

UNIVERZITA PALACKÉHO V OLOMOUCI  
PŘÍRODOVĚDECKÁ FAKULTA

**BAKALÁŘSKÁ PRÁCE**

Analýza úmrtnosti na základě úmrtnostních  
tabulek



**Katedra matematické analýzy a aplikací matematiky**

Vedoucí bakalářské práce: **RNDr. Ondřej Pavlačka, Ph.D.**

Vypracoval(a): **Zdeněk Zlámal**

Studijní program: B1103 Aplikovaná matematika

Studijní obor Matematika–ekonomie se zaměřením na bankovníctví/pojišťovnictví

Forma studia: prezenční

Rok odevzdání: 2020

## BIBLIOGRAFICKÁ IDENTIFIKACE

**Autor:** Zdeněk Zlámal

**Název práce:** Analýza úmrtnosti na základě úmrtnostních tabulek

**Typ práce:** Bakalářská práce

**Pracoviště:** Katedra matematické analýzy a aplikací matematiky

**Vedoucí práce:** RNDr. Ondřej Pavlačka, Ph.D.

**Rok obhajoby práce:** 2020

**Abstrakt:** Obsahem této práce bude analyzovat vybrané ukazatele a to pravděpodobnost úmrtí a střední délku života, které jsou počítány v úmrtnostních tabulkách. Na základě pravděpodobností úmrtí si sestavíme model odvíjející se od trendů v časových řadách. Pomocí tohoto modelu pak budeme moci odhadnout pravděpodobnost úmrtí v libovolném věku v jakémkoli čase. Tyto odhady pak můžeme využít v pojistné matematice, abychom tak při výpočtu pojištění dostali výsledek, který bude záviset právě na vývoji pravděpodobnosti úmrtí v čase. Na střední délce života si analyzujeme její vývoj v Česku, v jednotlivých krajích a v zahraničí.

**Klíčová slova:** úmrtnostní tabulka, pravděpodobnost úmrtí, časová řada, střední délka života

**Počet stran:** 48

**Počet příloh:** 1

**Jazyk:** český

## BIBLIOGRAPHICAL IDENTIFICATION

**Author:** Zdeněk Zlámal

**Title:** Analysis of mortality based on mortality tables

**Type of thesis:** Bachelor's

**Department:** Department of Mathematical Analysis and Application of Mathematics

**Supervisor:** RNDr. Ondřej Pavlačka, Ph.D.

**The year of presentation:** 2020

**Abstract:** The content of this work will be analyze selected indicators such as probability of deaths and life expectancy, which are calculated in mortality tables. Based on probability of deaths, we will create a model unfolding from trends in time lines. With the help of this model, we will be able to estimate the probability of deaths at any age and at any time. Then we can use this estimates in life insurance calculation, so we get a result, which will depend on the evolution of the probability of deaths in time. About life expectancy, we will analyze its evolution in the Czech Republic, regions and abroad.

**Key words:** mortality table, probability of deaths, time line, life expectancy

**Number of pages:** 48

**Number of appendices:** 1

**Language:** Czech

### **Prohlášení**

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně pod vedením pana RNDr. Ondřeje Pavlačky, Ph.D. a všechny použité zdroje jsem uvedl v seznamu literatury.

V Olomouci dne .....

.....

podpis

# Obsah

<b>Úvod</b>	<b>8</b>
<b>1 Úmrtnostní tabulky</b>	<b>9</b>
1.1 Konstrukce pravděpodobnosti úmrtí . . . . .	10
1.2 Popis úmrtnostní tabulky . . . . .	13
1.3 Dělení úmrtnostních tabulek . . . . .	17
<b>2 Predikce pravděpodobnosti úmrtí</b>	<b>18</b>
2.1 Načtení dat z úmrtnostních tabulek . . . . .	18
2.2 Vytvoření trendových funkcí . . . . .	19
2.2.1 Odhad parametrů exponenciální funkce . . . . .	22
2.3 Konstrukce modelu úmrtnosti . . . . .	26
2.4 Aplikace modelu při výpočtu pojištění důchodů . . . . .	28
<b>3 Analýza střední délky života</b>	<b>32</b>
3.1 Očekávaná délka života v historii . . . . .	32
3.2 Analýza střední délky života v krajích . . . . .	37
3.3 Analýza střední délky života v zahraničí . . . . .	41
<b>Závěr</b>	<b>45</b>
<b>Literatura</b>	<b>47</b>

# Seznam obrázků

2.1	Vývoj pravděpodobnosti úmrtí během prvního roku života . . . . .	19
2.2	Vývoj pravděpodobnosti úmrtí ve 30 letech . . . . .	20
2.3	Vývoj pravděpodobnosti úmrtí v 65 letech . . . . .	20
2.4	Proložení dat trendem během prvního roku . . . . .	23
2.5	Proložení dat trendem ve 30 letech . . . . .	24
2.6	Proložení dat trendem v 65 letech . . . . .	24
3.1	Vývoj střední délky života nově narozených jedinců. . . . .	33
3.2	Rozdíl v očekávané délce života u mužů při narození a při dožití se prvního roku. . . . .	33
3.3	Zbývající očekávaný počet let života napříč všemi generacemi v roce 2018. . . . .	34
3.4	Graf vyjadřující o kolik let se více v průměru dožijí narozené ženy oproti mužům. . . . .	35
3.5	Graf vyjadřující o kolik let se více v průměru dožijí ženy v 65 letech oproti mužům. . . . .	35
3.6	Očekávaná délka života narozených mužů v jednotlivých krajích. . . . .	37
3.7	Očekávaná délka života narozených žen v jednotlivých krajích. . . . .	38
3.8	Graf mezi průměrnou měsíční mzdou a střední délkou života v krajích. . . . .	40
3.9	Očekávaná délka života ve vybraných zemích v Evropě. . . . .	42
3.10	Graf mezi střední délkou života a průměrnou měsíční mzdou v zahraničí. . . . .	43

## **Poděkování**

Rád bych poděkoval panu RNDR. Ondřeji Pavlačkovi, Ph.D. za ochotu, vstřícný přístup a čas, během kterého dohlížel na vznik této práce.

# Úvod

Tématem této práce je analýza úmrtnosti a její aplikace z dostupných dat, které poskytují úmrtnostní tabulky. V první části práce bude popsána konstrukce nejdůležitějšího ukazatele a to pravděpodobnosti úmrtí, bez které bychom nebyli schopni dopočítat další ukazatele v úmrtnostní tabulce, jako je například očekávaná délka života člověka.

Ve druhé části práce přejdeme k analýze pravděpodobností úmrtí. Z úmrtnostních tabulek vybereme požadovaná data tvořící časové řady, ze kterých budeme schopni odvodit si rovnice trendů a odhadnout tak další možný vývoj v pravděpodobnosti úmrtí. Cílem bude na základě trendů vytvořit úmrtnostní model sloužící k predikci pravděpodobnosti úmrtí do budoucích let. Tento model pak můžeme aplikovat třeba v oblasti pojistné matematiky. Například si tak můžeme spočítat kolik korun bude stát pojištění důchodu a zjistit tak, jaká by měla být ona skutečná cena, která bude záviset na vývoji pravděpodobnosti úmrtí v čase.

Ve třetí části práce potom přejdeme k analýze očekávané délce života. Podíváme se na rozdíly očekávaného počtu let života u žen a mužů, abychom zjistili, kam se dané trendy v rozdílech ubírají a jaký by byl v budoucnu za předpokladu zachování trendů jejich důsledek. Také se podíváme na očekávanou délku života v jednotlivých krajích, abychom zjistili, kde se muži a ženy dožívají v průměru nejvíce let a kde naopak nejméně. V krajích také prozkoumáme vztah mezi průměrnou měsíční mzdou a očekávanou délkou života. V neposlední řadě pak srovnáme vývoj očekávané délky života v Česku s vybranými zahraničními zeměmi v rámci Evropy, u kterých rovněž prozkoumáme vztah mezi průměrnou měsíční mzdou v těchto zemích a očekávanou délkou života.



# 1. Úmrtnostní tabulky

V této kapitole se čtenář seznámí s pojmem úmrtnostní tabulka. Než se totiž pustíme do samotné analýzy úmrtnosti pomocí úmrtnostních tabulek, tak je potřeba si nejdříve říct, co to vlastně úmrtnostní tabulka je, jakým způsobem probíhá konstrukce jednotlivých ukazatelů a jak v konečném provedení může vypadat.

Počátek, kdy se dá považovat vznik první úmrtnostní tabulky, se datuje na konec 17. století, přesněji do roku 1693. V tomto roce známý astronom Edmond Halley studoval záznamy narození a úmrtí obyvatel z farních zápisů z města Vratislav. Ze svých poznatků pak vytvořil tabulku, která ukazovala počet přeživších v jakémkoliv věku z jedinců narozených ve stejném roce [4].

V dnešní době jsou úmrtnostní tabulky nástrojem, pomocí kterého můžeme namodelovat úmrtnost [1] v jednotlivých letech života člověka. Z toho důvodu patří k základním nástrojům a zároveň poskytují údaje k potřebným výpočtům v pojistné matematice v oblasti životních pojištění.

V České republice úmrtnostní tabulky vytváří a publikuje za celou populaci Český statistický úřad. Ten při konstrukci různých ukazatelů v tabulkách pracuje s celočíselnými roky života  $x$ . Základem pro všechny ukazatele v úmrtnostní tabulce jsou pravděpodobnosti úmrtí ve věku  $x$ , které se značí jako  $q_x$ . Například lze spočítat i pravděpodobnost přežití ve věku  $x$  označenou jako  $p_x$ , neboť lze využít vlastnosti pravděpodobnosti opačného jevu:

$$p_x = 1 - q_x. \tag{1.1}$$

## 1.1. Konstrukce pravděpodobnosti úmrtí

V této kapitole se zaměříme na popis metodiky používané Českým statistickým úřadem sloužící k odhadu pravděpodobnosti úmrtí. Po obecných informacích na úvod této podkapitoly zde poté budou představeny jednotlivé kroky, jak se z napozorovaných dat dá odhadnout pravděpodobnost úmrtí.

Konstrukce pravděpodobností  $q_x$  či  $p_x$  vychází z reálných dat, přičemž Český statistický úřad volí vhodnou metodiku k získání odhadů těchto pravděpodobností. Používá ji pro vytvoření úmrtnostních tabulek od roku 2018. Touto metodikou byl také proveden výpočet úmrtnostních tabulek v časové řadě od roku 1920 až do současnosti týkající se celé České republiky, dále pak od roku 2001 jsou sestaveny tabulky pro regiony, okresy i kraje. Pro regiony a kraje jsou zpracovány po dvou letech a za okresy a správní obvody obcí s rozšířenou působností po pěti letech [3].

Základními vstupními daty v metodice jsou napozorované údaje týkající se počtu zemřelých osob a středního stavu obyvatel rozlišených podle pohlaví v jednotlivých letech života v daném roce. V prvním kroku je tedy nutné z těchto vstupních dat spočítat tzv. míru úmrtnosti značenou  $m_x$  a to jako:

$$m_x = \frac{D_x}{1.7.P_x}, \quad (1.2)$$

kde  $D_x$  značí počet zemřelých ve věku  $x$  a  $P_x$  je skutečný počet přeživších osob ve věku  $x$  k datu 1. 7. daného roku (střední stav obyvatel).

V dalším kroku se pozorovaná míra úmrtnosti vyrovnává z důvodu výskytu náhodných výkyvů. Takové výkyvy se objevují převážně v letech, kde je úmrtnost poměrně nízká a nebo je zde nízký počet osob. Pro vyhlazení se využívají zobecněné adaptivní modely, které se ještě kombinují s tzv. P-spliny, spolu označováno jako metoda P-GAM [7]. V rámci adaptivního modelu se vyhlazení přizpůsobí na počet úmrtí, kdy nižší počet zemřelých je vyhlazen více, než vyšší počet a to vzhledem k zachování původního trendu.

Vzhledem k tomu, že údaje týkající se středního stavu obyvatel v těch

nejvyšších letech můžou být značně nespolehlivé kvůli nižším počtům zemřelých, neboť vyšších let se obecně dožívá méně lidí, tak se nahrazuje diskrétní pozorovaná míra úmrtnosti jejím modelovým protějškem. Modelová míra úmrtnosti se značí řeckým  $\mu_x$  a na rozdíl od té napozorované se jedná o veličinu spojitou. Mezi mírami platí přibližný vztah [8]:

$$m_x \cong \mu_x + 0,5. \quad (1.3)$$

Nyní je potřeba si intenzitu úmrtnosti  $\mu_x$  namodelovat. K tomu je vhodné využít logistickou křivku, která poměrně dobře zachycuje trend v úmrtnosti a zároveň zohledňuje zpomalení intenzity úmrtnosti v těch nejvyšších letech. Danou funkci lze matematicky zapsat ve tvaru:

$$\mu_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}} + c, \quad (1.4)$$

$\alpha, \beta$  se nazývají parametry logistické funkce a  $c$  je konstanta. Odhad parametrů se provádí pomocí metody maximální věrohodnosti [2]. Tuto metodu na odhad parametrů můžeme použít díky tomu, že předpokládáme, že počet zemřelých se řídí Poissonovým rozdělením pravděpodobnosti, protože Poissonovo rozdělení se týká jevů, které mají malou pravděpodobnost výskytu [2]. Na základě tohoto předpokladu se maximalizuje logaritmus věrohodnostní funkce Poissonova rozdělení  $L$  ve tvaru:

$$\ln L = \sum_x D_x \ln(\mu_x) - E_x \mu_x, \quad (1.5)$$

kde  $E_x$  je počet obyvatel ve věku  $x$ , ze kterých se sleduje počet zemřelých  $D_x$ . Pro odhad parametrů bereme věkové rozpětí 70 až 90 let. Jakmile máme odhadnuté parametry, zpětným dosazením do (1.4) dostaneme modelovou míru úmrtnosti při různých vstupech  $x$  v daném rozmezí.

Předpokládá se, že věk, od kterého modelové hodnoty popisují úmrtnost lépe než vyrovnané hodnoty, je vyšší než 75 let včetně. Pro hladký průběh nahrazení se nejprve hledá věk  $y \geq 75$  tak, aby platilo

$$y = \min |m_{x,vyrov.} - m_{x,model.}|. \quad (1.6)$$

Po nalezení tohoto věku se pak provede postupná graduace vyrovnaných měř na modelové pro věky  $x \in [(y - 4), (y + 4)]$  předpisem:

$$m_{x,grad.} = \left(1 - \frac{x - y + 5}{10}\right) \cdot m_{x,vyrov.} + \left(\frac{x - y + 5}{10}\right) \cdot m_{x,model.} \quad (1.7)$$

Pro výpočet pravděpodobnosti úmrtí se dále využívá kombinace specifických měř úmrtnosti:

- Pro věk  $0 \dots$  reálné hodnoty.
- Pro věk  $1$  až  $(y - 5) \dots$  vyrovnané hodnoty.
- Pro věk  $(y - 4)$  až  $(y + 4) \dots$  graduované hodnoty.
- Pro věk  $(y + 5) \dots$  modelové hodnoty.

Nyní již máme všechny podklady pro převod ze specifických měř úmrtnosti ( $m_x$ ) na pravděpodobnost úmrtí ve věku  $x$  ( $q_x$ ). Je ale navíc nutné předpokládat rovnost mezi skutečnou a tabulkovou mírou úmrtnosti a stabilitu měř úmrtnosti, pak lze pro převod použít rovnici [9]:

$$q_x = \frac{m_x}{1 + (1 - a_x) \cdot m_x}, \quad (1.8)$$

kde  $a_x$  je parametr udávající průměrný počet tzv. "člověkoroků" prožitých v daném věkovém intervalu zemřelými jedinci. S ohledem na předpoklad področní linearitě úmrtnosti [1] je brána jako výchozí hodnota  $0,5$  pro všechny věky kromě právě narozených a posledního uvažovaného věku. Průměrný věk zmeřelých kojenců  $a_0$  se získá ze skutečných dat jako podíl průměrného věku zemřelých kojenců ve dnech na celém kalendářním období [10].

**Poznámka 1.** Při zpracování úmrtnostních tabulek byl mezi roky 1920 až 1985 kvůli nedostatečné dostupnosti dat parametr  $a_0$  nastaven na  $0,1$ .

U posledního uvažovaného věku se parametr  $a_{105+}$  vypočítá jako převrácená hodnota k  $m_{105+}$ .

## 1.2. Popis úmrtnostní tabulky

Samotná tabulka, jejíž stručnou ukázkou reprezentuje tabulka 1.1, se skládá z devíti sloupců.

První sloupec věk představuje jednotlivé celočíselné roky života  $x$  (stáří jedince). V druhém sloupci hodnoty  $D_x$  představují skutečný počet zemřelých osob ve věku  $x$ , slouží jako podkladová data ke konstrukci odhadu pravděpodobnosti  $q_x$ .

Následně v dalším sloupci hodnoty  $P_x$  představují střední stav obyvatel pro celou Českou republiku, regiony, kraje a okresy, který je definován jako počet žijících osob k 1. 7. daného roku.

Poté následují hodnoty  $q_x$ , což jsou odhadnuté pravděpodobnosti, že jedinec ve věku  $x$  zemře před dosažením věku  $x + 1$ .

**Poznámka 2.** Při věku  $\omega = 105$  je  $q_x = 1$  (jedná o jev jistý), přestože to ještě není maximální věk, kterého je člověk schopen se dožít. Ovšem kvůli nízkým počtům doživších se jedinců již další roky neuvažujeme.

Další ukazatele již využívají napočítaných hodnot  $q_x$ . V případě  $l_0, l_1, \dots, l_\omega$  se jedná o posloupnosti fiktivního souboru osob, u kterého sledujeme počet jedinců doživších se  $x$  let. Konstrukce probíhá tak, že si nejprve zvolíme libovolný umělý počáteční počet jedinců  $l_0$ , čemuž říkáme **kořen** (radix) úmrtnostní tabulky. Postupně sledujeme kolik jedinců z kořene se dožije dalších let života  $x = 1, 2, 3, \dots, \omega$ . Protože ale nevíme, jaká je pravděpodobnost, že právě narozený jedinec se dožije daného věku  $x$  (lze označit jako  ${}_x p_0$ ), je nutné tuto posloupnost tvořit rekurentně, tedy po jednotlivých letech  $x = 0, 1, 2, \dots, \omega - 1$  způsobem:

$$l_{x+1} = l_x \cdot p_x, \quad (1.9)$$

kde  $p_x$  z rovnice (1.1) je pravděpodobnost, že jedinec ve věku  $x$  se dožije  $x + 1$  let. Tímto způsobem zkonstruujeme celou posloupnost  $l_x$ , pro kterou bude platit, že je nerostoucí (v dalších letech nikoho nového nepřidáváme):

$$l_0 \geq l_1 \geq l_2 \geq \dots \geq l_\omega. \quad (1.10)$$

Ženy 2018									
věk (x)	Dx	Px	mx	qx	lx	dx	Lx	Tx	ex
0	133	55681	0,002389	0,002384	100000	238	99795	8189175	81,89
1	15	55618	0,000297	0,000297	99762	30	99747	8089380	81,09
2	7	55002	0,000114	0,000114	99732	11	99726	7989633	80,11
3	4	54627	0,000081	0,000081	99721	8	99717	7889907	79,12
4	6	53891	0,000074	0,000074	99713	7	99709	7790190	78,13
5	2	53568	0,000079	0,000079	99705	8	99701	7690481	77,13
6	4	53731	0,000091	0,000091	99697	9	99693	7590780	76,14
7	6	56030	0,000104	0,000104	99688	10	99683	7491087	75,15
8	9	59048	0,000109	0,000109	99678	11	99673	7391404	74,15
9	8	59900	0,000097	0,000097	99667	10	99662	7291731	73,16
10	2	59176	0,000076	0,000076	99657	8	99654	7192069	72,17
11	1	55567	0,000062	0,000062	99650	6	99647	7092415	71,17
12	3	51667	0,000060	0,000060	99644	6	99641	6992769	70,18
13	5	49126	0,000069	0,000069	99638	7	99634	6893128	69,18
14	3	46987	0,000087	0,000087	99631	9	99626	6793494	68,19
15	7	45924	0,000111	0,000111	99622	11	99617	6693867	67,19
16	5	45291	0,000138	0,000138	99611	14	99604	6594250	66,20
17	6	44772	0,000166	0,000166	99597	17	99589	6494646	65,21
18	11	44764	0,000190	0,000190	99581	19	99571	6395057	64,22
19	11	45099	0,000209	0,000209	99562	21	99552	6295486	63,23
20	10	45700	0,000224	0,000224	99541	22	99530	6195934	62,24
21	7	46145	0,000237	0,000237	99519	24	99507	6096404	61,26
22	12	47713	0,000248	0,000248	99495	25	99483	5996897	60,27
23	15	51877	0,000255	0,000255	99471	25	99458	5897415	59,29
24	15	58030	0,000260	0,000260	99445	26	99432	5797957	58,30
25	18	61724	0,000262	0,000262	99419	26	99406	5698524	57,32
26	17	63981	0,000262	0,000262	99393	26	99380	5599118	56,33
27	19	66715	0,000262	0,000262	99367	26	99354	5499738	55,35
28	16	67234	0,000266	0,000266	99341	26	99328	5400384	54,36
29	14	68434	0,000277	0,000277	99315	28	99301	5301055	53,38
30	20	69160	0,000298	0,000298	99287	30	99273	5201754	52,39
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	60	155	0,512203	0,407772	1043	425	830	1878	1,80
101	40	108	0,551356	0,432206	618	267	484	1047	1,69
102	26	69	0,589953	0,455570	351	160	271	563	1,60
103	32	57	0,627542	0,477665	191	91	145	292	1,53
104	17	37	0,663723	0,498342	100	50	75	147	1,47
105	25	48	0,698153	1,000000	50	50	72	72	1,43

Tabulka 1.1: Úmrtnostní tabulka pro ženy ČR 2018

**Poznámka 3.** Při volbě  $l_0$  většinou volíme kulaté číslo jako například 10 000, 100 000, ...

Jakmile máme zkonstruovanou rekurentní posloupnost  $l_x$ , můžeme si odhadnout, s jakou pravděpodobností se jedinec ve věku  $x$  dožije dalších  $n$  let. Taková pravděpodobnost se značí  ${}_n p_x$  a při jejím vyjádření vycházíme ze vztahu (1.9), kde jedničku nahradíme obecným  $n \in \mathbb{N}$ :

$$l_{x+n} = l_x \cdot {}_n p_x, \quad (1.11)$$

$${}_n p_x = \frac{l_{x+n}}{l_x}. \quad (1.12)$$

Analogicky lze spočítat i pravděpodobnost úmrtí před dosažením věku  $x+n$  jako:

$${}_n q_x = 1 - {}_n p_x = 1 - \frac{l_{x+n}}{l_x} = \frac{l_x - l_{x+n}}{l_x}. \quad (1.13)$$

Následující ukazatel  $d_x$  přímo vychází z předchozího. Jedná se o počet jedinců z kořene  $l_0$  osob zemřelých ve věku  $x = 0, 1, 2, \dots, \omega$  a tento ukazatel se snadno spočítá jako:

$$d_x = l_x - l_{x+1}. \quad (1.14)$$

Ukazatel v sousedním sloupci  $L_x$  taktéž navazuje na počet přeživších osob. Představuje počet let prožitých jedinci ve věku  $x$ . Jinak řečeno je to střední počet tzv. "člověkoroků", které ve věku  $x$  prožije  $l_x$  osob z úmrtnostní tabulky. Pro výpočet tohoto ukazatele pro roky  $x \neq 0$  se použije aproximace:

$$L_x = l_{x+1} + \frac{1}{2} \cdot d_x = \frac{l_x + l_{x+1}}{2}, \quad (1.15)$$

kde se využil vztah z rovnice (1.14). Výraz se násobí  $\frac{1}{2}$  kvůli předpokladu področní linearity úmrtnosti, viz [1]. Pro poslední uvažující věk  $\omega = 105$  hodnotu  $l_{x+1}$  vynecháváme. Problémem je ale omezení týkající se právě narozených jedinců, u kterých nelze vyčíslit střední počet člověkoroků pomocí rovnice (1.15). Je to z

důvodu kojenecké úmrtnosti, která je naakumulována v prvních týdnech života, tudíž je porušen předpoklad področní linearit y úmrtnosti. Proto je nutné hodnotu  $\frac{1}{2}$  nahradit číslem menším. Například pro tabulku 1.1 je hodnota  $a_0 = 0,14$ . Pak je

$$L_0 = l_1 + a_0 \cdot d_0. \quad (1.16)$$

Předposlední sloupec  $T_x$  využívá hodnoty z předcházejícího ukazatele. Jedná se o počet zbylých let života jedinců doživších se věku  $x$ . Taky se dá říct, že je to střední počet "člověkoroků", které do konce svého života ještě prožije  $l_x$  osob a spočítá se jako součet:

$$T_x = L_x + L_{x+1} + \dots + L_\omega. \quad (1.17)$$

Důvodem zavedení tohoto ukazatele je aproximativní vztah, který bude sloužit pro výpočet posledního ukazatele v úmrtnostní tabulce.

Tím je sloupec  $e_x$  představující střední délku života jedince ve věku  $x$  let, která se spočítá jako:

$$e_x = \frac{T_x}{l_x}. \quad (1.18)$$



## 1.3. Dělení úmrtnostních tabulek

Tabulka 1.1 ukazuje pouze jeden z možných způsobů jak úmrtnostní tabulka může vypadat. Obecně je lze dělit na:

- Úplné a zkrácené:
  1. Úplné - Jsou to tabulky, které obsahují všechny celočíselné jednotlivé roky života člověka, tedy  $x = 0, 1, 2, \dots, \omega$ , kde  $\omega$  je maximální věk, který se při konstrukci stanoví, například  $\omega = 105$ .
  2. Zkrácené - U těchto tabulek máme delší věkové intervaly, může to být například  $x = 0, 1, 5, 10, 15, \dots$
- generační a průřezové:
  1. Generační - Tyto tabulky sledují jednu populaci současně narozených jedinců. To znamená, že si vybereme libovolný rok, podíváme se kolik jedinců se v tomto roce narodilo a sledujeme jejich úmrtí a dožití se v dalších letech jejich života. Potom úmrtnostní tabulka končí s posledním zemřelým, který se narodil v roce, jež jsme začali sledovat. Nevýhodou generačních tabulek je časová náročnost jejich konstrukce. Naopak výhodou je, že sledujeme pouze jednu generaci, což nám může dát lepší informace ohledně vývoje úmrtnosti pro danou generaci.
  2. Průřezové - U těchto tabulek se sleduje rok, ve kterém se zjistí, kolik osob žije ve vybraném roce ve věku  $x$  k 1.7. a kolik jich ve věku  $x$  zemřelo. Výhodou těchto tabulek je rychlá konstrukce, ale nevýhodou je, že se nám smíchají údaje pro několik populací dohromady.

S úmrtnostními tabulkami pracují kromě Českého statistického úřadu také například pojišťovny. Ty pro své účely používají úplné průřezové tabulky s celočíselnými věky [1].

## 2. Predikce pravděpodobnosti úmrtí

V této kapitole se zaměříme na podrobnější analýzu úmrtnostních tabulek. Na základě časových řad a modelování trendů zde bude odvozen model pro výpočet pravděpodobnosti úmrtí v budoucích letech v libovolně zvoleném věku. Poté si ukážeme, že takový model se dá využít při výpočtech pojistného v životním pojištění či důchodů - tedy kolik si žena nebo muž musí začít spořit od určitého okamžiku, aby jim byl pak v budoucnu vyplacen důchod ve stanovené výši. To dokážeme spočítat na základě průřezových tabulek. Ty ovšem mají tu nevýhodu, že nezohledňují vývoj pravděpodobnosti úmrtí v čase. Při výpočtu se tak berou v potaz hodnoty, které se v budoucnu mohou měnit. Pokud si však sestavíme zmíněný model pro budoucí odhady, dokážeme určit výši pojistného nebo důchodu právě na základě odhadovaného vývoje pravděpodobnosti úmrtí v čase. Na závěr kapitoly porovnáme výsledky mezi použitím průřezové tabulky a odvozeným modelem.

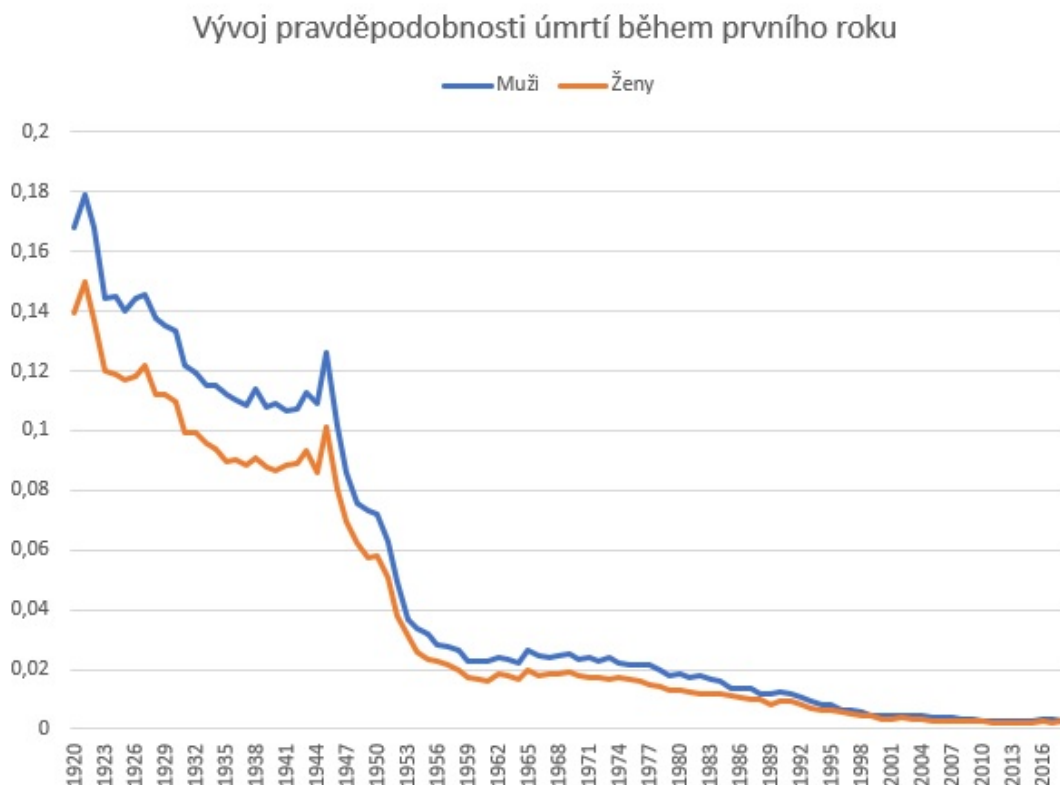
### 2.1. Načtení dat z úmrtnostních tabulek

Aby se vůbec dal vytvořit model, který bude predikovat budoucí hodnoty, je potřeba nejdříve vytáhnout z úmrtnostních tabulek ta správná data. Budeme vycházet z pravděpodobností úmrtí, které byly napozorovány v minulých letech. Tyto údaje jsou zaznamenány v úmrtnostních tabulkách od roku 1920 do roku 2018, které jsou k dispozici na stránkách Českého statistického úřadu. Nyní je potřeba si vynést do grafu požadované hodnoty tak, aby vznikla časová řada, která bude popisovat vývoj pravděpodobnosti úmrtí v daném věku. Protože jsou

ale tabulky dělané průřezově, tak je nutné vzít postupně z každé tabulky od roku 1920 hodnotu pravděpodobnosti odpovídající dané věkové kategorii. Něco takového je ale ručně příliš pracné a zdouhavé, proto je lepší pro ulehčení práce použít skript s názvem "UT-skript" napsaný v softwaru R, který je přiložený na příslušném CD i s komentáři.

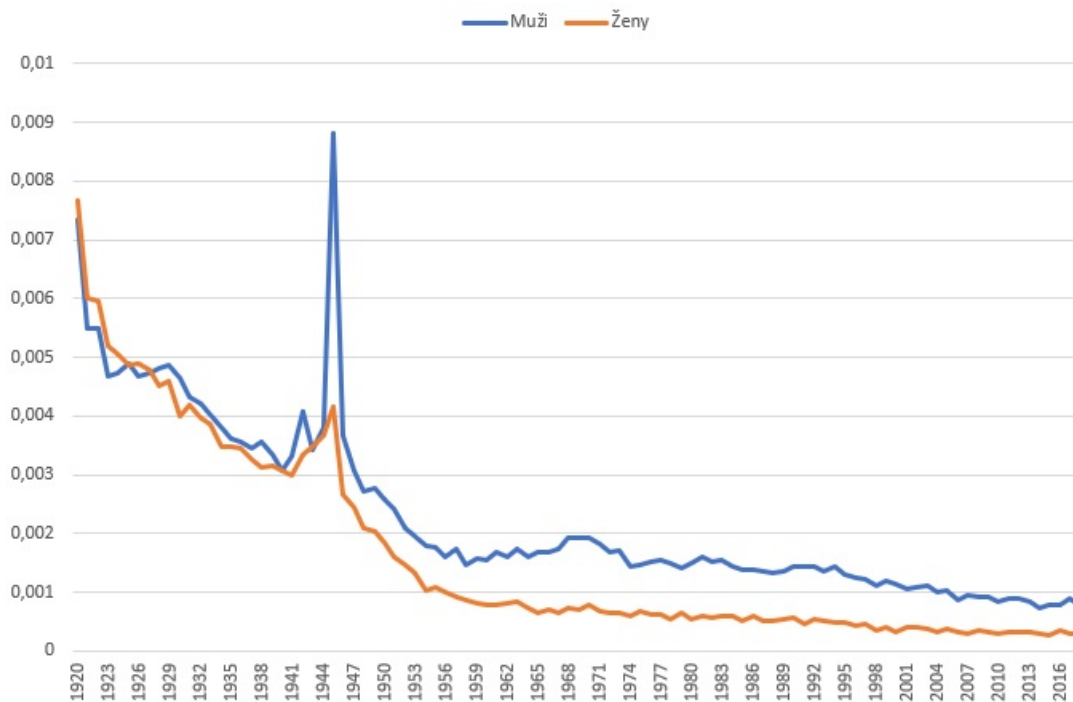
## 2.2. Vytvoření trendových funkcí

Z předchozí části víme jak získat podklady pro vytvoření modelu na odhad pravděpodobnosti úmrtí. Pro jeho konstrukci ovšem nemusíme brát všechny jednotlivé věky a prokládat je trendem, neboť pravděpodobnost úmrtí se v blízkém okolí jednotlivých věků chová podobně. Proto bude stačit, když vezmeme například každý pátý věk, tedy  $x = 0, 5, 10, \dots, 100$ . V těchto věcích určíme tren-



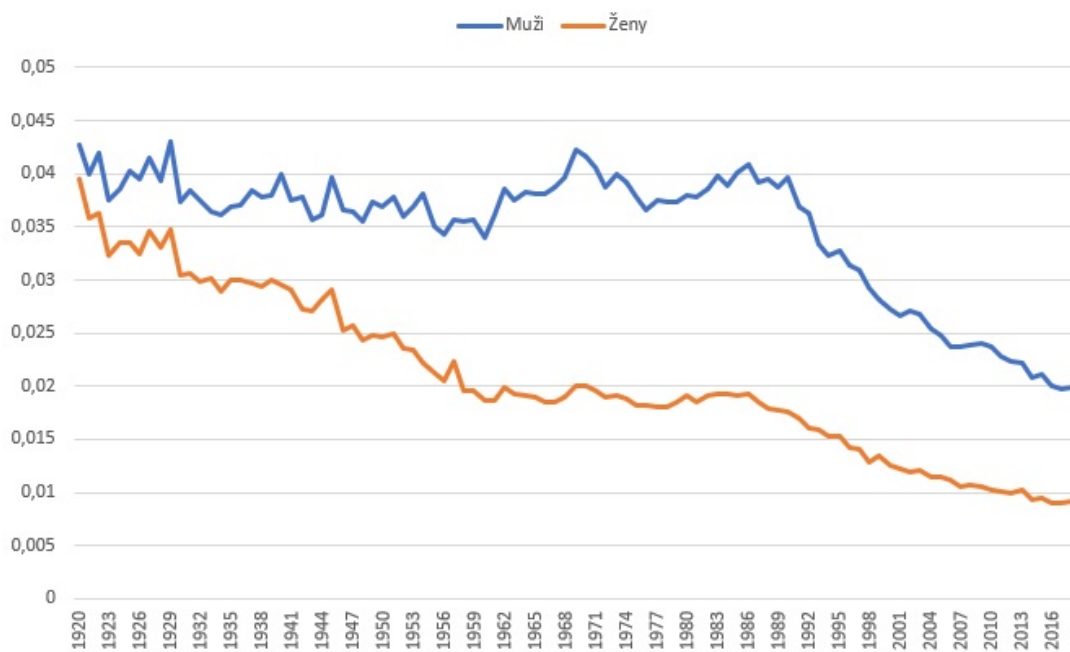
Obrázek 2.1: Vývoj pravděpodobnosti úmrtí během prvního roku života

Vývoj pravděpodobnosti úmrtí ve 30 letech



Obrázek 2.2: Vývoj pravděpodobnosti úmrtí ve 30 letech

Vývoj pravděpodobnosti úmrtí v 65 letech



Obrázek 2.3: Vývoj pravděpodobnosti úmrtí v 65 letech

dové funkce, které dokážou vyčíslit budoucí odhad  $q_x$  v čase  $t$ , kde  $t$  představuje jednotlivé kalendářní roky. Pro všechny věkové kategorie potom  $q_x$  v čase  $t$  vypočítáme jako vážený průměr mezi dvěma funkcemi.

Pro lepší přehled si získaná data můžeme uložit do jednoho MS excelovského sešitu, kde na každém listu vyneseme pravděpodobnosti  $q_x$  pro jednotlivé věky  $x = 0, 5, 10, \dots, 100$  pro muže a ženy dohromady. Poté si hodnoty pro obě pohlaví vykreslíme do jedné časové řady, kde na vodorovné ose budou roky  $t$  stejné pro všechny věkové kategorie a na svislé ose budou hodnoty  $q_0, q_5, q_{10}, \dots, q_{100}$ , jak je znázorněno ve vybraných grafech 2.1, 2.2, 2.3. Na nich je rovněž vidět, že pravděpodobnost úmrtí se od roku 1920 vyvíjí poměrně dynamicky. Během první světové války a krátce po ní se pravděpodobnost úmrtí dočasně zvýšila, příčinou kromě samotné války, byla rovněž španělská chřipka či jiné infekce nebo špatná výživa. V dalších letech je zřejmý pokles pravděpodobnosti úmrtí a to převážně u mladších lidí. Pokles trvá až do období druhé světové války, kde opět dochází k lokálnímu nárůstu. Nejvyšší nárůst se týká mladších mužů, ale i u mladších žen však dochází k dočasnému nárůstu. Oproti tomu například ženy kolem 65 let i za války zažívají pokles pravděpodobnosti úmrtí, jak je vidět na 2.3. Po druhé světové válce se pravděpodobnost úmrtí vrací na původní hodnoty a dále v čase klesá. To bylo způsobeno rozvoji antibiotik a pokrokem ve zdravotnictví. Od 60. let opět dochází ke zvýšení nebo stagnaci a to v souvislosti s civilizačními chorobami, špatným stavem životního prostředí a nezdravého životního stylu. To se nejvíce projevuje u starší populace, která je obecně náchylnější k civilizačním chorobám. Zlepšení nastává až v 90. letech a pokles pravděpodobnosti úmrtí trvá až do dnešní doby [13].

Nebylo by moc vhodné takové časové řady prokládat funkcemi, jelikož by zachycovaly historické výkyvy. Tím pádem by nemusely umět zachytit současný trend a přesnější odhad v budoucím vývoji. Proto je potřeba určit rok, od kterého bude jasně viditelný trend zachycující současný stav a zároveň pravděpodobný vývoj do budoucích let. Český statistický úřad ve své zprávě o budoucím vývoji [11] na základě detailnější analýzy bere v prognóze jako optimální období od roku 1980 až

2015, které je nejvíce relevantní. Z uvedených grafů 2.1, 2.2, 2.3, ale i z ostatních uložených v souboru "qx" na přiloženém CD, je vidět během tohoto období zlom, kdy ve všech věkových kategoriích dochází k poklesu pravděpodobnosti úmrtí. Rok 1980 se tedy zdá být vhodným počátkem časové řady.

V dalším kroku je potřeba odvodit z tvaru grafů trendové funkce, které by je proložily. V prvních letech života je vidět ukázkový příklad exponenciálního trendu, viz 2.4. V dalších letech dochází převážně u mužů k výraznějšímu poklesu na přelomu století a následně k jeho zpomalení, čemuž by taky odpovídal exponenciální trend. U žen, které mají nižší pravděpodobnost úmrtí, není sice pokles tak prudký jako u mužů, ale i tak je vidět, jak v toku času pokles pomalu zpomaluje, viz 2.5, 2.6. Podobný trend je vidět i na všech ostatních grafech uložených v souboru "trendy" na přiloženém CD. Na všechny grafy tedy použijeme proložení exponenciální funkcí, která má zároveň výhodu v tom, že zohledňuje zpomalení poklesu pravděpodobnosti úmrtí v čase a nikdy se nedostane do záporných hodnot.

### 2.2.1. Odhad parametrů exponenciální funkce

Pro vyjádření parametrů exponenciálního trendu budeme uvažovat rovnici:

$$y_i = \alpha \cdot e^{\beta \cdot t_i}, \quad (2.1)$$

kde  $y_i$  je  $i$ -tá hodnota trendu v čase  $t_i = 0, 1, 2, \dots, i = 1, 2, \dots, n$ , kde  $n$  je celkový počet pozorování,  $\alpha$  a  $\beta$  jsou parametry funkce a  $e$  je konstanta. Parametr  $\alpha$  představuje hodnotu trendu v čase  $t_1 = 0$ , v našem případě v roce 1980.  $\beta$  pak představuje násobnou změnu trendu za jednotku času. Je-li  $\beta > 1$ , jedná se o exponenciální růst, pokud je parametr  $\beta \in (0, 1)$ , bude se jednat o exponenciální pokles [6]. Odhad parametrů bude proveden metodou nejmenších čtverců, jen je potřeba rovnici (2.1) převést na tvar lineárního trendu zlogaritmováním:

$$\ln(y_i) = \ln(\alpha \cdot e^{\beta \cdot t_i}). \quad (2.2)$$

Dále použijeme pravidla pro počítání s logaritmy a rovnici převedeme do lineárního

tvaru:

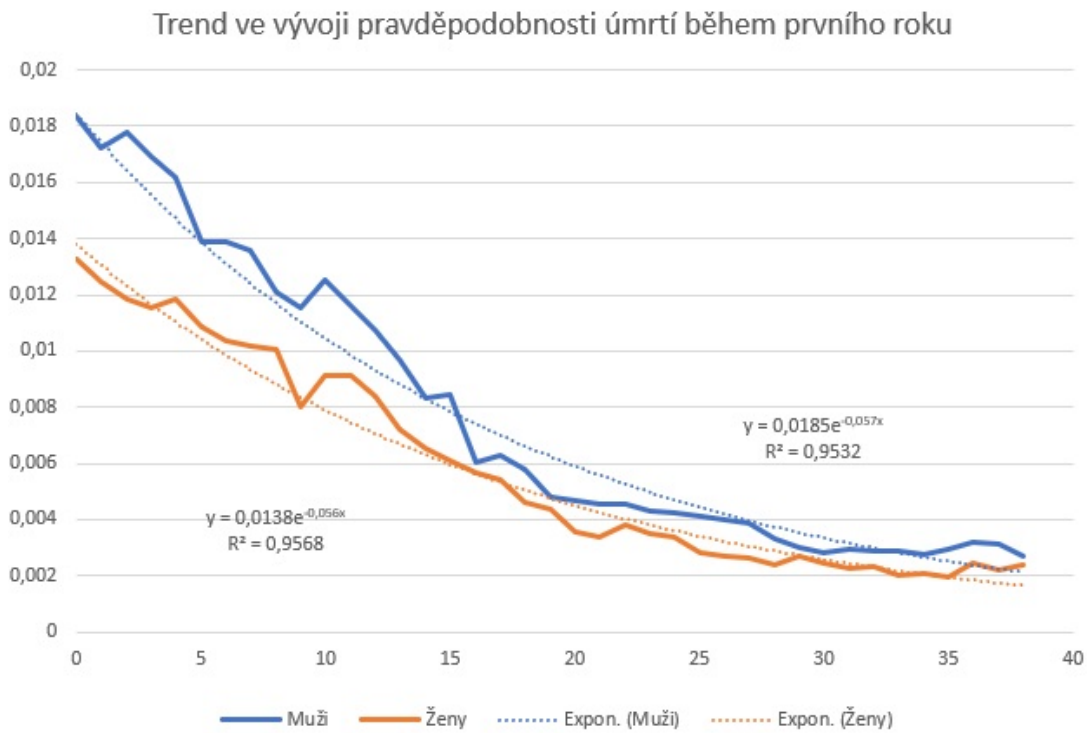
$$f_i = \alpha_1 + \beta \cdot t_i, \quad (2.3)$$

kde jsme nahradili  $f_i = \ln(y_i)$  a  $\alpha_1 = \ln(\alpha)$ . Nyní máme rovnici (2.3) v požadovaném tvaru a můžeme odhadnout parametry pomocí metody nejmenších čtverců [2].

Výsledný vztah pro odhad parametrů je pak:

$$\hat{\alpha}_1 = \frac{(\sum_{i=1}^n f_i) \cdot (\sum_{i=1}^n t_i^2) - (\sum_{i=1}^n t_i) \cdot (\sum_{i=1}^n t_i \cdot f_i)}{n \cdot \sum_{i=1}^n t_i^2 - (\sum_{i=1}^n t_i)^2}, \quad (2.4)$$

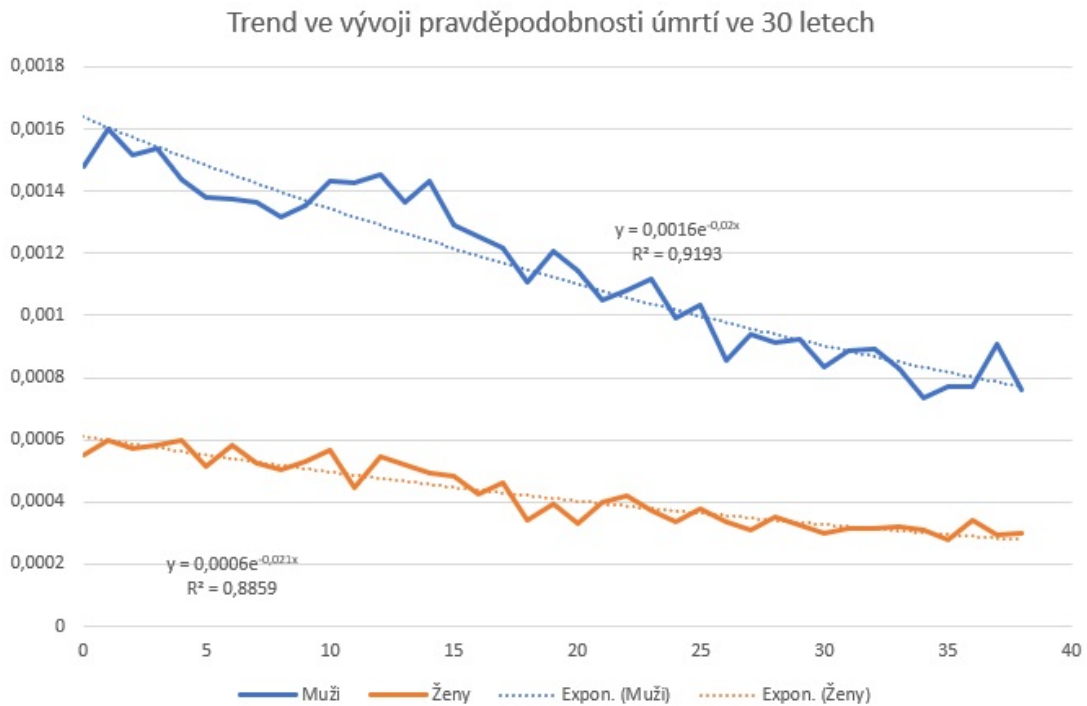
$$\hat{\beta} = \frac{n \cdot \sum_{i=1}^n t_i \cdot f_i - (\sum_{i=1}^n t_i) \cdot (\sum_{i=1}^n f_i)}{n \cdot \sum_{i=1}^n t_i^2 - (\sum_{i=1}^n t_i)^2}. \quad (2.5)$$



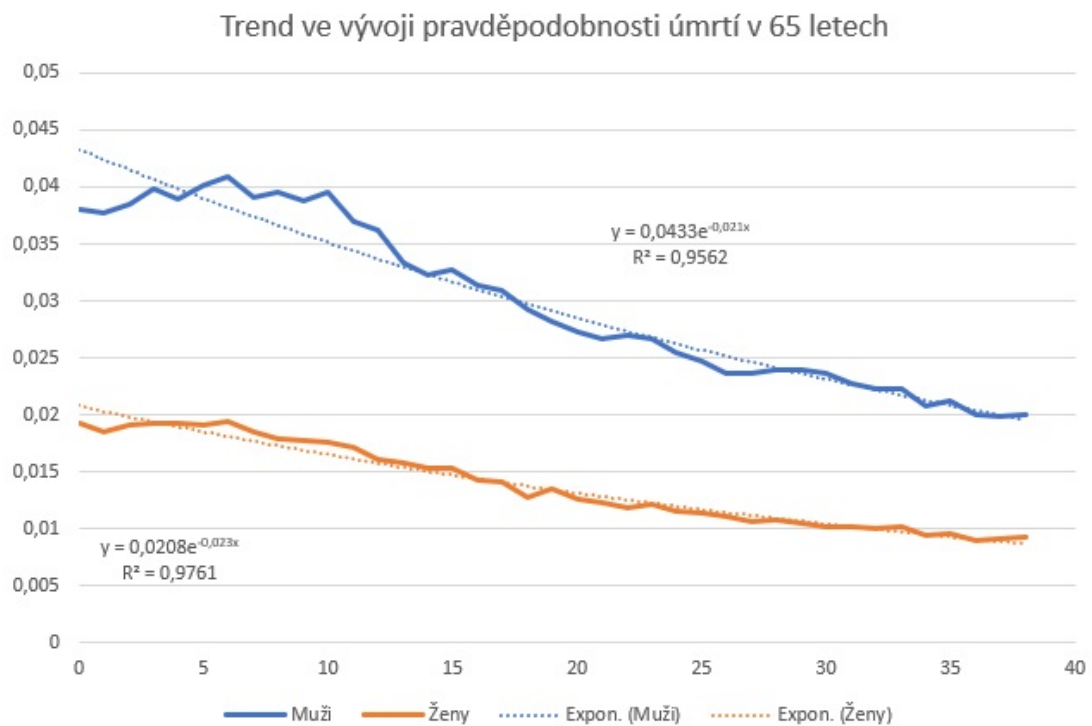
Obrázek 2.4: Proložení dat trendem během prvního roku

Tímto jsme odhadli parametry  $\hat{\beta}$  a  $\ln(\hat{\alpha})$ . Abychom mohli dosadit do rovnice (2.1), je ještě potřeba odstranit logaritmus:

$$\hat{\alpha} = e^{\hat{\alpha}_1} \quad (2.6)$$



Obrázek 2.5: Proložení dat trendem ve 30 letech



Obrázek 2.6: Proložení dat trendem v 65 letech



Muži				Ženy			
Věk	$\hat{\alpha}$	$\hat{\beta}$	$R^2$	Věk	$\hat{\alpha}$	$\hat{\beta}$	$R^2$
0	0,0185	-0,057	0,9532	0	0,01318	-0,056	0,9568
5	0,0005	-0,051	0,9411	5	0,0003	-0,041	0,8876
10	0,0004	-0,038	0,9034	10	0,0002	-0,033	0,8255
15	0,0006	-0,031	0,8709	15	0,0003	-0,021	0,6763
20	0,0012	-0,014	0,6884	20	0,0004	-0,019	0,8186
25	0,0014	-0,016	0,8518	25	0,0005	-0,019	0,7562
30	0,0016	-0,02	0,9193	30	0,0006	-0,021	0,8859
35	0,0023	-0,022	0,9233	35	0,001	-0,021	0,889
40	0,004	-0,026	0,9325	40	0,0016	-0,021	0,9114
45	0,0069	-0,026	0,9164	45	0,0028	-0,021	0,9362
50	0,0114	-0,025	0,9309	50	0,0045	-0,02	0,961
55	0,0182	-0,023	0,9404	55	0,0073	-0,02	0,9728
60	0,0289	-0,022	0,9731	60	0,0121	-0,021	0,9852
65	0,0433	-0,021	0,9562	65	0,0208	-0,023	0,9761
70	0,0656	-0,021	0,9637	70	0,0363	-0,025	0,9738
75	0,1016	-0,02237	0,98258	75	0,0647	-0,025	0,9843
80	0,1524	-0,02	0,9883	80	0,1101	-0,022	0,9848
85	0,2167	-0,016	0,9817	85	0,1748	-0,017	0,9846
90	0,2903	-0,01	0,9291	90	0,2579	-0,011	0,9667
95	0,3688	-0,005	0,7119	95	0,35	-0,006	0,8512
100	0,4428	-0,002	0,2117	100	0,4396	-0,002	0,4362

Tabulka 2.1: Odhadnuté parametry trendu a kvalita aproximace.

Dále můžeme také posoudit kvalitu našeho trendu pomocí koeficientu determinace, což je číslo udávající, jak dobře funkce proloží naměřená data [2]. Čím blíže bude koeficient 1, tím přesnější je naše aproximace. Výpočet vychází z rozptylů:

$$S_y^2 = \frac{1}{n} \cdot \sum_{i=1}^n (y_i - \bar{y})^2, \quad (2.7)$$

kde  $S_y^2$  je celkový rozptyl,  $y_i$  jsou původní hodnoty a  $\bar{y}$  je průměr původních hodnot. Druhým rozptylem bude rozptyl vyrovnaných hodnot:

$$S_{\hat{y}}^2 = \frac{1}{n} \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (2.8)$$

kde  $\hat{y}_i$  jsou hodnoty vypočítané pomocí odhadnuté funkce a  $\bar{y}$  je opět průměr

původních hodnot. Výsledný koeficient determinace pak dostaneme jako:

$$R^2 = \frac{S_{\hat{y}}^2}{S_y^2} \quad (2.9)$$

**Poznámka 4.** *Odhady parametrů exponenciálních trendů pomocí metody nejmenších čtverců a koeficienty determinace můžeme dostat pomocí nástrojů grafu v MS Excelu a křivku si do grafu rovnou vykreslit. U některých věků je koeficient determinace malý. Není to však tím, že by daný trend nevyhovoval, ale tím, že v těchto věcích jsou naměřené velké výkyvy, které snižují kvalitu aproximace.*

### 2.3. Konstrukce modelu úmrtnosti

V této části již přejdeme k finální podobě modelu pro budoucí odhady  $q_x$ . Model bude sestaven jako po částech exponenciální funkce. Budou v něm zahrnuty dvě proměnné: jedna proměnná bude věk  $x = 0, 1, 2, \dots, 100$  a druhá proměnná bude čas  $t$ . Na proměnné  $x$  bude záviset výběr funkce a také váhy  $(w_1, w_2) \in \langle 0, 1 \rangle$  potřebné pro výpočet váženého průměru dvou funkcí. Váhy budou lineárně rozloženy mezi věkové rozmezí pěti let, čemuž odpovídají věky s odhadnutými trendy. V tomto rozmezí se budou váhy měnit v závislosti na tom, jaký věk  $x$  z daného intervalu zvolíme. Pokud například vezmeme  $x = 6$ , tak bude mít logicky větší váhu funkce u věku 5 let a menší váhu funkce u 10 let. Obecně se váhy  $w_1, w_2$  vypočítají způsobem:

$$w_1 = \frac{x - x_{dolni}}{5} \quad (2.10)$$

$$w_2 = 1 - \frac{x - x_{dolni}}{5}, \quad (2.11)$$

kde  $x_{dolni}$  značí poslední předcházející věk, u kterého byla odhadnuta trendová funkce, platí tedy  $x_{dolni} < x$ . Například pokud  $x = 7$ , pak  $x_{dolni} = 5$  nebo  $x = 30$ , potom  $x_{dolni} = 25$ . Pokud zvolíme pro výpočet přímo věk s odhadnutým trendem,

pak váha funkce odpovídající tomuto věku bude  $w_1 = 1$ , pro druhou funkci bude  $w_2 = 0$ . Z toho rovněž plyne, že mezi dvěma funkcemi dostane váhu  $w_1$  funkce odpovídající horní věkové hranici a  $w_2$  naopak funkce odpovídající nižšímu věku  $x_{dolni}$ .

**Poznámka 5.** *Speciálním případem bude věk  $x = 0$ , protože před ním už není žádný s trendovou funkcí. V tomto případě položíme  $x = x_{dolni}$  a dopočítáme váhy  $w_1 = 0$  a  $w_2 = 1$ , přiřazení vah jednotlivým funkcím však zůstává stejné.*

Na druhé proměnné  $t$  pak bude záviset výstupní hodnota z exponenciálního trendu v časové řadě. Číslování v časové řadě ovšem není od hodnoty 1980, nýbrž od 0. Přesto si ale za  $t$  můžeme dosadit přímo rok a od něj odečíst konstantu 1980.

Výsledná funkce pro muže označená jako  $f_m$  pak vypadá následovně:

$$f_m(x, t) = \left\{ \begin{array}{ll} \left( \frac{x-0}{5} \right) \cdot 0,0005 \cdot e^{-0,051 \cdot (t-1980)} + \left( 1 - \frac{x-0}{5} \right) \cdot 0,0185 \cdot e^{-0,057 \cdot (t-1980)} & x \in (0, 5), \\ \left( \frac{x-5}{5} \right) \cdot 0,0004 \cdot e^{-0,038 \cdot (t-1980)} + \left( 1 - \frac{x-5}{5} \right) \cdot 0,0005 \cdot e^{-0,051 \cdot (t-1980)} & x \in (5, 10), \\ \left( \frac{x-10}{5} \right) \cdot 0,0006 \cdot e^{-0,031 \cdot (t-1980)} + \left( 1 - \frac{x-10}{5} \right) \cdot 0,0004 \cdot e^{-0,038 \cdot (t-1980)} & x \in (10, 15), \\ \left( \frac{x-15}{5} \right) \cdot 0,0012 \cdot e^{-0,014 \cdot (t-1980)} + \left( 1 - \frac{x-15}{5} \right) \cdot 0,0006 \cdot e^{-0,031 \cdot (t-1980)} & x \in (15, 20), \\ \left( \frac{x-20}{5} \right) \cdot 0,0014 \cdot e^{-0,016 \cdot (t-1980)} + \left( 1 - \frac{x-20}{5} \right) \cdot 0,0012 \cdot e^{-0,014 \cdot (t-1980)} & x \in (20, 25), \\ \left( \frac{x-25}{5} \right) \cdot 0,0016 \cdot e^{-0,02 \cdot (t-1980)} + \left( 1 - \frac{x-25}{5} \right) \cdot 0,0014 \cdot e^{-0,016 \cdot (t-1980)} & x \in (25, 30), \\ \left( \frac{x-30}{5} \right) \cdot 0,0023 \cdot e^{-0,022 \cdot (t-1980)} + \left( 1 - \frac{x-30}{5} \right) \cdot 0,0016 \cdot e^{-0,02 \cdot (t-1980)} & x \in (30, 35), \\ \left( \frac{x-35}{5} \right) \cdot 0,004 \cdot e^{-0,026 \cdot (t-1980)} + \left( 1 - \frac{x-35}{5} \right) \cdot 0,0023 \cdot e^{-0,022 \cdot (t-1980)} & x \in (35, 40), \\ \left( \frac{x-40}{5} \right) \cdot 0,0069 \cdot e^{-0,026 \cdot (t-1980)} + \left( 1 - \frac{x-40}{5} \right) \cdot 0,004 \cdot e^{-0,026 \cdot (t-1980)} & x \in (40, 45), \\ \left( \frac{x-45}{5} \right) \cdot 0,0114 \cdot e^{-0,025 \cdot (t-1980)} + \left( 1 - \frac{x-45}{5} \right) \cdot 0,0069 \cdot e^{-0,026 \cdot (t-1980)} & x \in (45, 50), \\ \left( \frac{x-50}{5} \right) \cdot 0,0182 \cdot e^{-0,023 \cdot (t-1980)} + \left( 1 - \frac{x-50}{5} \right) \cdot 0,0114 \cdot e^{-0,025 \cdot (t-1980)} & x \in (50, 55), \\ \left( \frac{x-55}{5} \right) \cdot 0,0289 \cdot e^{-0,022 \cdot (t-1980)} + \left( 1 - \frac{x-55}{5} \right) \cdot 0,0182 \cdot e^{-0,023 \cdot (t-1980)} & x \in (55, 60), \\ \left( \frac{x-60}{5} \right) \cdot 0,0433 \cdot e^{-0,021 \cdot (t-1980)} + \left( 1 - \frac{x-60}{5} \right) \cdot 0,0289 \cdot e^{-0,022 \cdot (t-1980)} & x \in (60, 65), \\ \left( \frac{x-65}{5} \right) \cdot 0,0656 \cdot e^{-0,021 \cdot (t-1980)} + \left( 1 - \frac{x-65}{5} \right) \cdot 0,0433 \cdot e^{-0,021 \cdot (t-1980)} & x \in (65, 70), \\ \left( \frac{x-70}{5} \right) \cdot 0,1016 \cdot e^{-0,02237 \cdot (t-1980)} + \left( 1 - \frac{x-70}{5} \right) \cdot 0,0656 \cdot e^{-0,021 \cdot (t-1980)} & x \in (70, 75), \\ \left( \frac{x-75}{5} \right) \cdot 0,1524 \cdot e^{-0,02 \cdot (t-1980)} + \left( 1 - \frac{x-75}{5} \right) \cdot 0,1016 \cdot e^{-0,02237 \cdot (t-1980)} & x \in (75, 80), \\ \left( \frac{x-80}{5} \right) \cdot 0,2167 \cdot e^{-0,016 \cdot (t-1980)} + \left( 1 - \frac{x-80}{5} \right) \cdot 0,1524 \cdot e^{-0,02 \cdot (t-1980)} & x \in (80, 85), \\ \left( \frac{x-85}{5} \right) \cdot 0,2903 \cdot e^{-0,01 \cdot (t-1980)} + \left( 1 - \frac{x-85}{5} \right) \cdot 0,2167 \cdot e^{-0,016 \cdot (t-1980)} & x \in (85, 90), \\ \left( \frac{x-90}{5} \right) \cdot 0,3688 \cdot e^{-0,005 \cdot (t-1980)} + \left( 1 - \frac{x-90}{5} \right) \cdot 0,2903 \cdot e^{-0,01 \cdot (t-1980)} & x \in (90, 95), \\ \left( \frac{x-95}{5} \right) \cdot 0,4428 \cdot e^{-0,002 \cdot (t-1980)} + \left( 1 - \frac{x-95}{5} \right) \cdot 0,3688 \cdot e^{-0,005 \cdot (t-1980)} & x \in (95, 100). \end{array} \right.$$

Analogicky se vypočítají pravděpodobnosti  $q_x$  v čase  $t$  i pro ženy po obměně příslušných parametrů exponenciálních funkcí. Pro usnadnění výpočtu lze použít funkci sestavenou v MS Excel v souboru "qx-koeficienty" uloženém na CD.

## 2.4. Aplikace modelu při výpočtu pojištění důchodů

V této části použijeme vytvořený model na výpočet pojištění důchodů. Důchodovým životním pojištěním rozumíme pojištění pro případ dožití, kdy se pojistné plnění vyplácí ve formě životních důchodů. Ty můžeme dělit na různé typy. Pokud osoba obdrží výplatu důchodu na začátku daného období (například na počátku roku), jedná se o *předlhůtní* důchod. Jestliže výplatu důchodu dostává až na konci uplynulého období (například konec roku), jde o důchod *polhůtní*. Dále pokud je důchod vyplácen hned od zaplacení pojištění důchodu, tak se jedná o důchod *bezprostřední*. Jestli je však mezi zaplaceným pojištěním důchodu a výplatou doba odkladu  $k$  let (a je jedno jestli je pojištění důchodu zapláceno jednorázově nebo se platí oněch  $k$  let běžně), jde o důchod *odložený* o  $k$  let. Další dělení je podle frekvence doby výplaty. Může být vyplácen *ročně* nebo *področně* (každý měsíc, každého půl roku atd.). Také lze důchody dělit podle doby výplaty a to na *doživotní*, kdy je důchod vyplácen až do smrti člověka, a nebo *dočasný*, který trvá po dobu  $n$  let [14].

Princip odvození počáteční hodnoty pojištění důchodu, tedy kolik musí osoba jednorázově zaplatit ve věku  $x$ , aby pak mohla pobírat svůj stanovený důchod, je u všech typů stejný. Vezměme například důchod *předlhůtní bezprostřední dočasný roční* vyplácený ve výši 1 koruny od okamžiku  $x$  let. Nás bude zajímat, kolik korun musíme zaplatit na počátku ve věku  $x$ . Musíme jít v čase postupně zpátky až o  $n$  let, budeme tedy provádět diskontování a to pomocí technické úrokové míry  $i$  [1]. Diskontní faktor si označíme jako  $v = \frac{1}{1+i}$ . Abychom do výpočtu zahrnuli také pravděpodobnost úmrtí, tak z úmrtnostní tabulky vezmeme ukazatel  $l_x$ . Na posloupnosti budeme pohlížet jako na fiktivní počet lidí, kteří si sjednali stejné důchodové pojištění ve výši 1 koruny. Pak budeme chtít vědět, kolik lidí z  $l_x$  se dožilo první výplaty důchodu, vezeme tedy hodnotu  $l_x$ , dál počet doživších se

druhé výplaty důchodu  $l_{x+1}$ , třetí  $l_{x+2}$ , až do poslední  $l_{x+n-1}$ . V každém roce zdiskontujeme korunu v čase, kterou pak pronásobíme očekávaným počtem přeživších v daném roce a podělíme počtem žijících osob  $l_x$ . Tím dostaneme počáteční hodnotu pojištění důchodu vztaženou na jednu osobu při výplatě pouze v jednom roce. Protože se ale důchod vyplácí každým rokem, je potřeba vše sečíst. Odvozený výpočet lze zapsat následovně:

$$\ddot{a}_{x,n} = \frac{l_x + l_{x+1} \cdot v + l_{x+2} \cdot v^2 + \dots + l_{x+n-1} \cdot v^{n-1}}{l_x}. \quad (2.12)$$

Na závěr hodnotu  $\ddot{a}_{x,n}$  vynásobíme požadovaným ročním důchodem a dostaneme, kolik musí osoba ve věku  $x$  zaplatit na pojištění důchodu, aby požadovaný důchod mohla potom pobírat.

V případě, kdy vycházíme z průřezových úmrtnostních tabulek, máme k dispozici již všechny hodnoty z posloupnosti  $l_x$  pro výpočet, protože posloupnost vychází z údajů napočítaných během jednoho roku. Ovšem v dalších letech může  $l_x$  vypadat jinak, protože se může změnit pravděpodobnost úmrtí. Z toho důvodu je vhodnější použít zkonstruovaný model, který dokáže zachytit odhadovaný vývoj v čase  $t$ . Hodnoty  $l_x$  v budoucích letech  $t$  lze rekurentně spočítat podle (1.9) jako:

$$l_{x+1}(t+1) = l_x(t) \cdot (1 - q_x(t)), \quad t = 2018, 2019, \dots, \quad x = 0, 1, 2, \dots, 100. \quad (2.13)$$

**Příklad 1.** *Uvažujme zvlášť muže a ženu ve věku 50 let, kteří si sjednali v roce 2018 bezprostřední pojištění důchodu na dobu 25 let do roku 2043. Kolik musí jednorázově zaplatit, aby ročně na důchodu dostávali 20 000 korun při technické úrokové míře  $i = 1,3 \%$ ?*

Nejdříve si napočítejme hodnoty posloupnosti  $l_x$  z rovnice (2.13). Jako výchozí hodnotu můžeme použít  $l_{50}(2018)$  z úmrtnostní tabulky pro rok 2018. Od tohoto základu pak už dopočítáme  $l_{51}(2019), l_{52}(2020), \dots, l_{75}(2043)$ , viz sloupce  $l_{x,muži}(t)$  a  $l_{x,zeny}(t)$  v tabulce 2.2. Potom už jen dosadíme potřebné hodnoty do

rok(t)	věk(x)	$q_{x,muzi}(t)$	$l_{x,muzi}(t)$	$UT_{muzi}(2018)$	$q_{x,zeny}(t)$	$l_{x,zeny}(t)$	$UT_{zeny}(2018)$
2018	50	0,004293	95183	95183	0,002018	97661	97661
2019	51	0,004924	94774	94774	0,00232	97464	97464
2020	52	0,005418	94308	94310	0,002525	97238	97252
2021	53	0,005889	93797	93794	0,002722	96992	97025
2022	54	0,006339	93244	93229	0,00291	96728	96778
2023	55	0,006769	92653	92616	0,003089	96447	96506
2024	56	0,007488	92026	91950	0,003383	96149	96204
2025	57	0,008174	91337	91227	0,003662	95824	95863
2026	58	0,00883	90590	90438	0,003927	95473	95477
2027	59	0,009456	89790	89575	0,004178	95098	95041
2028	60	0,010053	88941	88626	0,004416	94700	94552
2029	61	0,010962	88047	87580	0,004807	94282	94007
2030	62	0,011833	87082	86424	0,005175	93829	93406
2031	63	0,012667	86052	85148	0,00552	93343	92748
2032	64	0,013464	84962	83748	0,005844	92828	92031
2033	65	0,014227	83818	82225	0,006147	92286	91253
2034	66	0,015366	82625	80583	0,006688	91718	90410
2035	67	0,016452	81355	78830	0,007194	91105	89499
2036	68	0,017486	80017	76971	0,007666	90450	88512
2037	69	0,01847	78618	75001	0,008106	89756	87439
2038	70	0,019406	77166	72908	0,008515	89029	86265
2039	71	0,020631	75668	70681	0,009604	88271	84975
2040	72	0,021783	74107	68318	0,010634	87423	83553
2041	73	0,022863	72493	65828	0,011608	86493	81990
2042	74	0,023875	70835	63226	0,012527	85489	80278
2043	75	0,024822	69144	60531	0,013393	84418	78420

Tabulka 2.2: Hodnoty posloupnosti ze sestaveného modelu a úmrtnostní tabulky z roku 2018.

rovnice (2.12):

$$\ddot{a}_{50,25}^{muz} = 20000 \cdot \left( \frac{95183 + 94774 \cdot \frac{1}{1+0,013} + 94308 \cdot \left(\frac{1}{1+0,013}\right)^2 + \dots + 70835 \cdot \left(\frac{1}{1+0,013}\right)^{24}}{95183} \right) = 389\,710$$

$$\ddot{a}_{50,25}^{zena} = 20000 \cdot \left( \frac{97661 + 97464 \cdot \frac{1}{1+0,013} + 97238 \cdot \left(\frac{1}{1+0,013}\right)^2 + \dots + 85489 \cdot \left(\frac{1}{1+0,013}\right)^{24}}{97661} \right) = 411\,170$$

Je vidět, že pro ženu je počáteční hodnota pojištění důchodu dražší, než-li pro muže. To je způsobeno tím, že ženy mají větší pravděpodobnost dožití se postupných vyplácení důchodů, tím pádem na začátku musí zaplatit víc.

Ještě srovnajme výpočet pomocí úmrtnostní tabulky pro rok 2018, viz sloupce  $UT_{muži}(2018)$  a  $UT_{zeny}(2018)$  v tabulce 2.2. Analogickým dosazením do (2.12) dostaneme  $\ddot{a}_{50,25}^{muz} = 382\,188$  a  $\ddot{a}_{50,25}^{zena} = 406\,461$ . Z výpočtů vycházejících z modelu tak lze zjistit, že počáteční hodnota pojištění důchodů je kvůli předpokladu poklesu pravděpodobnosti úmrtí v čase o něco větší, než při výpočtech z průřezové úmrtnostní tabulky, protože při očekávaném delším dožití musí zaplatit více. Ze srovnání hodnot rovněž plyne, že ačkoliv ženy platí více, tak by si větší částku připlatili muži, což potvrzuje, že u mužů je trend v poklesu pravděpodobnosti úmrtí větší, než u žen.

## 3. Analýza střední délky života

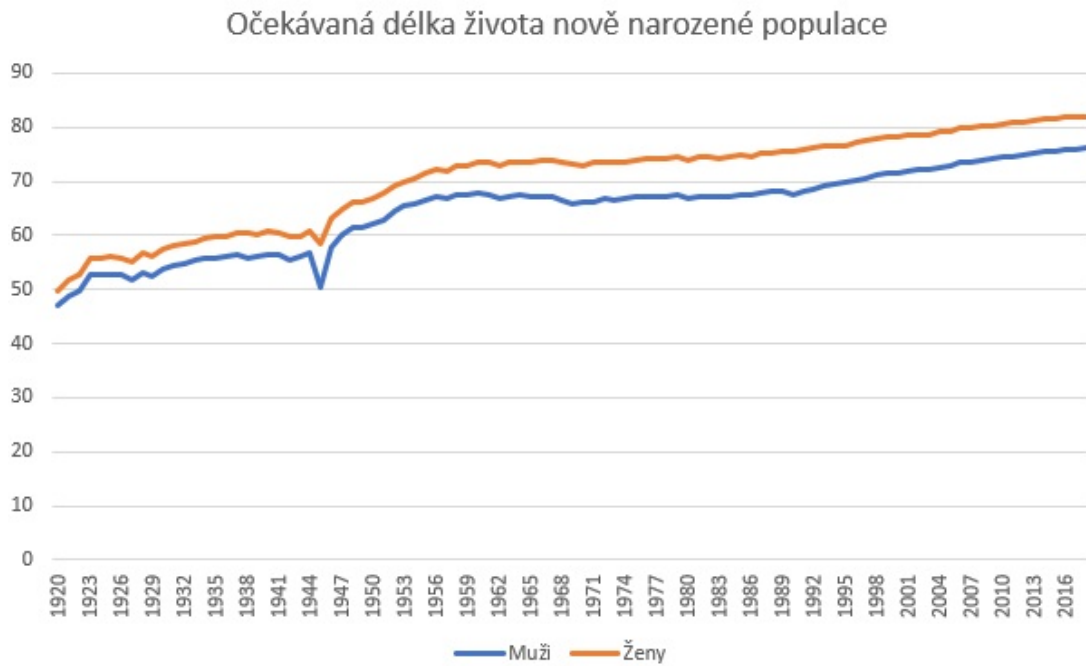
V této kapitole se zaměříme na další ukazatel z úmrtnostní tabulky a tím je střední délka života nebo také očekávaná délka života člověka. Připomeňme ještě z první části, co je střední délka života. V úmrtnostní tabulce najdeme očekávanou délku života pod označením  $e_x$ . Jedná se o počet očekávaných let, kterých se osoba ve věku  $x$  má ještě v průměru dožít. Nejvíce vypovídající hodnotu z posloupnosti  $e_x$  má první číslo  $e_0$ , což je očekávaný počet let, kterých se v průměru má dožít právě narozený člověk. Je to z toho důvodu, že se v tomto čísle skrývají veškeré napočítané pravděpodobnosti  $q_x$  za daný rok, a proto je to nejvhodnější hodnota pro analýzu.

### 3.1. Očekávaná délka života v historii

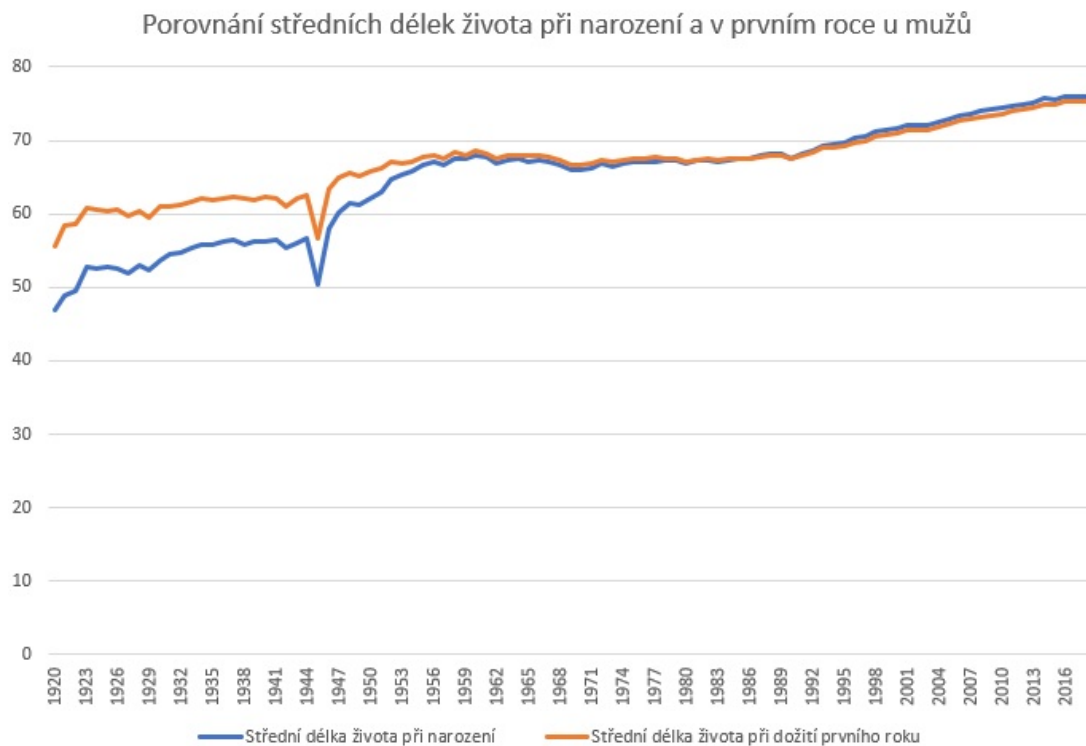
Jako první se podíváme na to, jak se vyvíjela střední délka života od roku 1920 do roku 2018 zvláště pro muže a ženy. Podobně jako v předchozí kapitole, tak i tady pro vybrání dat z úmrtnostních tabulek do časové řady se dá použít skript "UT-skript" uložený na přiloženém CD.

Pohledem na graf 3.1 bychom mohli říct, že střední délka života pro nově narozené se zvýšila v čase přibližně o třicet let. Uvědomme si však, jak velká v té době byla kojenecká úmrtnost, která je vidět na grafu z předchozí kapitoly 2.1. Tak vysoká pravděpodobnost úmrtí v prvním roce života se určitě musela podílet při výpočtu očekávané délky života na jejím zkrácení. Pokud je tomu opravdu tak, tak během doby, kdy byla vysoká kojenecká úmrtnost, nelze říct, že očekávaný počet zbylých let života jedinců ve věku jednoho roku, tedy  $e_1$ , bude nižší než  $e_0$ .

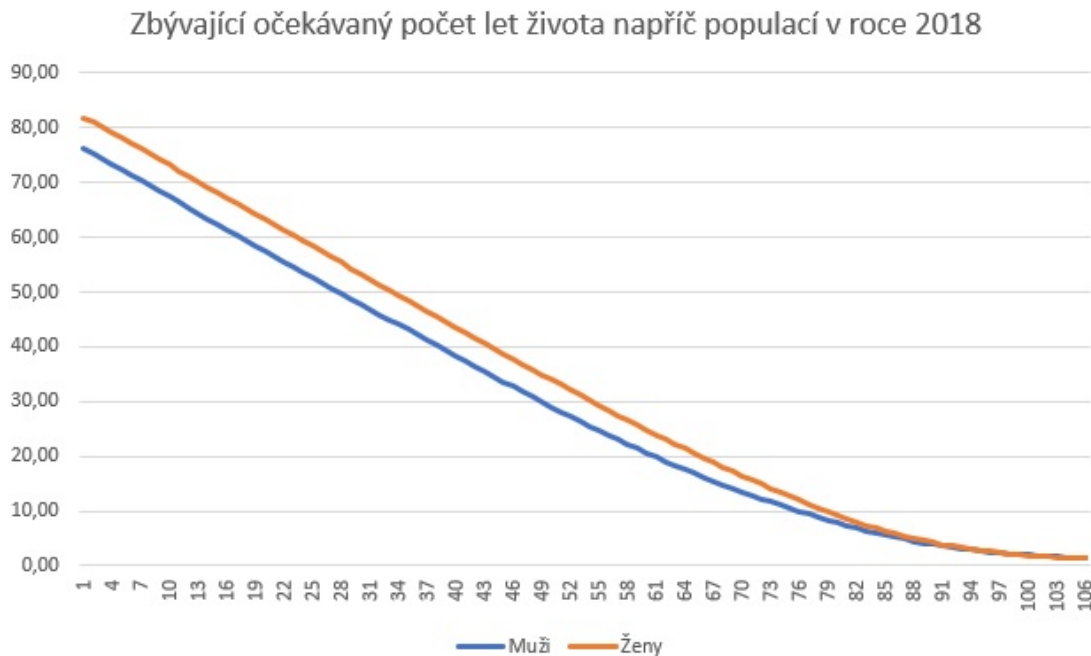




Obrázek 3.1: Vývoj střední délky života nově narozených jedinců.



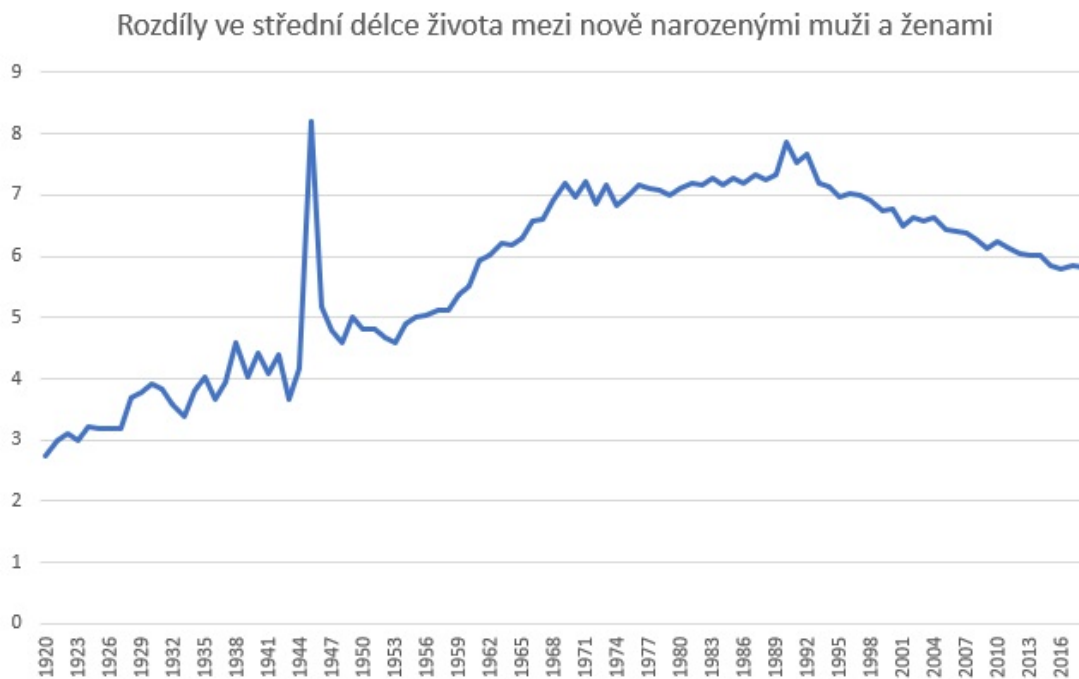
Obrázek 3.2: Rozdíl v očekávané délce života u mužů při narození a při dožití se prvního roku.



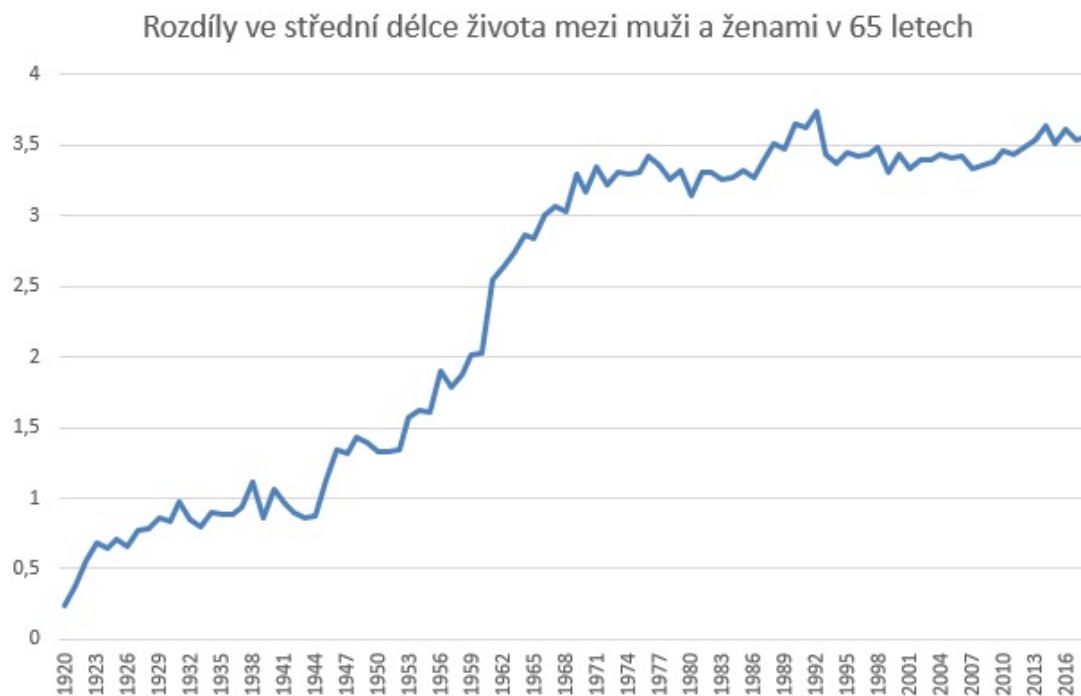
Obrázek 3.3: Zbývající očekávaný počet let života napříč všemi generacemi v roce 2018.

Jako důkaz si vykresleme dané hodnoty  $e_1$  a porovnejme je s  $e_0$ . Z grafu 3.2 vidíme, že jakmile se populace dožila prvního roku, tedy překonala kojeneckou úmrtnost v prvních týdnech života, najednou se zvedla i očekávaná délka života. Intuitivně bychom však spíše očekávali, že s přibývajícím věkem se bude střední délka života snižovat, a kdy nově narození by měli mít před sebou nejvíc let života. Z grafu 3.3 je znát, že v současné době v roce 2018 už tomu tak opravdu je. Na příkladu z historie ale můžeme vidět, že kvůli nadměrné kojenecké úmrtnosti tomu tak být nemusí a hodnota  $e_0$  nám tím pádem v dřívějších dobách nepředstavuje maximum z ukazatele  $e_x$ . Očekávaná délka života nově narozených je tak zkrslena právě vysokou pravděpodobností úmrtí zapříčiněnou kojeneckou úmrtností v prvních týdnech života.

Zůstaňme ještě u grafu 3.1 a podívejme se také na rozdíl v dožití se u mužů a žen. Je patrné, že daleko menší rozdíl, než je nyní, byl v první polovině 20. století. Od 60. let 20. století lze pak pozorovat větší rozdíly, které přetrvávají dodnes. Z grafu 3.1 však nedokážeme určit trend, kterým se rozdíly ubírají. Proto si napočítejme



Obrázek 3.4: Graf vyjadřující o kolik let se více v průměru dožijí narozené ženy oproti mužům.



Obrázek 3.5: Graf vyjadřující o kolik let se více v průměru dožijí ženy v 65 letech oproti mužům.

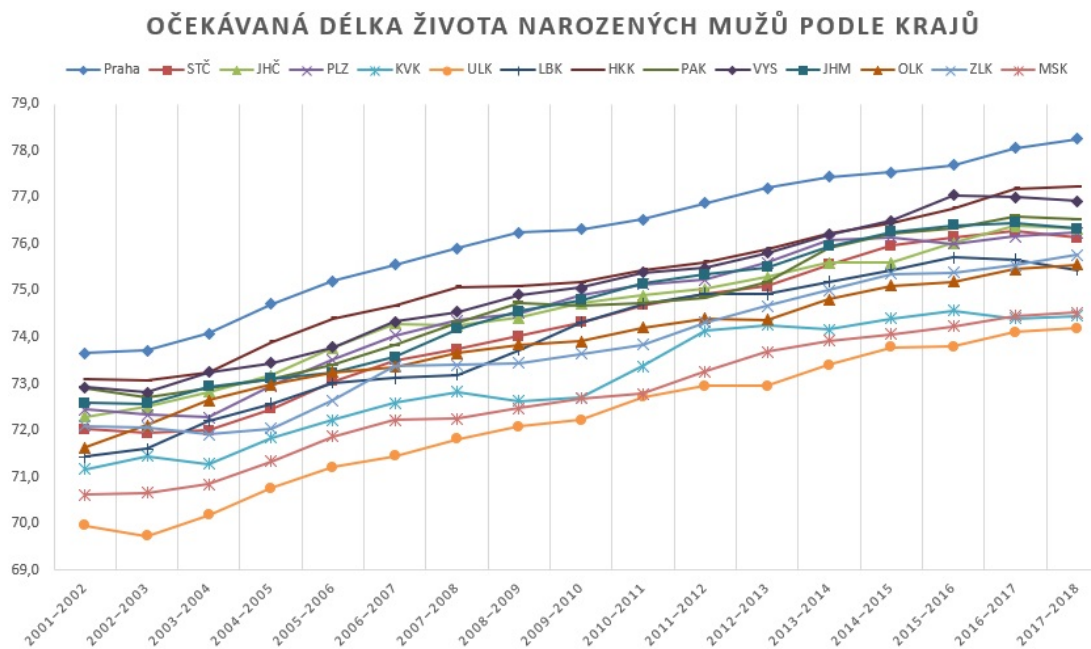
$e_{0,zeny} - e_{0,muzi}$  a vykresleme do samostatného grafu 3.4. Najednou jsme dostali naprosto odlišnou křivku, na které vidíme, že k pozvolnému zvyšování rozdílu docházelo už v 1. polovině 20. století, vrcholem pak bylo období od 60. let. V 90. letech nastává zlom, který pokračuje dále ve 21. století a začíná snižování rozdílu průměrné délky života nově narozených žen a mužů. Daný pokles je způsoben snižující se pravděpodobností úmrtí u mužů ve všech věkových kategoriích, viz tabulka s koeficienty 2.1. Při zachování současného trendu v grafu 3.4 by se tak nově narození muži měli čím dál tím víc přibližovat průměrné délce života nově narozených žen. Přitom muži v mladších letech by klidně mohli mít nižší pravděpodobnost úmrtí, podobnou jako mají ženy. U mužů ji od 15 do zhruba 50 let nadmíru navyšují vnější vlivy jako nehody, úrazy či sebevraždy [15]. Například od roku 2004 spáchá v průměru sebevraždu čtyřikrát větší počet mužů, než žen [18].

Podívejme se, jestli se stejně jako na obrázku 3.4 snižuje i rozdíl v očekávaném počtu zbývajících let u osob například v 65 letech značené  $e_{65}$ . To je počet let, které má ještě v průměru žít osoba, která dosáhla věku 65 let. Analogicky si hodnoty vykresleme do grafu 3.5. Z obrázku už je na první pohled jasné, že křivka s rozdíly v 65 letech má jiný tvar připomínající logistický trend a samotné napočítané rozdíly nejsou tak velké, jako v případě předchozí situace, protože nezahrnují celý život právě narozených jedinců, kdy muži mají daleko větší pravděpodobnost úmrtí i v mladších letech oproti ženám, ale pouze už jen část života od 65 let. Také nedochází v období kolem 21. století k poklesu, trend je místo toho konstantní a v posledních letech začíná dokonce mírně stoupat. To je způsobeno buď nedostatečným poklesem pravděpodobnosti úmrtí u mužů ve vyšších věcích nebo pravděpodobněji dochází k většímu poklesu pravděpodobnosti úmrtí ve vyšších letech i u žen, jak je vidět v tabulce s koeficienty 2.1, a proto zde nedochází k sestupné tendenci rozdílů. Pokud by současný trend byl zachován i v budoucnu, tak spolu s kombinací trendu, který vidíme u  $e_0$ , to pro mladší populaci mužů prakticky znamená, že by se měl snižovat počet úmrtí na vnější vlivy a muži by se tak měli více dožívat pokročilejšího věku podobně jako ženy. Avšak

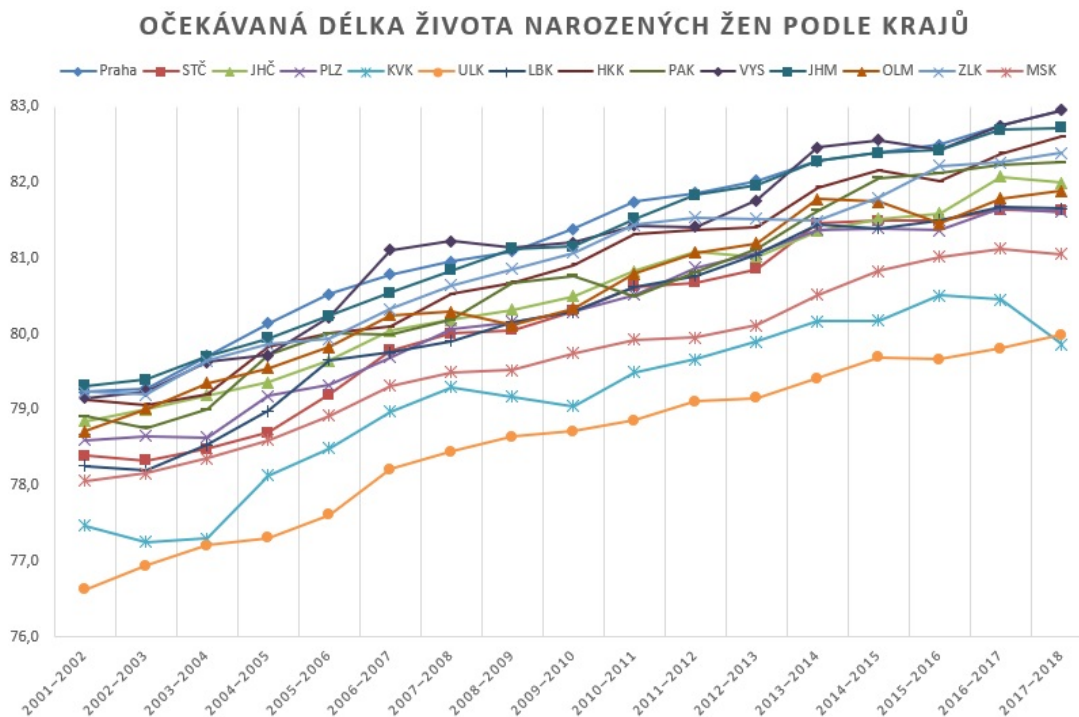
v důchodovém věku by se nic nezměnilo a ženy by se ve stáří pořád v průměru dožívaly vyšších let.

## 3.2. Analýza střední délky života v krajích

Zatím jsme jen zkoumali veškeré hodnoty z úmrtnostních tabulek napočítaných plošně za celou Českou republiku. Na začátku této práce jsme si řekli, že se konstruují i tabulky menších samosprávních celků. A těmi jsou například kraje nebo ještě menší okresy. Ostatně data za celou republiku můžeme roztrždit třeba podle krajů a veškeré ukazatele v úmrtnostní tabulce vypočítat jen v rámci jednoho kraje. Dostaneme tak přesnější hodnoty ukazatelů v rámci regionu. Vykresleme si tak hodnoty  $e_0$  pro všechny kraje vypočítané za dvouletí od roku 2001-2002 do roku 2017-2018 do grafu 3.6 pro muže a 3.7 pro ženy. Na obou grafech můžeme vidět dlouhodobý růst očekávané délky života ve všech krajích. Také je zřejmé, že v každém kraji se očekávaná délka života liší.



Obrázek 3.6: Očekávaná délka života narozených mužů v jednotlivých krajích.



Obrázek 3.7: Očekávaná délka života narozených žen v jednotlivých krajích.

Tím pádem lze říct, že počet zemřelých není v ČR rozložen rovnoměrně dle jednotlivých krajů, ale jsou zde kraje, kde se lidé v průměru dožívají více, respektive méně let, než v jiných krajích. U mužů můžeme vidět, že nejvyšší očekávaná délka života je v Praze, kde lze ale také očekávat dostupnější zdravotní péči [17]. Po větším odstupem následuje Královehradecký kraj s krajem Vysočina. U žen si lze všimnout, že očekávaný počet let života v Praze nemá až takový odstup od ostatních krajů. Podobně očekávaný počet let života jako v Praze lze pozorovat v kraji Vysočina nebo Jihomoravském kraji. Naopak nejmenší očekávaná délka života u obou pohlaví je v Ústeckém kraji. Před ním ještě s výraznou odchylkou od zbytku je Karlovarský kraj a Moravskoslezský kraj. Za zmínku také stojí výrazně viditelný pokles očekávané délky života u žen v Karlovarském kraji v roce 2018. S rozdílem  $-0,6$  let se jedná o největší meziroční pokles od počátku časové řady a také o největší pokles v daném roce mezi všemi kraji. Příčinou poklesu je, že v Karlovarském kraji byl v roce 2018 pozorován největší počet zemřelých z celé ČR na tisíc obyvatel. V tomto kraji byla také největší kojenecká úmrtnost z

celé ČR [16]. U mužů v Karlovarském ale kraji takový pokles pozorovaný nebyl. To se dá vysvětlit tím, že očekávaná délka života u mužů je nižší, a tak se vysoká úmrtnost moc neprojevila.

Zkusme vybrat například rok 2011 v časové řadě a zjistit, jestli souvisí střední délka života s některým ekonomickým ukazatelem, jakým je například výše průměrné měsíční mzdy v krajích za rok 2011 [21]. Budeme tedy počítat vztah mezi průměrnou měsíční mzdou a očekávanou délkou života v jednotlivých krajích v roce 2011. Uvidíme, jestli se prokáže tvrzení, že čím vyšší průměrná mzda je v kraji, tím je i větší průměrná délka života. Na odhalení vztahu mezi očekávanou délkou života a průměrnou mzdou použijeme Spearmanův korelační koeficient definovaný jako [2]:

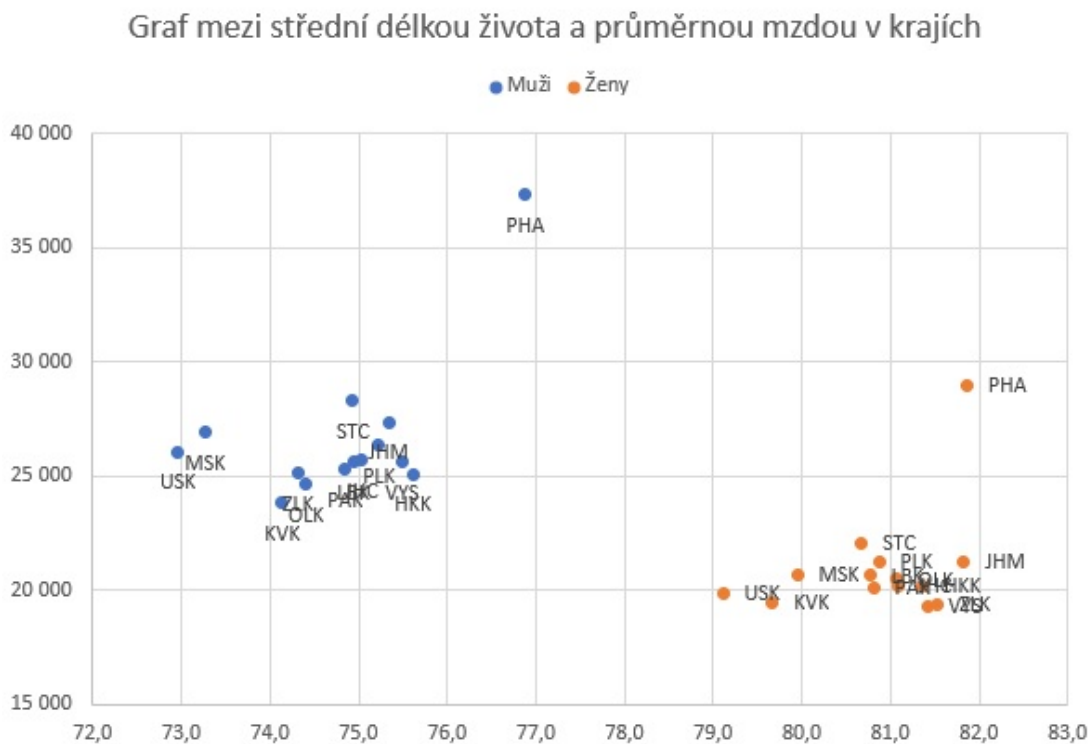
$$R_s = 1 - \frac{6}{n \cdot (n^2 - 1)} \cdot \sum_{i=1}^n (R_i - Q_i)^2, \quad (3.1)$$

kde  $R_i, i = 1, \dots, n$  představují pořadí náhodných veličin  $X_1, \dots, X_n$  a  $Q_i, i = 1, \dots, n$  jsou pořadí náhodných veličin  $Y_1, \dots, Y_n$ , jedná se tedy o čísla  $1, 2, \dots, n$ . Spearmanův korelační koeficient pracuje pouze s pořadími, a tak má výhodu v tom, že ho neovlivňují odlehlá pozorování. Také měří závislost dvou veličin, která nemusí být lineární, ale stačí nám obecný předpoklad monotonie. Korelační koeficient nabývá hodnot mezi  $\langle -1, 1 \rangle$ . Pokud koeficient nabývá kladných hodnot, bude v datech přímý vztah (s rostoucími hodnotami  $X$  roste i  $Y$ ), naopak pokud záporných hodnot, tak bude zde protichůdný vztah (s rostoucími hodnotami  $X$  klesá  $Y$  nebo naopak). Jestliže  $R_s = 0$ , pak jsou veličiny nekorelované. Také platí, že čím blíže jsme hodnotě  $-1$  nebo  $1$ , tím je vztah těsnější. Na základě realizace výpočtu Spearmanova korelačního koeficientu  $r_s$  lze otestovat, jestli jsou veličiny nekorelované. Testová statistika v případě Spearmanova korelačního koeficientu je jednoduchá. Stačí vzít pouze  $|R_s|$  a srovnat na hladině významnosti  $\alpha$  s příslušnou kritickou hodnotou, která je tabelována pro počet pozorování  $n < 30$  [2].

Korelační koeficient mezi průměrnou měsíční mzdou a  $e_0$  vyšel pro muže  $r_s = 0,266$  a pro ženy  $r_s = 0,138$ . Otestujme hypotézu, jestli průměrná mzda je nekorelovaná s očekávanou délkou života.

**Hypotéza 1.** *Střední délka života a průměrná měsíční hrubá mzda v krajích jsou nekorelované.*

Alternativou bude, že veličiny jsou korelované. Testování provedeme ve statistickém softwaru R, kde hladinu významnosti  $\alpha = 0,05$  srovnáme s p-hodnotou. Soubor s testem "mzda" i soubor s potřebnými podklady "mzdy-kraje", jsou uloženy na příloženém CD. Testování prokázalo, že v obou případech hypotézu nezamítáme. Střední délka života a průměrná měsíční mzda v krajích jsou nekorelované. Nelze tak tvrdit, že ženy, ani muži v krajích s vyšší průměrnou mzdou, mají také vyšší očekávanou délku života, což také lze vidět na obrázku 3.8.



Obrázek 3.8: Graf mezi průměrnou měsíční mzdou a střední délkou života v krajích.

Z grafu můžeme pozorovat také tři oddělené skupiny krajů. Nejvíce se u mužů odlišují Moravskoslezský a Ústecký kraj a ačkoliv to jsou kraje s nejkratší délkou

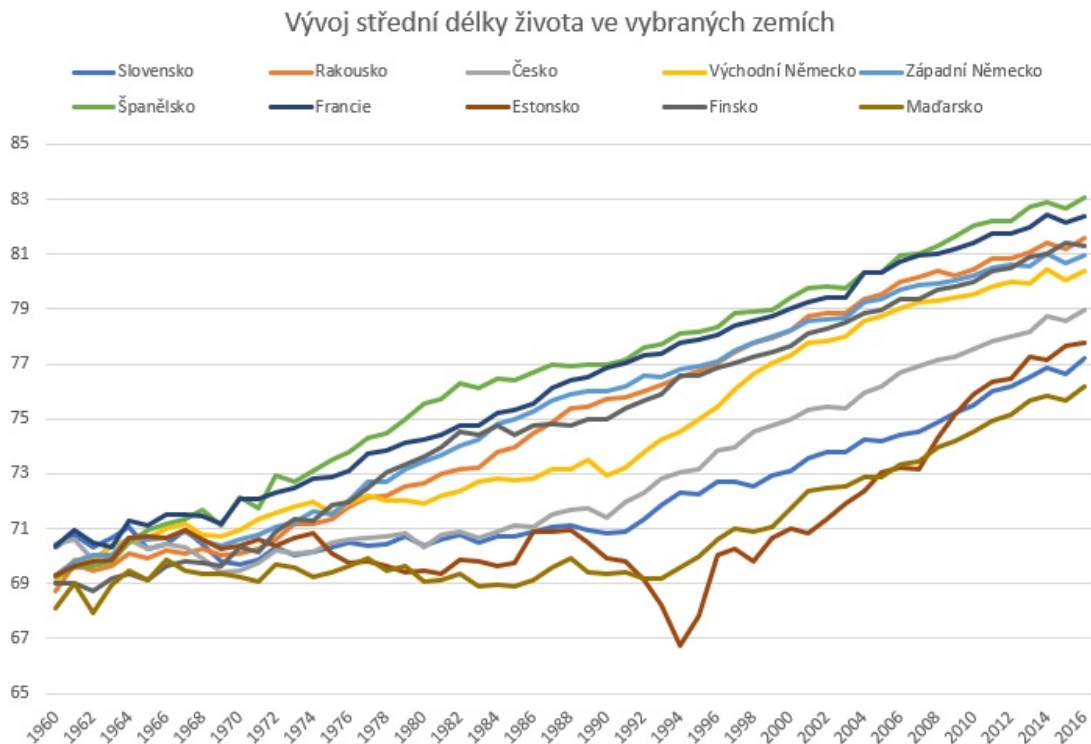


života, nejsou to kraje s nejnižšími příjmy. U žen bychom do této skupiny ještě zařadili Karlovarský kraj. Další samostatnou skupinu tvoří Praha, kde je nejvyšší délka života a příjmy u obou pohlaví. Do třetí skupiny spadají ostatní kraje, mezi kterými tak výrazné odchylky nepozorujeme.

### 3.3. Analýza střední délky života v zahraničí

V této části se podíváme na to, jak se vyvíjela očekávaná délka života v Česku ve srovnání s vybranými evropskými zahraničními zeměmi. Data pro toto srovnání vezmeme pro muže a ženy dohromady z internetové databáze Human mortality database [19]. Mezi vybranými zeměmi budou některé okolní státy sousedící s Českem jako Rakousko, Německo a Slovensko. V rámci Německa máme k dispozici data také zvlášť pro Východní Německo a Západní Německo. Také vyberme země, které jsou dál od Česka, například Španělsko, Francie, Estonsko, Maďarsko a Finsko. Údaje poskytované k jednotlivým zemím nejsou od jednoho časového okamžiku, proto zvolme jako začátek rok 1960 a časovou řadu ukončíme v roce 2016. Vykresleme si očekávanou délku života  $e_0$  ve zmíněných státech od roku 1960 do společného grafu 3.9.

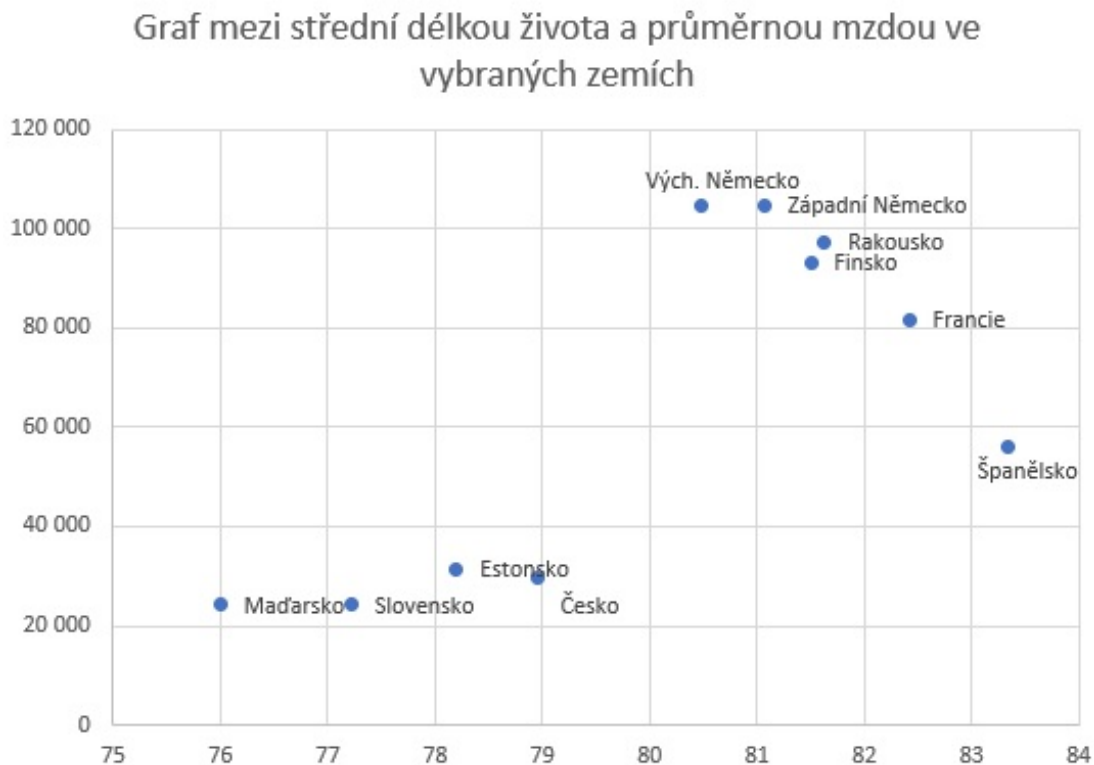
Z časové řady je vidět, že kolem roku 1960 ještě nebyl takový rozdíl v očekávaném dožití mezi jednotlivými státy. V Česku tehdy byla podobná délka života jako ve Francii či Španělsku. V dalším období v 70. letech, kdy ve východní a střední Evropě vládl komunistický režim a Česko, Slovensko, Východní Německo, Maďarsko a Estonsko (které v té době bylo ještě součástí Sovětského svazu), byly součástí východního bloku [20]. Je vidět, že očekávaná délka života v těchto státech v té době nerostla tolik, jako v zemích, které nebyly součástí onoho bloku. Příčinou je, že v kapitalistických zemích byl v té době uvědomělejší přístup ke zdraví, nedošlo k útlumu rozvoje zdravotnictví, a tak v těchto zemích klesala úmrtnost na kardiovaskulární nemoci a kladl se větší důraz na zdravější životní styl [13]. Ze zemí bývalého východního bloku je na tom nejlépe Východní Německo. Tam na začátku 70. let byla očekávaná délka života dokonce větší, než v Západním Německu, Rakousku a Finsku. Ke konci 70. let však nárůst v těchto zemích již přesáhl střední



Obrázek 3.9: Očekávaná délka života ve vybraných zemích v Evropě.

délku života ve Východním Německu, která se v dalších letech prodlužovala jen velmi zvolna. V Česku, Slovensku a Maďarsku byla v té době stagnace a na území Estonska byl zaznamenaný pokles. Zlom v zemích bývalého východního bloku přišel až s pádem komunismu a zlepšením životní úrovně v 90. letech. Ve Východním Německu nastává prudký nárůst střední délky života. Mezi lety 1991 až 2000 se průměrná délka života zvýšila o 4,42 let a Východní Německo se tak v délce života dokázalo po roce 2000 připojit ke státům západní Evropy. Avšak rozdíly mezi Východním a Západním Německem se zatím nesmazaly a převládají dodnes. Podobný trend jako ve Východním Německu nastává i v dalších postkomunistických zemích. V 90. letech a po roce 2000 přichází prudký růst očekávané délky života v Česku, Slovensku a Maďarsku. Za zmínku také stojí to, že v Česku se začala očekávaná délka života zvyšovat dříve, než na Slovensku a to ještě před rozdělením Československa, které přišlo až v roce 1993. Výjimkou v trendech je Estonsko, kde došlo mezi lety 1987 až 1995 k největšímu poklesu pozorovaném v

celé časové řadě a to celkově o  $-4,17$  let. Hned v dalších letech však v zemi začal strmý růst, při kterém Estonsko v průměrné délce života překonalo Slovensko a Maďarsko. Z hlediska celého období od roku 1960 do roku 2016 lze říct, že státy, které v historii nebyly součástí východního bloku, mají vyšší průměrnou délku života. Z vybraných států v grafu je to nejvíc ve Španělsku, Francii, Západním Německu, Rakousku a Finsku. V těchto zemích nedošlo od roku 1960 k žádnému prudkému nárůstu ani poklesu a trend, který by seděl pro daný vývoj, by tak nejspíš byl po celou dobu rostoucí lineární, zatímco v regionu střední Evropy, tedy v Česku, Slovensku a Maďarsku podobný trend nastává až po pádu komunismu, kdy dlouholetou stagnaci vystřídal prudký nárůst a převážně v Česku se začíná snižovat rozdíl v očekávaném počtu let života se západními státy.



Obrázek 3.10: Graf mezi střední délkou života a průměrnou měsíční mzdou v zahraničí.

Obdobně jako u krajů, tak i zde provedme výpočet korelačního koeficientu mezi

průměrnými mzdami přepočtenými na koruny [23] a očekávanou délkou života v jednotlivých státech pro rok 2017. Korelační koeficient vypočítáme stejným způsobem jako v předchozím případě, a to pomocí rovnice (3.1). Hodnota korelačního koeficientu je v tomto případě  $r_s = 0,584$ . Otestujme, že průměrná mzda není korelovaná s očekávanou délkou života v jednotlivých státech.

**Hypotéza 2.** *Střední délka života v jednotlivých státech a průměrná měsíční mzda v těchto státech jsou nekorelované.*

Alternativou k testované hypotéze bude, že průměrná mzda je korelovaná se střední délkou života. Testování proved'eme ve statistickém softwaru R ve stejném souboru, jako předchozí test pro kraje. Podklady pro testování jsou v souboru "mzdy-zahranici" uloženy na CD. Jestliže p-hodnotu = 0,07497 srovnáme s hladinou významnosti  $\alpha = 0,05$ , pak dojdeme k závěru, že hypotézu o nekorelovanosti mezi očekávanou délkou života a průměrnou mzdou ve vybraných zemích, nezamítáme. I když v tomto případě jsme velmi blízko jejímu zamítnutí, pořád nelze říct, že existuje korelovanost mezi průměrnou mzdou a očekávanou délkou života ve vybraných zemích. To potvrzuje i graf 3.10, na kterém vidíme, že bychom nemohli popsat lineární vztah mezi průměrnou mzdou a očekávanou délkou života mezi vybranými zeměmi žádnou funkcí. Na tomto grafu je také vidět oddělení postkomunistických zemí kromě Východního Německa od zbytku vybraných evropských států, které mají vyšší mzdy i očekávanou délku života. U těchto zemí lze dokonce pozorovat protichůdný vztah mezi očekávanou délkou života a průměrnou mzdou. Celkově však mezi všemi vybranými zeměmi neplatí tvrzení, že čím vyšší mzda, tím větší je očekávaná délka života.

# Závěr

V této práci jsme si na začátku vysvětlili, co je úmrtnostní tabulka, popsali jsme metodiku, kterou používá Český statistický úřad při konstrukci pravděpodobnosti úmrtí, jakož důležitého ukazatele, protože se od něj odvíjí další ukazatele v úmrtnostní tabulce, jejichž konstrukci a význam jsme si rovněž ukázali. Poté jsme přešli k samotné analýze pravděpodobnosti úmrtí, kdy bylo cílem vytvořit úmrtnostní model na predikci pravděpodobnosti úmrtí. Z dostupných úmrtnostních tabulek jsme si načetli historická data do časových řad. Zjistili jsme však, že tak dlouhá časová řada obsahuje spoustu historických výkyvů, a tak, abychom mohli namodelovat trend, který by neobsahoval vliv těchto výkyvů, museli jsme časové řady zkrátit. Zkrácením došlo k tomu, že jsme v časových řadách mohli pozorovat trend exponenciálního poklesu. Metodou nejmenších čtverců jsme odhadli parametry exponenciálních funkcí. Tímto jsme dokázali odhadnout budoucí hodnoty pravděpodobnosti úmrtí a zkonstruovat tak daný model. Cílem tohoto modelu pak bylo ukázat na příkladu pojištění důchodu z oblasti pojistné matematiky, jak se změní velikost pojištění důchodu, pokud k výpočtu použijeme zkonstruovaný model, který zachycuje budoucí odhad pravděpodobnosti úmrtí v čase. Výsledky plynoucí z použití modelu jsme rovněž porovnali s výpočtem při použití průřezové úmrtnostní tabulky. Z výsledků plyne, že při použití modelu dojde ke zdražení pojištění důchodu, protože na základě odhadnutých trendů bychom očekávali pokles pravděpodobnosti úmrtí, což obecně vede ke zdražení pojištění důchodu. Další na řadě byla analýza ukazatele střední délky života. Podívali jsme se, jak se liší rozdíl v Česku mezi nově narozenými ženami a muži a na základě předpokladu zachování trendů i v budoucnu jsme

dospěli k tomu, že by se měl rozdíl v očekávané délce života mužů a žen dále snižovat. Ovšem graf s rozdíly v 65 letech měl jiný trend, který neukazoval na snižování rozdílu ve starších letech, což vedlo k závěru, že při kombinaci a zachování obou trendů i v budoucnu by to pro mladší populaci mužů znamenalo, že se v průměru budou dožívat vyššího věku, než teď, ale v důchodovém věku by se pořád více let v průměru dožívaly ženy. Pak jsme se podívali na střední délku života v jednotlivých krajích. Vývoj ukázal, že největší délka života u mužů je v Praze, ale také na Vysočině nebo v Královéhradeckém kraji. Naopak nejméně v Ústeckém, Moravskoslezském a Karlovarském kraji. Ženy v Praze nemají takový odstup od zbytku krajů, jako mají muži. Podobně jako v Praze mají průměrnou délku života ženy na Vysočině a Jihomoravském kraji. Nejméně to bylo v Ústeckém kraji, následovaný Karlovarským, kde byl zaznamenán v roce 2018 největší meziroční pokles v časové řadě a dostal se tak v ten rok ještě pod Ústecký a dále je nízká očekávaná délka života také v Moravskoslezském kraji. Také jsme zjistili, že neplatí premisa, že čím větší je v kraji průměrná mzda, tím je větší očekávaná délka života. Jako poslední jsme porovnali vývoj střední délky života v Česku s vybranými zahraničními státy. Zjistili jsme, že státy, které dříve nebyly součástí tzv. východního bloku, mají vyšší očekávanou délku života. Ovšem po pádu komunismu se v postkomunistických státech rovněž začala prodlužovat očekávaná délka života, a tak pomalu dochází k poklesu rozdílů mezi východem a západem Evropy. Rovněž jsme odhalili nekorelovanost mezi průměrnou měsíční mzdou a střední délkou života ve vybraných státech.

# Literatura

- [1] Cipra T., *Pojistná matematika: Teorie a praxe*, Ekopress, Praha 2006
- [2] Hron K., Kunderová P., Vencálek O., *Základy počtu pravděpodobnosti a metod matematické statistiky*, Univerzita Palackého v Olomouci, Olomouc 2018
- [3] <https://www.czso.cz/csu/czso/umrtnostni-tabulky-metodika>
- [4] Nicolas Bacaër, *A short history of mathematical population dynamics*, Springer, Londýn 2010
- [5] Pavlačka O., *Přednášky z předmětu pojišťovnictví 1*
- [6] J. Seger, R. Hindls., *Statistické metody v tržním hospodářství*, Victoria Publishing, Praha, 1995
- [7] Eilers a Marks, 1996
- [8] Thatcher a kol., 1998
- [9] Chiang, 1984
- [10] Andreev Kingkade, 2015
- [11] [https://www.czso.cz/documents/10180/94500304/zprava\\_o\\_ocekavanem\\_+vyvoji.pdf/36c116ef-204f-4ee0-b7b8-f5c1a10a65e8?version=1.0](https://www.czso.cz/documents/10180/94500304/zprava_o_ocekavanem_+vyvoji.pdf/36c116ef-204f-4ee0-b7b8-f5c1a10a65e8?version=1.0)
- [12] <https://www.financnivzdelavani.cz/svet-financi/pojistovnictvi/jake-jsou-zakladni-pojistne-produkty/pojisteni-osob/duchodove-zivotni-pojisteni>
- [13] [http://www.demografie.info/?cz\\_umrtnosthistorie](http://www.demografie.info/?cz_umrtnosthistorie)
- [14] [http://oldwww.upol.cz/fileadmin/user\\_upload/knihovna/Skripta\\_FF/finan.pdf](http://oldwww.upol.cz/fileadmin/user_upload/knihovna/Skripta_FF/finan.pdf)
- [15] [https://www.irozhlas.cz/zpravy-domov/umrtnost-zdravotnictvi-nemoci-srdce-sebevrazdy-nehody\\_2001240600\\_jab](https://www.irozhlas.cz/zpravy-domov/umrtnost-zdravotnictvi-nemoci-srdce-sebevrazdy-nehody_2001240600_jab)
- [16] <https://www.kurzy.cz/zpravy/526580-zemreli-podle-pricin-umrti-v-usteckem-kraji-v-roce-2018-vydano-13-01-2020/>

- [17] <https://www.czso.cz/documents/10180/20542003/31003314sta.pdf/5ff48177-fd93-4996-aead-eb9f556252d8?version=1.0>
- [18] [https://www.czso.cz/csu/czso/sebevrazdy\\_zaj](https://www.czso.cz/csu/czso/sebevrazdy_zaj)
- [19] <https://www.mortality.org/>
- [20] <https://docplayer.cz/25135956-Vychodni-blok-staty-a-organizace-vychodniho-bloku.html>
- [21] [https://dk.upce.cz/bitstream/handle/10195/53587/SeibertovaP\\_Analyza\\_Vyvoje\\_PJ\\_2013.pdf;jsessionid=1FFD33E8A3D041FF8F6E16B47521FEE0?sequence=2](https://dk.upce.cz/bitstream/handle/10195/53587/SeibertovaP_Analyza_Vyvoje_PJ_2013.pdf;jsessionid=1FFD33E8A3D041FF8F6E16B47521FEE0?sequence=2)
- [22] <https://www.statista.com/statistics/263722/life-expectancy-in-spain/>
- [23] <https://www.finance.cz/509221-prumerna-mzda-vyssi-nez-100-tisic-kc/>