlepšování i ve vysokých věcích. Pro názornost uvádím obrázek 6.5, který zachycuje srovnání zlepšování celkových středních dob života a středních dob života podle příčin v čase pro ženy ve věku 65 let. Více hodnot pro střední doby života je uvedeno v příloze C.

Ženy	2011	2030	2050	2070
$e_0$	81,12	84,47	87,36	89,74
$e_{30}$	51,45	54,59	57,41	59,76
$e_{65}$	18,84	21,32	23,62	25,59

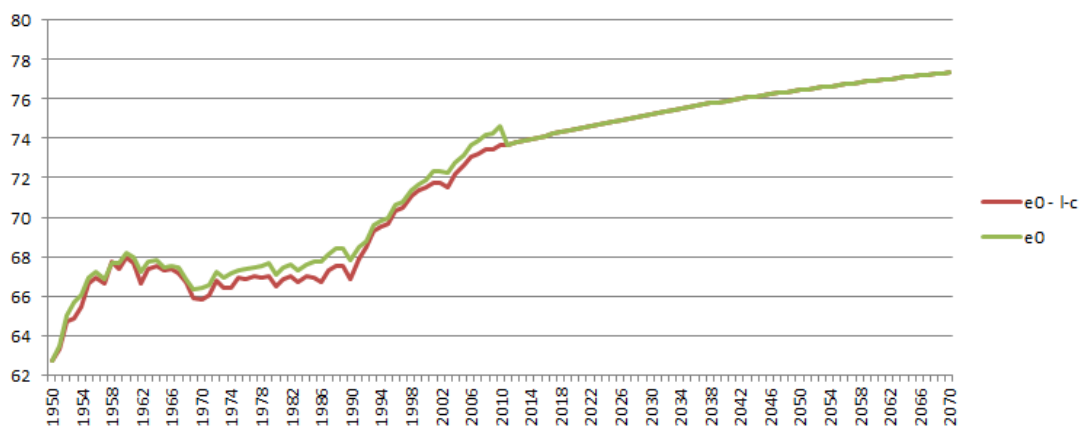
Tabulka 6.12: Střední doby života získané z celkových měr úmrtnosti a jejich projekcí pro vybrané kalendářní roky a věky pro ženy, zaokrouhlené na dvě desetinná místa.



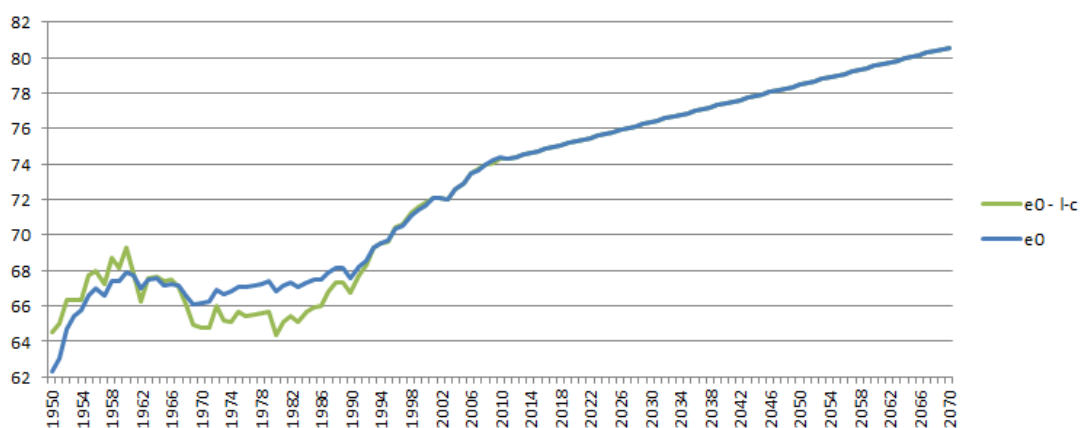
Obrázek 6.5: Srovnání průběhu celkových středních dob života a středních dob života podle příčin úmrtí pro ženy ve věku 65 let.

I u středních dob života je možné srovnat hodnoty, které jsou známé (období 1950–2010) s těmi, které bychom získali Lee-Carterovou metodou a zjistit tak, jak moc je Lee-Carterova metoda přesná. Nejdříve srovnám střední doby života získané z centrálních měr úmrtnosti agregovaných z měr rozdělených podle příčin úmrtí. U žen ve věku 0 je aproximace Lee-Carterovou metodou poměrně přesná, jak je tomu u mužů znázorňuje obrázek 6.6. Zelená křivka v grafu ukazuje vývoj středních dob života získaných pomocí aproximace Lee-Carterovou metodou, červená vystihuje vývoj středních dob života vypočítaných přímo z agregovaných měr úmrtnosti. Křivka získaná pomocí Lee-Cartera je nižší než křivka získaná ze známých hodnot a v roce 2011 tak dochází k propadu. Lee-Carterova metoda v tomto případě vede k pesimističtějším výsledkům, než jaké ve skutečnosti jsou, a jaké by v budoucnosti mohly být.

Nyní ještě stejným způsobem srovnám střední doby života získané z celkových centrálních měr úmrtnosti. I zde je aproximace u žen dobrá, pro muže opět přidávám obrázek 6.7. Aproximované hodnoty Lee-Carterovou metodou se od těch známých hodnot liší nejvíce o tři roky a v roce 2010 nedochází k žádným skokům, projektované hodnoty na známé hladce navazují.



Obrázek 6.6: Srovnání středních dob života získaných pomocí aproximace metodou Lee-Cartera a získaných přímo centrálních měr úmrtnosti podle příčin úmrtí pro muže ve věku 0.



Obrázek 6.7: Srovnání středních dob života získaných pomocí aproximace metodou Lee-Cartera a získaných přímo z celkových centrálních měr úmrtnosti pro muže ve věku 0.

# Závěr

V práci se nejdříve zabývám tím, jak se úmrtnost v posledním století změnila ve velkých zemích mimo Českou republiku a tím, proč se vlastně projekce úmrtnosti používají čím dál více a z jakých důvodů se vůbec používat začaly. Je jasné, že mnoho z těchto poznatků platí i pro Českou republiku. Například problémy s prodlužující se délkou života a stárnutím obyvatelstva a snahy o jejich řešení jsou vidět na nedávném pokusu o zavedení důchodové reformy. Ta má přimět obyvatele k tomu, aby nespolehali pouze na státní důchod, ale šetřili si na svůj život v důchodovém věku i soukromě. Projekce úmrtnosti jsou tedy velmi aktuální i pro českou populaci.

V dalších kapitolách se zabývám popisem některých projekčních metod, nejvíce jsem se zaměřila na Lee-Carterovu metodu. Tuto metodu jsem si vybrala pro vytvoření projekcí úmrtnosti jak podle příčin úmrtí tak celkových úmrtností pro česká data, protože je jednoduchá na pochopení, interpretaci i na samotné modelování vzhledem k nízkému počtu parametrů a absenci parametrů vnějších. Díky tomu všemu je ostatně tato metoda velmi rozšířená a oblíbená i v zahraničí.

K dispozici jsem měla data pro období kalendářních let 1950–2010, tedy 61 let. Všechny projekce jsem vytvářela pro období 2011–2070, tedy na 60 let do budoucnosti. Horizont projekce a základní období jsou přibližně stejně veliké, a získané projekce by neměly být příliš znehodnoceny malým počtem známých dat, ze kterých se vychází. V tomto ohledu by se projekce daly vylepšit buď prodloužením horizontu projekce nebo zkrácením základního období. Pro česká data ale není ideální prodloužovat horizont projekce, neboť pro období let 1938–1944 ČSÚ data nemá k dispozici z důvodu druhé světové války. Horizont projekce by tedy šel prodloužit maximálně o 5 let. Chyby jednotlivých projekcí jsou k nahlédnutí v příloze B.

Cílem práce bylo vytvořit projekce úmrtnosti podle příčin úmrtí a následně je porovnat s projekcemi celkové úmrtnosti. Já jsem si všechny příčiny úmrtí rozdělila do šesti skupin, se kterými jsem následně pracovala a vytvářela projekce pro každou skupinu zvlášť. K tomuto postupu existuje více alternativ. Bylo by možné dělit všechny příčiny úmrtí do více skupin a sledovat tak vývoj v budoucnosti pro více příčin úmrtí. Tento postup by nemusel být nutně lepší, jelikož při dělení na další skupiny příčin by v takových skupinách byly velmi nízké počty úmrtí, což by mohlo vést k horším výsledkům.

Další možností by bylo nedělit příčiny do skupin, ale vytvářet projekce pro každou příčinu zvlášť. To by u českých dat vedlo ještě k horším výsledkům. U tohoto postupu by byl problém i se samotnou klasifikací nemocí. Některé příčiny úmrtí zanikají, jiné vznikají, někdy se více příčin seskupuje do jedné a naopak se někdy jedna příčina rozdělí na více příčin. Pro vytváření celkových projekcí úmrtnosti by tento postup nebyl nejlepší, mohl by však být k užitku například ve zdravotnictví kvůli informaci o vývoji konkrétní nemoci. Projekce úmrtnosti podle příčin úmrtí by se daly vylepšit také tím, že bych nepředpokládala nezávislost jednotlivých příčin úmrtí. V takovém případě by se musely počítat vzájemné korelace mezi jednotlivými příčinami, v mém případě by se musely počítat korelace mezi jednotlivými skupinami příčin úmrtí, což by bylo složité. V případě závislostí mezi jednotlivými příčinami úmrtí by se mohly použít i kopuly.

Nakonec jsem v práci srovnala výsledky získané z projekcí úmrtnosti podle příčin úmrtí seskupených do celkových úmrtností a projekcí přímo celkových úmrtností. Z tohoto porovnání jsem zjistila, že tyto dva rozdílné způsoby vedou k různým výsledkům a nelze je zaměnit. Agregované projekce úmrtnosti vyšly pro muže i ženy obecně vyšší. Střední doby života ve věku 0 vychází pro agregované projekce úmrtnosti nižší pro obě pohlaví. Obecně lze říci, že projekce úmrtnosti podle příčin úmrtí vedou k méně optimistickým výsledkům, než projekce celkových úmrtností. To může být způsobeno například tím, že když vytváříme projekce pro celkové úmrtnosti, vychází se z úmrtností, které jsou vyrovnanější a výrazně neklesají nebo nestoupají, jako tomu může být u jednotlivých příčin úmrtí. Následné projekce jsou pak také vyrovnanější, než u projekcí úmrtností podle jednotlivých příčin. Při projekcích podle příčin vznikají markantnější rozdíly, mnohem větší poklesy nebo růsty v průběhu let, což při agregaci vede k výraznějším výsledkům, než které pozorujeme u projekcí z celkových úmrtností. Obecně lze říci, že rozdělení úmrtí podle jednotlivých příčin a modelování Lee-Carterovou metodou vede k méně optimistickým výsledkům. Nelze však říci, který z těchto dvou způsobů je správnější.

# Seznam použité literatury

- [1] ANDĚL, Jiří. *Statistické metody*. Praha: Matfyzpress, 2007. ISBN 80-7378-003-8, 2003.
- [2] BOOTH, Heather. Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 2006, 22.3: 547-581.
- [3] BOOTH, Heather; TICKLE, Leonie. Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 2008, 3.1-2: 3-43.
- [4] BOWLEY, A. L. Births and population in Great Britain. *The Economic Journal*, 1924, 34.134: 188-192.
- [5] BRASS, W. Mortality models and their uses in demography. *Transactions of the Faculty of Actuaries*, 1974, 33.1971-1973: 123-142.
- [6] CANNAN, Edwin. The probability of a cessation of the growth of population in England and Wales during the next century. *The Economic Journal*, 1895, 5.20: 505-515.
- [7] GIROSI, Federico; KING, Gary. *Understanding the Lee-Carter mortality forecasting method (2007)*. Harvard University, 2007.
- [8] KOISSI, Marie-Claire; SHAPIRO, Arnold F. The Lee-Carter Model Under The Condition Of Variables Age-Specific Parameters. *In: Actuarial Research Conference, Regina, Canada*. 2008.
- [9] LEE, Ronald D.; CARTER, Lawrence R. Modeling and Forecasting US Mortality. *Journal of the American Statistical Association*, 1992, 87.419: 659-671.
- [10] MAXDORF, Velký lékařský slovník [online], c2008. [cit. 2013-07-04]. Dostupné na World Wide Web: <http://lekarske.slovniky.cz>
- [11] PROCTOR, R. N. Commentary: Schairer and Schoeniger's forgotten tobacco epidemiology and the Naziquest for racial purity. *International Journal of Epidemiology*, 2001, 30, 31-34.
- [12] RICHARDS, S. J.; et al. Two dimensional mortality data: patterns and projections. *British Actuarial Journal*, 2007, 13.03: 479-536.
- [13] RIDSDALE, Brian; et al. Mortality by cause of death and by socio-economic and demographic stratification. Paper for ICA2010, 2010.
- [14] RODU, B. Swedish tobacco use: smoking, smokeless, and history. *American Council on Science and Health*, 2004.
- [15] RODU, Brad; COLE, Philip. The burden of mortality from smoking: comparing Sweden with other countries in the European Union. *European journal of epidemiology*, 2004, 19.2: 129-131.



- [16] SHRYOCK, Henry S.; SIEGEL, Jacob S.; LARMON, Elizabeth A. *The Methods and Materials of Demography*. US Dept. of Commerce, Bureau of the Census: for sale by the Supt. of Docs. US Govt. Print. Off., 1980.
- [17] RYAN, Daniel; SINGLETON, Matt. *A window into the future: Understanding and predicting longevity (2011)*. Swiss Re Media Production Zurich.
- [18] TABEAU, Ewa. A review of demographic forecasting models for mortality. In: *Forecasting Mortality in Developed Countries*. Springer Netherlands, 2002. p. 1-32.
- [19] VIVANTIS A.S., celostnimedica.cz [online], c2004. [cit. 2013-07-04]. Dostupné na World Wide Web: <http://www.celostnimedica.cz/>
- [20] WHELPTON, Pascal K. Population of the United States, 1925 to 1975. *American Journal of Sociology*, 1928, 253-270.

# Seznam tabulek

- Tabulka 6.1: Příklady nemocí pro jednotlivé skupiny příčin úmrtí. (Strana 27.)
- Tabulka 6.2: Verze MKN v jednotlivých obdobích v rozmezí let 1950 – 2010. (Strana 28.)
- Tabulka 6.3: Správnost aproximace prvním řádem SVD pro muže, ženy a jednotlivé příčiny úmrtí. (Strana 29.)
- Tabulka 6.4: Kalendářní roky výskytů nereálných řešení rovnice 3.6 u mužů. (Strana 31.)
- Tabulka 6.5: Kalendářní roky výskytů nereálných řešení rovnice 3.6 u žen. (Strana 31.)
- Tabulka 6.6: Výsledné odhady driftů z modelu náhodné procházky pro jednotlivé příčiny úmrtí, ženy a muže. (Strana 33.)
- Tabulka 6.7: Hodnoty  $\sigma_\theta^2$  a  $\sigma_{rw}^2$  pro české muže a jednotlivé příčiny úmrtí. (Strana 36.)
- Tabulka 6.8: Hodnoty  $\sigma_\theta^2$  a  $\sigma_{rw}^2$  pro české ženy a jednotlivé příčiny úmrtí. (Strana 36.)
- Tabulka 6.9: Střední doby života získané z agregovaných měř úmrtnosti rozdělených podle příčin úmrtí a jejich projekcí pro vybrané kalendářní roky a věky pro muže, zaokrouhlené na dvě desetinná místa. (Strana 39.)
- Tabulka 6.10: Střední doby života získané z agregovaných měř úmrtnosti rozdělených podle příčin úmrtí a jejich projekcí pro vybrané kalendářní roky a věky pro ženy, zaokrouhlené na dvě desetinná místa. (Strana 39.)
- Tabulka 6.11: Střední doby života získané z celkových měř úmrtnosti a jejich projekcí pro vybrané kalendářní roky a věky pro muže, zaokrouhlené na dvě desetinná místa. (Strana 39.)
- Tabulka 6.12: Střední doby života získané z celkových měř úmrtnosti a jejich projekcí pro vybrané kalendářní roky a věky pro ženy, zaokrouhlené na dvě desetinná místa. (Strana 40.)

# Přílohy

Příložené DVD obsahuje tyto přílohy:

- A. Projekce.xls - obsahuje výpočty projekcí centrálních měř úmrtnosti podle příčin a pravděpodobností úmrtí podle příčin včetně grafů pro jednotlivé příčiny, pro muže i ženy.
- B. Chyby.xlsx - obsahuje hodnoty chybové složky Lee-Carterovy metody pro jednotlivé skupiny příčin úmrtí, pro ženy i muže včetně grafů s konfidenčními intervaly.
- C. Zbyvajici doby zivota.xlsm - obsahuje hodnoty celkových středních dob života a středních dob života podle příčin úmrtí pro ženy i muže včetně grafů.