



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Marek Šešulka

Stochastické modely úmrtnosti v kontextu kvantifikace vybraných pojistněmatematických rizik

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Radek Hendrych, Ph.D.

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2022

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

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 má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Chtěl bych poděkovat RNDr. Radku Hendrychovi, Ph.D. za jeho odborné vedení, rady a profesionální přístup. Rád bych také poděkoval moji rodině za jejich podporu během mých studií. V neposlední řadě děkuji mým nejdražším kamarádům, kteří mi vždy pomohli, povzbudili a poradili během krušných časů. Markéto, Tereško, Dominiku, Stano s Vámi je studium a život lehčí!

Název práce: Stochastické modely úmrtnosti v kontextu kvantifikace vybraných pojistněmatematických rizik

Autor: Marek Šešulka

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Radek Hendrych, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce se věnuje stochastickým úmrtnostním modelům v kontextu pojistněmatematických rizik. V teoretické části definuje pět úmrtnostních modelů společně s přístupem k predikci. Zavádí testy, které mají za cíl ohodnocení a posouzení věrohodnosti predikcí. Dále představuje pojistněmatematická rizika vztahující se k úmrtnostní tematice. Zabývá se zajištěním rizika dlouhověkosti pomocí finančního instrumentu zvaného longevity dluhopis skrze Wangovu transformaci. V empirické části nejprve odhaduje, popisuje a interpretuje všechny modely na základě českých dat. Poté je provedeno vyhodnocení predikční modelů, pro obě pohlaví a dva zvolené věky. V poslední části práce je zpracováno ocenění longevity dluhopisu a s ním spojené ceny rizika.

Klíčová slova: longevity dluhopis, pojistněmatematická rizika, úmrtnostní model, zajištění

Title: Stochastic mortality models within the quantification of selected actuarial risks

Author: Marek Šešulka

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Radek Hendrych, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The thesis focuses on the stochastic mortality models in the context of actuarial risks. In the theoretical part, the thesis defines five mortality models with approaches to prediction. After that, it follows the description of selected actuarial risks in the context of mortality. Hedging of longevity risk is obtained through a financial instrument called a longevity bond. Pricing of that bond is delivered via the Wang transformation approach. In the empirical part of the thesis, there are carried out estimates, interpretations, and comments for models based on the Czech mortality data. Later, prediction tests are executed for an individual model, both sex and two chosen ages. The last part of the empiric section deals with the data driven pricing of longevity bond and the price of risk.

Keywords: actuarial risks, hedge, longevity bond, mortality model

Obsah

Úvod	3
1 Historický vývoj a motivace	4
2 Základní pojmy, definice, věty	5
2.1 Demografické pojmy a tvrzení	5
2.2 Časové řady	6
2.3 Zobecněný lineární a nelineární model	6
3 Stochastické úmrtnostní modely	8
3.1 Předmluva a značení	8
3.2 Podstata modelů a jejich třídy	8
3.2.1 Lee-Carterův model	8
3.2.2 Cairns-Blake-Dowdův model	10
3.3 Přístupy k předpovídání modelů	12
3.3.1 Predikce v Lee-Carterově modelu	13
3.3.2 Predikce v Cairns-Blake-Dowd modelu	14
3.4 Diagnostika modelů	16
3.4.1 Test konvergence zužujícího horizontu	17
3.4.2 Test dlouhé predikce	17
3.4.3 Test posouvajícího se okna predikce	18
4 Pojistněmatematická rizika v kontextu úmrtností	19
4.1 Riziko úmrtnosti a katastrofy	19
4.2 Riziko dlouhověkosti	20
4.3 Riziko úrovně	20
4.4 Riziko trendu	21
4.5 Selekční riziko	21
5 Zajištění rizika dlouhověkosti	22
5.1 Longevity dluhopis	22
5.2 Wangova transformace	24
6 Empirická část	26
6.1 Odhady modelů	27
6.1.1 Lee-Carterův model	27
6.1.2 Lee-Carterův model s kohortním členem	29
6.1.3 Cairns-Blake-Dowdův model	31
6.1.4 Cairns-Blake-Dowdův model s kohortním členem	33
6.1.5 Cairns-Blake-Dowdův model s variačním členem	35
6.2 Empirická diagnostika modelů	37
6.2.1 Test konvergence zužujícího horizontu	38
6.2.2 Test dlouhé predikce	44
6.2.3 Test posouvajícího se okna predikce	50
6.2.4 Shrnutí	55
6.3 Ocenění longevity dluhopisu	56

Závěr	64
Seznam použité literatury	66
Seznam obrázků	68
Seznam tabulek	72
A Přílohy	73
A.1 Odhady modelů a residuí pro ženskou populaci	73

Úvod

Stochastické úmrtnostní modely tvoří fundamentální základ v odhadování a predikci úmrtnosti. Jejich podrobnější výzkum si vyžaduje stále se zlepšující střední očekávaná doba dožití, která se za posledních několik dekad změnila způsobem, který by málokdo odhadoval. Pojišťovny, soukromé penzijní společnosti, ale i státní instituce musí správně odhadovat a predikovat vývoj úmrtnosti. Pomocí těchto odhadů stanovují pojistné, penzijní pilíře a analyzují stabilitu sociální ekonomiky. Věrohodné odhady úmrtnosti a středního očekávaného věku dožití jsou pro tyto subjekty klíčové.

Tato práce se zabývá stochastickými úmrtnostními modely v kontextu kvantifikace vybraných pojistněmatematických rizik. Nejprve uvedeme stručný historický vývoj úmrtnosti a motivaci k tomuto tématu. Následuje krátký přehled definic a tvrzení, o které se práce opírá v jednotlivých sekcích.

Ve třetí kapitole si ujasníme základní myšlenky ztvárňující naši představu o úmrtnosti. Následně zavedeme pět stochastických úmrtnostních modelů: Lee-Carterův model a jeho rozšíření o kohortní člen, Cairns-Blake-Dowdův model s jeho dvěma rozšířeními o kohortní a variační člen. Dále se zaměříme na přístupy k modelování predikcí. Tuto sekci zakončíme definicemi diagnostických testů, které nám poslouží jako hodnotící kritéria v empirické části práce.

Čtvrtá kapitola se zabývá pojistněmatematickými riziky v kontextu úmrtnosti. Nejprve si představíme rizika vycházející z takzvaného standardního vzorce, který slouží pro výpočet solventnostního rizikového kapitálu pojišťoven. Jedná se o rizika úmrtnosti, katastrofy a dlouhověkosti. Poté popisujeme rizika spojená přímo s odhadem modelů: riziko úrovně, trendu a selekční.

Pátá kapitola se věnuje zajištění rizika dlouhověkosti, které považujeme za nejvíce materiální pro životní pojišťovny, penzijní společnosti a penzijní státní fondy. Sekci zahájíme s definicí longevity dluhopisu a vysvětlením jeho principu a peněžních toků. Poté se práce zabývá přístupem ke stanovení spravedlivé ceny tohoto finančního kontraktu. K tomu je využita takzvaná Wangova transformace, která zavádí parametr ceny tržního rizika, který je klíčový pro stanovení férové ceny longevity dluhopisu.

Poslední šestá kapitola je empirickou částí práce. Aplikujeme zde vybrané teoretické sekce na českých datech úmrtnosti. Začneme s odhadem všech stochastických úmrtnostních modelů, kde popisujeme jednotlivé odhady parametrů a proložení skutečných realizací pro mužskou populaci. Věnujeme se také základní diagnostice residuí. V další části provádíme predikční testy na jednotlivé modely. Vždy pro obě pohlaví a pro věky 65 a 85. Pro každý test a pohlaví uvádíme tabulku úspěšnosti daných modelů. V poslední části oceňujeme longevity dluhopis pomocí Wangovy transformace a tržní ceny rizika. Dodáváme rozpady spravedlivé ceny kontraktu pro pojišťovnu, investora a odhadnutou cenou tržního rizika.

1. Historický vývoj a motivace

Modelování a předpovídání úmrtnosti je v historii matematiky poměrně mladé téma. Zájem o podrobnější zkoumání tohoto tématu vzešel do popředí koncem minulého století. Počátky měření a analýzy úmrtnosti se však datují v poměrně daleké minulosti, a to konkrétně do druhé poloviny 17. století, kdy John Graunt přednesl *Natural And Political Observations Made upon the Bills of Mortality*. Jeho projev se považuje za základ moderní představy o chování úmrtnosti.

Motivace k podrobnějšímu bádání, které začalo koncem 20. století, pramení z poměrně strmého globálního nárůstu očekávané doby dožití novorozenců, který je důsledkem zkvalitňování zdravotní péče a obecným zájmem o zdravější životní styl. Tento pozitivní fakt prodlužování věku dožití s sebou přináší i problematické výzvy. Zejména pak nutnost podrobnějšího výzkumu a modelování úmrtností, a to hlavně z důvodu penzijních politik, jež tvoří hlavní sociální pilíř moderních ekonomik. Pojišťovny či penzijní fondy, ať státní nebo soukromé, jsou zatíženy problémy udržitelnosti z důvodu významného růstu věku dožití, se kterým nepočítaly při kalkulaci spravedlivých pojistných částek několik dekad zpět. Tehdejší představa o očekávaném věku dožití značně podhodnocuje dnešní realitu. Problém těchto systémů pochází z nepřesnosti extrémně dlouhé predikce očekávaného věku dožití, konkrétně znatelné podhodnocení skutečné realizace. Každý rok, který konkrétní pojištěné kohorty žijí navíc oproti očekávání, vytváří obrovské finanční ztráty institucím, které takovýto závazek přijaly. Závažnost těchto finančních ztrát, jež dosahují gigantických rozměrů, motivují k podrobnějšímu zkoumání vývoje a predikce úmrtnosti.

Hlavním cílem této práce je představit několik úmrtnostních modelů, popsat jejich vlastnosti a interpretovat dané parametry. Následně modely otestovat pomocí stanovených testů predikce na základě českých historických dat. Porovnat jejich výsledky a zkoumat zdroje jejich nepřesnosti. Práce si také klade za cíl představení vybraných pojistněmatematických rizik a zaměření se na nejvíce materiální riziko v kontextu úmrtnosti při výplatě doživotních důchodů, a to jest riziko dlouhověkosti.

2. Základní pojmy, definice, věty

Tato kapitola slouží pro zavedení základního značení, definic a vět, které budeme používat v celé práci. První část je věnována demografii, poté následuje krátký úsek k časovým řadám a na konci si shrneme zobecněný lineární a nelineární model.

2.1 Demografické pojmy a tvrzení

Zde si zadefinujeme základní pojmy a značení, které budeme používat v celé práci. Začneme úvodním značením demografických náhodných veličin a pravděpodobností. Poznamenejme, že pojmy a značení čerpáme převážně z Cipra (2006). Uvažujme žijící osobu ve věku x , kterou budeme označovat zkráceně jako věk x . Zavedme:

- T_x označuje náhodnou veličinu zbývající délky života věku x .
- Pravděpodobnost, že věk x zemře do $t > 0$ let, značíme jako ${}_tq_x := P(T_x \leq t)$.
- Pravděpodobnost, že věk x přežije alespoň $t > 0$ let, značíme jako ${}_tp_x := P(T_x > t)$.
- Očekávanou střední dobou dožití věku x rozumíme ${}^0e_x = \mathbb{E}[T_x]$.

Intenzitu úmrtnosti věku $x + t$ pro náhodnou veličinu věku dožití T_x definujeme jako:

$$\mu_{x,t} = \lim_{dt \rightarrow 0^+} \frac{P[T_x \leq t + dt | T_x > t]}{dt}. \quad (2.1)$$

Centrální úmrtnostní míru definujeme jako:

$$m_x = \frac{D_x}{N_x}, \quad (2.2)$$

kde D_x je náhodná veličina počtu smrtí věku x a N_x je náhodná veličina počtu žijících věků x . Poznamenejme, že m_x se vztahuje vždy ke konkrétnímu kalendářnímu roku.

Teoretické modely typicky definujeme pomocí spojité intenzity úmrtnosti. Pro účely empirických odhadů z reálných dat musíme tuto spojitou míru aproximovat diskrétní náhodnou veličinou. Platí následující aproximace:

$$m_x \simeq \mu_{x+1/2}. \quad (2.3)$$

Odvození tohoto vztahu lze nalézt v Pollard (1972). Této aproximace využijeme v kapitole 6. *Empirická část*. Dodejme, že pro účely přehlednosti budeme intenzitu úmrtnosti značit jako $\mu_{x,y}$, kde y je daný kalendářní rok. Formálně tímto zápisem rozumíme intenzitu úmrtnostní v půli roku a tedy i věku: $\mu_{x+1/2,y+1/2}$.

Dle Brouhns a kol. (2002) zavedeme předpoklad, že náhodná veličina D_x má Poissonovo rozdělení:

$$D_x \sim Po(E_x^c \mu_x), \quad (2.4)$$

kde E_x^c je centrální expozice věků x . Tento předpoklad pro nás bude esenciální při odhadování modelů pomocí metod zobecněných (ne)lineárních modelů.

Posledním vztahem této sekce je takzvaný princip ekvivalence, kterým rozumíme rovnicí:

$$\mathbb{E}PV(\text{pojistné}) = \mathbb{E}PV(\text{pojistné plnění}).$$

Z této rovnice vychází pojistitel při netto oceňování pojistných smluv a stanovení spravedlivého netto pojistného pramenící z povahy konkrétního pojištění.

2.2 Časové řady

Pro účely predikce vybraných stochastických modelů budeme využívat koncept časových řad, které si nyní zdefinujeme. Definice čerpáme z Prášková a Lachout (1998).

Normálním bílým šumem rozumíme náhodný proces $\{\varepsilon_t, t \in \mathbb{Z}\}$ nekorelovaných náhodných veličin s nulovou střední hodnotou a nenulovým konečným rozptylem pocházející z normálního rozdělení.

Náhodný proces $\{X_t, t \in \mathbb{Z}\}$ splňuje ARIMA(p, d, q) model, platí-li:

$$\Delta^d X_t + a_1 \Delta^d X_{t-1} + \dots + a_p \Delta^d X_{t-p} + \mu = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}, \quad t \in \mathbb{Z}, \quad (2.5)$$

kde $\{\varepsilon_t\}$ je normální bílý šum, $\mu \in \mathbb{R}$, a_1, \dots, a_p a b_1, \dots, b_q , jsou reálné konstanty, $a_p \neq 0$, $b_q \neq 0$ a $\Delta^d X_t$ reprezentuje d -tou diferenci procesu X_t .

2.3 Zobecněný lineární a nelineární model

Také budeme v další kapitole potřebovat definici a řešení takzvaného zobecněného lineárního modelu. Vycházíme ze článku Nelder a Wedderburn (1972). Řekneme, že data (Y_i, X_i) splňují zobecněný lineární model (GLM), pokud:

1. Y_1, \dots, Y_n jsou nezávislá a rozdělení Y_i závisí na X_i skrze regresní koeficienty $\beta = (\beta_1, \dots, \beta_p)^T$.

2. Podmíněná hustota Y_i při daném X_i se dá zapsat ve formátu:

$$f(y, \theta_i, \phi) = \exp\left(\frac{y\theta_i - b(\theta_i)}{\phi} + c(y, \phi)\right),$$

kde $b(\cdot)$ je známá dvakrát diferencovatelná funkce a $\phi > 0$ je konstanta.

3. Parametr θ_i závisí na X_i a β skrze lineární prediktor $\eta_i := X_i^T \beta$.

4. Existuje známá striktně monotonní dvakrát spojitě diferencovatelná linková funkce g taková, že $g(\mu_i) = \eta_i$, kde $\mu_i = \mathbb{E}[Y_i|X_i]$.

Nechť Y_1, \dots, Y_n jsou nezávislá pozorování, taková že $Y_i \sim Po(\lambda_i)$. Nechť λ_i závisí na X_i skrze identitu $\log \lambda_i = X_i^T \beta$. Poissonovo rozdělení s parametrem $\lambda_i > 0$ je exponenciálního typu, tj. splňuje druhý bod předchozí definice, s $\mu_i = \lambda_i, \theta_i = \log(\lambda_i), \phi = 1, b(\theta_i) = e^{-\theta_i}$, neboť platí:

$$f(y, \theta_i, \phi) = \exp(y \log(\lambda_i) - \lambda_i - \log(y!)) = \frac{\lambda_i^y}{y!} e^{-\lambda_i}$$

Linková funkce jest logaritmus: $g(\mu_i) = X_i^T \beta$. Odhad parametru vychází z maximálně věrohodného odhadu a metody iterativních vážených nejmenších čtverců.

Log-věrohodnost pro Poissonovo rozdělení z n pozorování jest:

$$L_n(\beta|Y_1, \dots, Y_n) = \sum_{i=1}^n Y_i \log(\lambda_i) - \lambda_i - \log(Y_i!).$$

Podle Kulich (2021) vede tato log-věrohodnost skrze metodu iterativních vážených nejmenších čtverců k řešení ve tvaru:

$$\hat{\beta}_n^{(k+1)} = (\mathbf{X}^T \hat{\mathbf{W}}^{(k)} \mathbf{X})^{-1} (\mathbf{X}^T \hat{\mathbf{W}}^{(k)} \hat{\mathbf{Z}}^{(k)}), \quad (2.6)$$

kde

$$\hat{\mathbf{W}}^{(k)} = \text{diag}(\hat{\lambda}_1^{(k)}, \dots, \hat{\lambda}_n^{(k)}),$$

$$\hat{\mathbf{Z}}_i^{(k)} = X_i^T \hat{\beta}^{(k)} + \frac{Y_i - \hat{\lambda}_i^{(k)}}{\hat{\lambda}_i^{(k)}}.$$

Odvození odhadu včetně použitého algoritmu je možné nalézt v Kulich (2021).

(GLM) zobecňuje striktní lineární model a přidává linkovou funkci, umožňuje obecnější hustotu odezvy, nicméně stále vyžaduje, aby η byl lineární prediktor. Pro účely odhadu Lee-Cartetova (LC) modelu si nevystačíme s lineárním vztahem, proto definujme Zobecněný nelineární model (GNM), jako (GLM), kde třetí podmínku nahradíme následovně:

3. Parametr θ_i závisí na X_i a β skrze nelineární funkci $\eta_i(X_i, \beta)$.

Takto definovaný model umožňuje, aby prediktory byly například v multiplikačním tvaru, čehož využijeme v (LC) i v jeho rozšíření (LC-C). Řešení (GNM) využívá také algoritmus iterativních vážených nejmenších čtverců, s použitím další metod jako je QR dekompozice, singulární rozklad i Gauss-Newtonova metoda. Práce Lee a Carter (1992) používá pro odhad (LC) metodu singulárního rozkladu.

3. Stochastické úmrtnostní modely

V této kapitole se budeme věnovat stochastickým úmrtnostním modelům. Nejprve si ujasníme základní myšlenky a značení. Následně si zadefinujeme primární modely a jejich jednotlivá rozšíření. Poté budeme rozebírat přístupy k predikci u konkrétních modelů a v poslední sekci si zavedeme diagnostické testy. Tato kapitola je teoretickým fundamentem pro kapitolu 6. *Empirická část.*

3.1 Předmluva a značení

Pro účel důkladného pochopení esence stochastických úmrtnostních modelů je třeba si nejprve rozebrat základní myšlenku jejich významu. Podstata takovýchto modelů je nalezení vhodného vztahu mezi věkem, časem a úmrtností. Apriorně předpokládáme, že hlavními faktory vysvětlujícími úmrtnost jsou věk a čas. Existují také další faktory určující úmrtnost, avšak výše uvedené jsou bezesporu těmi nejdůležitějšími. Základní otázka, kterou chceme, aby nám tyto modely odpověděly, je následující: Jaký je nejlepší odhad úmrtnosti v závislosti na vstupních parametrech a jaká je predikce úmrtnosti do budoucnosti?

V celé práci budeme pracovat s roky a věky. Konzistentně budeme značit věky proměnou x a kalendářní roky písmenem y . Uvažujeme tedy hypotetickou datovou sadu s věky: x_1, \dots, x_{n_x} a roky y_1, \dots, y_{n_y} . Celkem tudíž máme n_x věků a n_y roků. Než si představíme jednotlivé modely, tak je třeba se rozhodnout pomocí jaké veličiny chceme měřit úmrtnost. Jedním z nejběžnějších přístupů je užití intenzity úmrtnosti $\mu_{x,y}$, kterou jsme si definovali v (2.1). Naším cílem v následující sekci bude zkoumat modely, které budou mít na levé straně intenzitu úmrtnosti a na pravé straně parametry vysvětlující úmrtnost. Poznamenejme, že u každého modelu budeme odhadovat logaritmickou transformaci intenzity úmrtnosti.

3.2 Podstata modelů a jejich třídy

Za první stochastický model úmrtnosti se považuje Lee-Carterův model publikovaný v práci Lee a Carter (1992). Na něhož navazuje několik modifikací a také další příbuzné modely. Některé z nich budou prezentovány v této podkapitole, která vychází primárně z Macdonald a kol. (2018).

3.2.1 Lee-Carterův model

Lee-Carterův model je formulován následovně:

$$(LC) : \log(\mu_{x,y}) = \alpha_x + \beta_x \kappa_y, \quad x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y},$$

kde α_x je absolutní člen reprezentující efekt věku, součin $\beta_x \kappa_y$ představuje efekt roku a věku v multiplikativním vztahu. Takto definovaný (LC) model nemá jednoznačné určené řešení. Proto je třeba zavést podmínky lokace a měřítka, například dle Lee a Carter (1992) následovně:

$$\sum_{i=1}^{n_x} \beta_{x_i} = 1, \quad \sum_{i=1}^{n_y} \kappa_{y_i} = 0. \quad (3.1)$$

Tento model se dá řešit vícero metodami, například skrze singulární rozklad. Nicméně, my využijeme zobecněný nelineární model (GNM), neboť druhý člen je nelineární. Tato metoda odhadne (LC) model pomocí přístupu maximální věrohodnosti, avšak tento odhad nemusí splňovat podmínky lokace a měřítka (3.1). Splnění těchto podmínek lze docílit pomocí reparametrizace odhadnutých koeficientů. Necht $\hat{\alpha}_E, \hat{\beta}_E$ a $\hat{\kappa}_E$ jsou libovolné odhady (LC) modelu metodou (GNM). Definujme:

$$\tilde{\kappa} = \frac{1}{n_y} \sum_{i=1}^{n_y} \hat{\kappa}_{y_i, E}, \quad \tilde{\beta} = \frac{1}{n_x} \sum_{i=1}^{n_x} \hat{\beta}_{x_i, E}. \quad (3.2)$$

Pak koeficienty definované následovně:

$$\begin{aligned} \hat{\alpha} &= \hat{\alpha}_E + \tilde{\kappa} \hat{\beta}_E \\ \hat{\kappa} &= n_x \tilde{\beta} (\hat{\kappa}_E - \tilde{\kappa} \mathbf{1}) \\ \hat{\beta} &= \frac{\hat{\beta}_E}{n_x \tilde{\beta}}, \end{aligned} \quad (3.3)$$

zřejmě splňují podmínky lokace a měřítka (3.1) a platí:

$$\hat{\alpha} + \hat{\kappa} \hat{\beta} = \hat{\alpha}_E + \hat{\beta}_E \hat{\kappa}_E.$$

Pro tento základní model existuje několik rozšíření. Jedním z nejdůležitějších je adaptace modelu pro reflexi takzvaného efektu kohorty. Na přelomu tisíciletí vyšlo na toto téma několik prací a článků, které tento efekt představují. Například práce CMI (2002), která ukazuje významnost kohortního efektu v modelování úmrtnosti. Efekt kohorty v modelování úmrtnosti se dá popsat jako zvláštní chování skupiny jedinců, jež jsou narozeni ve stejném roce oproti jedincům narozených v rocích předchozích nebo následujících. Jinými slovy je to speciální charakteristika jednotlivých generací. Existence efektu kohorty se vysvětluje více faktory charakteristické pro konkrétní generace. Typickým příkladem jsou kohorty narozené během druhé světové války, u kterých je zřejmé specifické chování, které je odlišuje od před a poválečných dob.

Lee-Carterův model obohacený o efekt kohorty můžeme definovat dle Dowd a kol. (2010) následovně:

$$(LC - C) : \log(\mu_{x,y}) = \alpha_x + \beta_x \kappa_y + \beta_x^{(2)} \gamma_c, \quad x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y},$$

kde γ_c je efekt kohorty a $c = y - x$ je rok narození dané kohorty. Obdobně jako u základního (LC) modelu musíme definovat podmínky jednoznačnosti, například jako:

$$\sum_{i=1}^{n_x} \beta_{x_i} = 1, \quad \sum_{i=1}^{n_x} \beta_{x_i}^{(2)} = 1, \quad \sum_{i=1}^{n_y} \kappa_{y_i} = 0, \quad \sum_{c=y_1-x_{n_x}}^{y_{n_y}-x_1} \gamma_c = 0. \quad (3.4)$$

Řešení modelu (LC-C) opět vychází z přístupu (GNM), ale odhadnuté koeficienty nemusí splňovat výše definované podmínky jednoznačnosti, a proto znovu přistoupíme k reparametrizaci. Necht $\hat{\alpha}_E, \hat{\beta}_E, \hat{\beta}_E^{(2)}, \hat{\kappa}_E$ a $\hat{\gamma}_E$ jsou libovolné odhady (LC-C) modelu metodou (GNM). Definujme:

$$\tilde{\gamma} = \frac{1}{n_y + n_x - 1} \sum_{c=y_1-x_{n_x}}^{y_{n_y}-x_1} \hat{\gamma}_{c,E}, \quad \tilde{\beta}^{(2)} = \frac{1}{n_x} \sum_{i=1}^{n_x} \hat{\beta}_{x_i,E}^{(2)}, \quad (3.5)$$

$\tilde{\kappa}$ a $\tilde{\beta}$ mají tvar (3.2) Pak koeficienty definované následovně:

$$\begin{aligned} \hat{\alpha} &= \hat{\alpha}_E + \tilde{\kappa} \hat{\beta}_E + \tilde{\gamma} \hat{\beta}_E^{(2)}, \\ \hat{\kappa} &= n_x \tilde{\beta} (\hat{\kappa}_E - \tilde{\kappa} \mathbf{1}), \\ \hat{\beta} &= \frac{\hat{\beta}_E}{n_x \tilde{\beta}}, \\ \hat{\beta}^{(2)} &= \frac{\hat{\beta}_E^{(2)}}{n_x \tilde{\beta}^{(2)}}, \\ \hat{\gamma} &= n_x \tilde{\beta}^{(2)} (\hat{\gamma}_E - \tilde{\gamma} \mathbf{1}), \end{aligned} \quad (3.6)$$

zřejmě splňují podmínky lokace a měřítka (3.4) a platí:

$$\hat{\alpha} + \hat{\beta} \hat{\kappa} + \hat{\beta}^{(2)} \hat{\gamma} = \hat{\alpha}_E + \hat{\beta}_E \hat{\kappa}_E + \hat{\beta}_E^{(2)} \hat{\gamma}_E.$$

3.2.2 Cairns–Blake–Dowdův model

Další ze známých stochastických modelů úmrtnosti je takzvaný Cairns-Blake-Dowdův model, jenž se dá definovat dle Macdonald a kol. (2018) následovně:

$$(CBD) : \log(\mu_{x,y}) = \kappa_y^{(1)} + \kappa_y^{(2)}(x - \bar{x}), \quad x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y}.$$

Povšimneme si, že tento model popisuje intenzitu úmrtnosti jako lineární funkci dvou parametrů: absolutního a lineárního příslušného rozdílu daného věku a průměrného věku. Poznamenejme, že pro pevné y představuje model věku x takzvaný Gompertzův zákon úmrtnosti. Díky jednoduchosti (CBD) modelu je možné ho odhadnout pomocí zobecněného lineárního modelu (GLM), který oproti (GNM) nevyžaduje dodatečné podmínky jednoznačnosti. Všechny různých kombinací x a y je celkem $n_x \times n_y$, tím pádem vektor odezvy a modelová matice pro model (CBD) mají právě $n_x \times n_y$ řádků. Označme:

$$\boldsymbol{\kappa}^{(k)} := (\kappa_{y_1}^{(k)}, \dots, \kappa_{y_{n_y}}^{(k)})^T, \quad k = 1, 2,$$

pak $\boldsymbol{\kappa} := (\boldsymbol{\kappa}^{(1)T}, \boldsymbol{\kappa}^{(2)T})^T$ je $2n_y$ vektor neznámých parametrů. Modelovou matici (CBD) označme jako $\mathbf{X} := [\mathbf{X}_1, \mathbf{X}_2]$, kde $\mathbf{X}_1, \mathbf{X}_2$ jsou blokově diagonální matice dimenze $(n_x \times n_y) \times n_y$. Jednotlivé bloky obou matic jsou tvořeny vektory o délce n_x následovně:

$$\mathbf{X}_1 = \mathbf{I}_{n_y} \otimes \mathbf{1}_{n_x} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} x_1 - \bar{x} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ x_{n_x} - \bar{x} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & x_1 - \bar{x} \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & x_{n_x} - \bar{x} \end{bmatrix}.$$

(CBD) můžeme tedy přepsat následovně:

$$\log(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\kappa},$$

kde $\boldsymbol{\mu} = (\mu_{x_1, y_1}, \mu_{x_2, y_1}, \dots, \mu_{x_{n_x}, y_1}, \dots, \mu_{x_{n_x}, y_{n_y}})^T$. Řešení předchozí rovnice vychází z iterativních vážených nejmenší čtverců a je ve tvaru (2.6), kde $\boldsymbol{\beta} = \boldsymbol{\kappa}$.

I tento model se dá jednoduše rozšířit o efekt kohorty, například jako:

$$(CBD - C) : \log(\mu_{x,y}) = \kappa_y^{(1)} + \kappa_y^{(2)}(x - \bar{x}) + \gamma_c, \\ x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y},$$

kde $c = y - x$. Modelová matice tohoto modelu je tvaru $\mathbf{X}' := [\mathbf{X}, \mathbf{X}_3]$, kde \mathbf{X} je modelová matice příslušná (CBD). Matice \mathbf{X}_3 je typu $(n_x \times n_y) \times (n_y + n_x - 1)$, neboť celkový počet kohort pro věky n_x a roky n_y jest právě $(n_y + n_x - 1)$. Matice \mathbf{X}_3 má v prvním sloupci jednu jednotku na pozici $[n_x]$, ve druhém sloupci má dvě jednotky na pozicích $[n_x - 1]$ a $[2n_x]$, iterativně má v n_x -tém sloupci k krát jednotku. Od sloupce n_x do sloupce n_y obsahuje každý sloupec matice přesně n_x jednotek od sloupce $n_y + 1$ začne počet jednotek klesat až k posledním sloupci, kde je opět jen jedna jednotka. Tedy formálně k -tý sloupec má jednotku na pozicích:

$$[n_x - k + 1], [n_x \times 2 - (k - 1) + 1], \dots, [n_x \times k], 1 \leq k \leq n_x - 1, \\ [n_x \times (k + 1 - n_x) + 1], [n_x \times (k + 2 - n_x) + 2], \dots, [n_x \times (k) + n_x], n_x \leq k \leq n_y, \\ [n_x \times n_y + 1], [n_x \times (k + 2 - n_x) + 2], \dots, [n_x \times (k) + n_x], n_y \leq k \leq n_y + n_x - 1.$$

Graficky je možné si \mathbf{X}_3 představit jako:

$$\mathbf{X}_3 = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \cdots \\ 0 & \ddots & \ddots & \ddots & 0 & \cdots \\ 1 & \ddots & \ddots & \ddots & 1 & \cdots \\ 0 & \ddots & \ddots & \ddots & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \cdots \\ 0 & 1 & 0 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Máme-li modelovou matici \mathbf{X}' , pak můžeme obdobně přepsat (CBD-C) jako:

$$\log(\boldsymbol{\mu}) = \mathbf{X}'\boldsymbol{\beta},$$

kde $\boldsymbol{\beta} := (\boldsymbol{\kappa}^T, \boldsymbol{\gamma}^T)^T$. Řešení dostane opět skrze (GLM) stejně jako pro (CBD).

Složitější pěti-parametrické rozšíření (CBD) modelu o efekt kohorty definujeme jako:

$$(CBD - V) : \log(\mu_{x,y}) = \kappa_y^{(1)} + \kappa_y^{(2)}(x - \bar{x}) + \kappa_y^{(3)}|(x - \bar{x}) - \varsigma_x| + \gamma_c, \\ x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y},$$

kde ς_x je směrodatná odchylka věku. Takto definovaný model reflektuje nejen kohortní efekt, který je typicky významný při modelování úmrtností, ale zohledňuje také odchylky věku od průměru poníženou směrodatnou odchylkou. Tento člen se empiricky ukazuje jako statistický významný.

Modelovou matici sestavíme jako $\mathbf{X}'' := [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_{2var}, \mathbf{X}_3]$, kde $\mathbf{X}_1, \mathbf{X}_2$ a \mathbf{X}_3 jsou stejné jako v (CBD), respektive v (CBD-C). Matice $\mathbf{X}_{2var} := \mathbf{I}_{n_y} \otimes \mathbf{x}_{2var}$ má dimenzi $(n_x \times n_y) \times n_y$, kde

$$\mathbf{x}_{2var} := (|(x_1 - \bar{x}) - \varsigma_x|, \dots, |(x_{n_x} - \bar{x}) - \varsigma_x|)^T.$$

Identicky jako v předchozích případech zapíšeme (CBD-V) následovně:

$$\log(\boldsymbol{\mu}) = \mathbf{X}''\boldsymbol{\beta},$$

kde $\boldsymbol{\beta} := (\boldsymbol{\kappa}_1^T, \boldsymbol{\kappa}_2^T, \boldsymbol{\kappa}_3^T, \boldsymbol{\gamma}^T)^T$ a opět řešíme skrze (GLM).

3.3 Přístupy k předpovídání modelů

Esenciální vlastnost, kterou požadujeme od úmrtnostních modelů, jest poskytnout věrohodný odhad budoucích realizací úmrtnosti. Zde si ujasníme, jak budeme přistupovat k predikci u všech modelů definovaných výše. Pomocí metod zavedených v této sekci budeme zkoumat empirické vlastnosti predikcí jednotlivých modelů v poslední kapitole.

Pro účel predikce stochastických modelů zvolíme přístup modelování pomocí časových řad, konkrétně podle třídy (více rozměrných) modelů ARIMA(p, d, q) pro vhodně zvolené parametry (p, d, q).

3.3.1 Predikce v Lee-Carterově modelu

Idea za predikcí (LC) modelu je pro každé x zafixovat parametry α_x, β_x jako odhady, které jsme získali, následně vhodně přeparametrizovali, z metody (GNM), a predikovat pouze parametr měnící se s y , tedy κ_y . Nejčastějším přístupem je modelovat κ_y jako náhodnou procházku s driftem, která je speciálním případem ARIMA(p, d, q) modelu, definovaného v (2.5), kde $p = 1$ a $d = 0, q = 0$. Položme:

$$\kappa_y = \mu + \kappa_{y-1} + \varepsilon_y, \quad y = 2, \dots, n_y, \quad (3.7)$$

kde ε_y je normální bílý šum, tedy ε_y je náhodný výběr z $N(0, \sigma^2)$ a σ, μ jsou neznámé parametry. Odhad takto definovaného modelu je přímočarý. Definujme náhodnou posloupnost prvních rozdílů modelu ($\Delta\kappa_y$):

$$\begin{aligned} \Delta\kappa_y &= \kappa_y - \kappa_{y-1} & y &= 2, \dots, n_y \\ &= \mu + \varepsilon_y. \end{aligned}$$

Zřejmě platí:

$$\mathbb{E}[\Delta\kappa_y] = \mu + 0, \quad \text{Var}[\Delta\kappa_y] = \text{Var}[\varepsilon_y] = \sigma^2.$$

Protože je (ε_y) je nezávislá posloupnost, pak $(\Delta\kappa_y)$ je náhodný výběr z $N(\mu, \sigma^2)$. Tím pádem nestranný a konzistentní odhad parametru μ je výběrový průměr a pro parametr σ^2 je to výběrový rozptyl ($\Delta\kappa_y$):

$$\hat{\mu} = \frac{1}{n_y - 1} \sum_{y=2}^{n_y} \Delta\kappa_y, \quad \hat{\sigma}^2 = \frac{1}{n_y - 2} \sum_{y=2}^{n_y} (\Delta\kappa_y - \hat{\mu})^2.$$

Pro účely predikce budeme potřebovat znát i standardní chybu odhadu $\hat{\mu}$:

$$SE(\hat{\mu}) = (\text{Var}[\hat{\mu}])^{\frac{1}{2}} = (\text{Var}[\frac{1}{n_y - 1} \sum_{y=2}^{n_y} \Delta\kappa_y])^{\frac{1}{2}} = (\frac{n_y - 1}{(n_y - 1)^2} \sigma^2)^{\frac{1}{2}} = \frac{\sigma}{(n_y - 1)^{\frac{1}{2}}}. \quad (3.8)$$

Naše situace je následující: odhadli jsme pomocí (GNM) metody řadu parametrů $\kappa_{y_1}, \dots, \kappa_{y_{n_y}}$ a poté je vhodně přeparametrizovali tak, aby byly splněny podmínky jednoznačnosti. Tím pádem můžeme podmínit touto známou řadou a predikovat m budoucích hodnot κ_y pro $y = y_{n_y+1}, \dots, y_{n_y+m}$ dle (3.7) jako:

$$\begin{aligned} \hat{\kappa}_{y_{n_y}+m} &= \hat{\mu} + \kappa_{y_{n_y}+m-1} + \varepsilon_{y_{n_y}+m} = 2\hat{\mu} + \kappa_{y_{n_y}+m-2} + \varepsilon_{y_{n_y}+m} + \varepsilon_{y_{n_y}+m-1} \\ &= m\hat{\mu} + \kappa_{y_{n_y}} + \sum_{j=1}^m \varepsilon_{y_{n_y}+j}. \end{aligned}$$

Naším cílem je odhadnout střední hodnotu $\hat{\kappa}_{y_{n_y}+m}$ podmíněnou známou řadou $(\kappa_{y_1}, \dots, \kappa_{y_{n_y}})$:

$$\begin{aligned} \mathbb{E}[\hat{\kappa}_{y_{n_y}+m} | (\kappa_{y_1}, \dots, \kappa_{y_{n_y}})] &= m\hat{\mu} + \mathbb{E}[\kappa_{y_{n_y}} | (\kappa_{y_1}, \dots, \kappa_{y_{n_y}})] + \sum_{y=1}^m \mathbb{E}[\varepsilon_{y_{n_y}+j}] \\ &= m\hat{\mu} + \kappa_{y_{n_y}} + m \times 0 = m\hat{\mu} + \kappa_{y_{n_y}} \end{aligned} \quad (3.9)$$

a dopouštíme se tedy predikční chyby $m\sigma^2$. Poznamenejme, že lineární řád chyby způsobuje rychle klesající věrohodnost odhadu z (LC) modelu. Upřesněme, že výsledek (3.9) můžeme použít jak ke konstrukci konfidenčních intervalů pomocí standardní chyby (3.8) přenásobenou délkou predikce m , ale taky i ke konstrukci predikčních intervalů pomocí odmocniny z predikčních chyby rovnu $\sqrt{m}\sigma$. Konfidenční interval zachycuje chybu v odhadu parametru, oproti tomu predikční interval obsahuje chybu, jež vychází ze zanedbání náhodných disturbancí (ε_y).

Zabýváme-li se rozšířeným (LC-C) modelem, tak musíme obdobně jako u parametru κ_y předpokládat časovou řadu i pro efekt kohorty γ_c . Pro parametr γ_c předpokládáme, že se řídí procesem ARIMA(1,1,0), tedy:

$$\Delta\gamma_c = \mu + \psi(\Delta\gamma_{c-1}) + \varepsilon_c, \quad (3.10)$$

kde ε_c je náhodný výběr z $N(0, \sigma^2)$ a μ, ψ a σ jsou neznámé parametry. Předpokládejme, že známe odhady parametrů $\hat{\mu}, \hat{\psi}, \hat{\sigma}$, pak m -tý krok predikce je:

$$\begin{aligned} \Delta\gamma_{y_{n_y}+x_{n_x}-1+m} &= \mu + \psi\Delta\gamma_{y_{n_y}+x_{n_x}-2+m} + \varepsilon_{y_{n_y}+x_{n_x}-1+m} = \\ &= m\mu + \psi^m\Delta\gamma_{y_{n_y}+x_{n_x}-1} + \sum_{k=1}^{m-1} \psi^{(m-k)}\Delta\gamma_{y_{n_y}+x_{n_x}-1+k} + \\ &+ \sum_{j=1}^m \varepsilon_{y_{n_y}+x_{n_x}-1+j}. \end{aligned} \quad (3.11)$$

Naší m -tou predikci dostaneme z rovnice (3.11), kde položíme všechny členy řady (ε_c) rovné nule a nahradíme parametry jejich odhadem:

$$\Delta\hat{\gamma}_{y_{n_y}+x_{n_x}-1+m} = m\hat{\mu} + \hat{\psi}^m\Delta\gamma_{y_{n_y}+x_{n_x}-1} + \sum_{k=1}^{m-1} \hat{\psi}^{(m-k)}\Delta\hat{\gamma}_{y_{n_y}+x_{n_x}-1+k}. \quad (3.12)$$

Predikční rozptyl odhadu (3.12) určíme jako:

$$\begin{aligned} \mathbb{E}[\Delta\hat{\gamma}_{c+m} - \Delta\gamma_{c+m}]^2 &= \mathbb{E}\left[\sum_{j=1}^m \varepsilon_{y_{n_y}+x_{n_x}-1+j}\right]^2 = \\ &= \sum_{j=1}^m \mathbb{E}[\varepsilon_{y_{n_y}+x_{n_x}-1+j}]^2 = m\sigma^2, \end{aligned} \quad (3.13)$$

a tedy predikční chybu jako $\sqrt{m}\sigma$.

3.3.2 Predikce v Cairns–Blake–Dowd modelu

V této části se bude věnovat predikci (CBD) modelu a jeho dvou rozšíření, které jsme si zadefinovali výše. V případě základního modelu uvažujeme dva parametry závislé na času: $\kappa_y^{(1)}$ a $\kappa_y^{(2)}$. Oproti původnímu (LC) modelu zde máme dvě proměnné, které na sobě již mohou být závislé a typicky také budou, proto je nutné je modelovat jako dvourozměrnou časovou řadu. Podle Blake a kol. (2006) je vhodné tento náhodný vektor modelovat přes dvourozměrnou náhodnou procházku s driftem. Povšimne si paralely s jednorozměrným případem u (LC) modelu.

Definujeme:

$$\boldsymbol{\kappa}_y = \boldsymbol{\mu} + \boldsymbol{\kappa}_{y-1} + \boldsymbol{\varepsilon}_y \quad y = 2, \dots, n_y, \quad (3.14)$$

kde $\boldsymbol{\varepsilon}_y$ je dvourozměrný náhodný výběr z $N_2(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$ je neznámý vektor a $\boldsymbol{\Sigma}$ je kovarianční matice. Opět zavedeme první diferenci vektorové řady ($\boldsymbol{\kappa}_y$) jako:

$$\begin{aligned} \Delta \boldsymbol{\kappa}_y &= \boldsymbol{\kappa}_y - \boldsymbol{\kappa}_{y-1} \quad y = 2, \dots, n_y \\ &= \boldsymbol{\mu} + \boldsymbol{\varepsilon}_y. \end{aligned}$$

Obdobně jako u predikce parametru (LC) lze ukázat, že $\Delta \boldsymbol{\kappa}_y$ je náhodný výběr z $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a můžeme označit prvky jeho matice jako:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{Var}[\Delta \kappa^{(1)}] & \text{Cov}[\Delta \kappa^{(1)}, \Delta \kappa^{(2)}] \\ \text{Cov}[\Delta \kappa^{(1)}, \Delta \kappa^{(2)}] & \text{Var}[\Delta \kappa^{(2)}] \end{bmatrix}.$$

Odhady parametrů získáme stejnou cestou jako v předchozí podkapitole:

$$\hat{\boldsymbol{\mu}} = \frac{1}{n_y - 1} \sum_{j=2}^{n_y} \Delta \boldsymbol{\kappa}_j, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n_y - 2} \sum_{j=2}^{n_y} (\Delta \boldsymbol{\kappa}_j - \hat{\boldsymbol{\mu}})(\Delta \boldsymbol{\kappa}_j - \hat{\boldsymbol{\mu}})^T, \quad (3.15)$$

Uvažujme situaci, kdy máme odhadnuté parametry (CBD) modelu pomocí metody (GLM). Známe tedy časovou řadu:

$$\begin{bmatrix} \kappa_1^{(1)} \\ \kappa_1^{(2)} \end{bmatrix} \begin{bmatrix} \dots \\ \dots \end{bmatrix} \begin{bmatrix} \kappa_{n_y}^{(1)} \\ \kappa_{n_y}^{(2)} \end{bmatrix}.$$

Pak z (3.14) dostáváme pro m -tý krok predikce:

$$\boldsymbol{\kappa}_{n_y+m} = \boldsymbol{\mu} + \boldsymbol{\kappa}_{n_y+m-1} + \boldsymbol{\varepsilon}_{n_y+m} = m\boldsymbol{\mu} + \boldsymbol{\kappa}_{n_y} + \sum_{j=n_y+1}^{m+n_y} \boldsymbol{\varepsilon}_j, \quad (3.16)$$

kde $\boldsymbol{\kappa}_{n_y}$ je poslední známá hodnota. Naší centrální m -tou predikci dostaneme z rovnice (3.16), kde položíme všechny členy řady ($\boldsymbol{\varepsilon}_y$) rovné nulovému vektoru a nahradíme $\boldsymbol{\mu}$ jeho odhadem:

$$\hat{\boldsymbol{\kappa}}_{n_y+m} = m\hat{\boldsymbol{\mu}} + \boldsymbol{\kappa}_{n_y}.$$

Nechť $\mathbf{X}^{(1)}$ a $\mathbf{X}^{(2)}$ jsou matice modelu (CBD) pro predikci m let do budoucnosti:

$$\begin{aligned} \mathbf{X}^{(1)} &= \mathbf{I}_m \otimes \mathbf{1}_{n_x} \\ \mathbf{X}^{(2)} &= \mathbf{I}_m \otimes \mathbf{x}_m \\ \mathbf{X} &= (\mathbf{X}^{(1)}, \mathbf{X}^{(2)}), \end{aligned}$$

kde $\mathbf{x}_m = (x_1 - \bar{x}, \dots, x_{n_x} - \bar{x})^T$. Abychom mohli zkonstruovat konfidenční interval musíme spočítat rozptyl m -té predikce. Označme celou řadu predikcí počínaje $n_y + 1$ až $n_y + m$ maticí typu $m \times 2$:

$$\hat{\boldsymbol{\kappa}}_F = (\kappa_{n_y}^{(1)} \mathbf{1}_m + \hat{\mu}_1 \mathbf{c}_m, \kappa_{n_y}^{(2)} \mathbf{1}_m + \hat{\mu}_2 \mathbf{c}_m)^T,$$

kde $\mathbf{c}_m = (1, \dots, m)^T$. Spočtěme tedy kovarianční matici predikce podmíněnou poslední známou hodnotou ($\boldsymbol{\kappa}_{n_y}$):

$$\begin{aligned} \text{Var}[\mathbf{X}\hat{\boldsymbol{\kappa}}_F | (\boldsymbol{\kappa}_{n_y})] &= \mathbf{X} \text{Var} \begin{bmatrix} \hat{\mu}_1 \mathbf{c}_m \\ \hat{\mu}_2 \mathbf{c}_m \end{bmatrix} \mathbf{X}^T \\ &= \mathbf{X} \text{Var} \begin{bmatrix} \mathbf{c}_m \text{Var}[\hat{\mu}_1] \mathbf{c}_m^T & \mathbf{c}_m \text{Cov}[\hat{\mu}_1, \hat{\mu}_2] \mathbf{c}_m^T \\ \mathbf{c}_m \text{Cov}[\hat{\mu}_1, \hat{\mu}_2] \mathbf{c}_m^T & \mathbf{c}_m \text{Var}[\hat{\mu}_2] \mathbf{c}_m^T \end{bmatrix} \mathbf{X}^T \\ &= \mathbf{X} \text{Var}[\hat{\boldsymbol{\mu}}] \otimes \mathbf{c}_m \mathbf{c}_m^T \mathbf{X}^T = \frac{1}{n_y - 1} \mathbf{X} \boldsymbol{\Sigma} \otimes \mathbf{c}_m \mathbf{c}_m^T \mathbf{X}^T, \quad (3.17) \end{aligned}$$

kde jsme v první a druhé rovnosti využili vlastnosti kovarianční matice vzhledem ke konstantě a násobení nenáhodnou maticí nebo vektorem, ve třetí rovnici definici Kroneckerova součinu a v poslední rovnosti kovarianční matici $\hat{\boldsymbol{\mu}}$ vycházející z (3.15). Nyní již můžeme jednoduše postavit konfidenční interval, a to tak, že odmocníme jednotlivé prvky diagonály (3.17) a s hodnotou centrální predikce určíme ramena intervalu.

Predikční chybu odhadu (3.16) určíme jako:

$$\mathbb{E}[\hat{\boldsymbol{\kappa}}_{n_y+m} - \boldsymbol{\kappa}_{n_y+m}]^2 = \mathbb{E} \left[\sum_{j=n_y+1}^{m+n_y} \boldsymbol{\varepsilon}_j \right]^2 = \sum_{j=n_y+1}^{m+n_y} \mathbb{E}[\boldsymbol{\varepsilon}_j]^2 = m \boldsymbol{\Sigma}. \quad (3.18)$$

Pro konstrukci predikčních intervalů modelu (CBD) musíme predikční chybu (5.7) přenásobit z obou stran vektorem $(1, x - \bar{x})^T$ a vytvořit řadu pomocí vektoru \mathbf{c}_m :

$$(1, x - \bar{x})^T \boldsymbol{\Sigma} (1, x - \bar{x}) \mathbf{c}_m.$$

Kohortní člen γ_c v modelu (CBD-C) modelujeme stejně jako u (LC-C), a to tedy dle (3.10). Predikce v modelu (CBD-V) je analogickým rozšířením predikce (CBD) o třetí dimenzi.

3.4 Diagnostika modelů

V této sekci si zadefinujeme tři predikční testy, které budeme následně aplikovat v podkapitole 6.2 *Empirická diagnostika modelů* na reálných datech. Dané testy sledují vlastnosti predikcí, každý z nich má za cíl ověřit jinou specifickou vlastnost. Sekce vychází z Dowd a kol. (2010). Poznamenejme, že existuje mnoho způsobů testování správnosti predikce stochastický úmrtnostních modelů. Nicméně my považujeme následující tři testy jako dostatečné k posouzení věrohodnosti predikce modelu.

Hlavním účelem výzkumu stochastických úmrtnostních modelů je snaha o co nejlepší možný odhad budoucích měř úmrtnosti pro jednotlivé věky z dostupných dat. Proto se naše testy budou zaměřovat na schopnost modelů věrohodně odhadnout úmrtnost. Pokud se chystáme analyzovat vlastnosti odhadů modelů je třeba rozlišovat dva druhy odhadů: *odhad ve vzorku* a *odhad mimo vzorek*. Diagnostikou *odhadu ve vzorku* rozumíme schopnost modelu vhodně prokládat data,

z nichž byly odhadnuty jeho koeficienty, naproti tomu diagnostikou *odhadu mimo vzorek* chápeme jako schopnost modelu vhodně predikovat budoucí body, ze kterých nejsou modelové koeficienty spočteny.

Poznamenejme, že dobrý model, vztaženo k naším požadavku, by měl poskytovat oba odhady v dobré kvalitě. Zaměření těchto testů je zkoumat vlastnosti *odhadu mimo vzorek*, který je v kontextu pojišťovny primárním zájmem. Atributy *odhadu ve vzorku* budeme analyzovat v podkapitole 6.1 *Odhady modelů*.

3.4.1 Test konvergence zužujícího horizontu

V prvním testu budeme zkoumat, zda predikce modelu konzistentně konverguje ke skutečné hodnotě při postupném zužování horizontu předpovědi. Zvolíme si jeden konkrétní rok Y_1 a odhadneme n_1 -krát model k rokům $Y_1 - 1, Y_1 - 2, \dots, Y_1 - n_1$. Pro každý odhadnutý model budeme predikovat hodnotu roku Y_1 . Jinými slovy dostaneme n_1 predikcí hodnoty roku Y_1 , kde nejkratší předpověď je 1-roční a nejdelší n_1 -roční. Model považujeme za úspěšný v tomto testu pokud se zužujícím horizontem predikce konzistentně konverguje ke skutečné hodnotě. Dodejme, že konzistence konvergence je klíčová vlastnost, proto se na ní primárně zaměříme v empirické části testu. Tento test je empiricky vyhotoven v podkapitole 6.2.1 *Test konvergence zužujícího horizontu*.

3.4.2 Test dlouhé predikce

V tomto testu budeme analyzovat schopnost modelu predikovat delší časové predikce od fixního roku Y_2 do roku $Y_2 + n_2$. Předchozí test zkoumal predikci jednoho konkrétního roku Y_1 při zkracujícím se okně predikce. Naopak v tomto testu budeme porovnávat jednotlivé budoucí realizace v letech $Y_2 + 1, Y_2 + 2, \dots, Y_2 + n_2$ s predikcí založenou na datech do roku Y_2 . Hlavní rozdíl oproti předchozím testu je zrušení fixace konkrétního roku Y_1 . K vyhodnocení testu si zavedeme pomocné statistiky. Budeme počítat, kolik skutečných hodnot leží pod centrální predikcí, kolik nad, a také kolik jich bude ležet pod dolním intervalem predikce a kolik nad horním intervalem predikce. Definujme tedy statistiky:

- $x_L :=$ počet skutečných hodnot ležících pod dolním intervalem predikce,
- $x_U :=$ počet skutečných hodnot ležících nad horním intervalem predikce,
- $x_{C_L} :=$ počet skutečných hodnot ležících pod centrální predikcí,
- $x_{C_U} :=$ počet skutečných hodnot ležících nad centrální predikcí.

U dobrého modelu očekáváme přibližnou shodu x_{C_L} a x_{C_U} . Pokud převažuje jedna statistika nad druhou, tak daný model buď nadhodnocuje ($x_{C_L} \gg x_{C_U}$), nebo podhodnocuje ($x_{C_L} \ll x_{C_U}$). Podstatný je i součet statistik $x_L + x_U$, je-li model spolehlivý, tak by tento součet měl být co nejmenší, přesněji $x_L + x_U < n_2\alpha$, uvažujeme-li $(1 - \alpha)\%$ -predikční interval. Tento test budeme empiricky vyhodnocovat v podkapitole 6.2.2 *Test dlouhé predikce*.

3.4.3 Test posouvajícího se okna predikce

V posledním testu chceme zkoumat vývoj a stabilitu predikce, když posouváme predikované okno dopředu. Konkrétněji, zvolíme fixní rok Y_3 , od kterého budeme predikovat, začátek predikčního okna n_{3s} a konec predikčního okna n_{3e} . Pro každý model provedeme $n_{3e} - n_{3s}$ predikcí. Jednotlivé predikce budou mít predikční okno n_{3s} dlouhé, vycházející z dat od $Y_3 - h + 1, Y_3 - h + 2, \dots, Y_3 - h + n_{3e} - n_{3s}$ do $Y_3, Y_3 + 1, \dots, Y_3 + n_{3e} - n_{3s} - 1$. Opět nás bude zajímat kolik skutečně realizovaných hodnot bude ležet v predikčních pásech, využijeme i čtveřici hodnotících statistik $[x_L, x_U, x_{C_L}, x_{C_U}]$, jež byly definovány v předchozím testu. Od dobrého modelu očekáváme obdobné vlastnosti jako u *Testu dlouhé predikce*. Daný test empiricky zkoumáme v kapitole 6.2.3 *Test posouvajícího se okna predikce*.

4. Pojistněmatematická rizika v kontextu úmrtností

V této kapitole se budeme zabývat riziky, kterým je pojišťovna vystavena a zároveň se pojí s úmrtností. Nejprve je třeba si ujasnit, co myslíme pod pojmem riziko. V kontextu této práce budeme *rizikem* rozumět jakoukoliv událost, která může mít negativní dopad na budoucí peněžní toky pojišťovny. Zákon o pojišťovnictví SII (2009) v sekci solventnosti ukládá pojišťovně povinnost dostát svých závazků v libovolném budoucím okamžiku. Proto je pro pojišťovnu klíčové správně definovat, měřit a zajišťovat se proti rizikům, které by mohly mít signifikantní dopady na její finanční stabilitu. *Zajištěním rizika* rozumíme libovolné přenesení rizika na další subjekt, který typicky za přijetí rizika požaduje kompenzaci ve formě rizikové prémie. Připomeňme, že čtenář může nalézt definice pojistně-technických pojmů v knize Cipra (2006).

Pojišťovna je vystavena velkému množství různorodých rizik, avšak jak již bylo zmíněno výše, my se budeme zabývat pouze vybranými riziky, jež přímo souvisejí s úmrtnostní tematikou. Začneme od rizik specifikovaných v životním sub-modulu standardního vzorce sloužícího pro výpočet solventnostního rizikového kapitálu dle regulatorního rámce *Solvency II: úmrtnosti, dlouhověkosti, katastrofické*. Tato část čerpá převážně z EU (2009). Další skupinou jsou rizika související přímo s odhadem a modelem: *úrovně a trendu*, které vycházejí z Macdonald a kol. (2018). Jednotlivá rizika si probereme v následujících podkapitolách.

4.1 Riziko úmrtnosti a katastrofy

Rizikem úmrtnosti rozumíme ztrátu nebo nepříznivou změnu hodnoty pojistného závazku vyplývající ze změn úrovně, vývoje nebo volatility měř úmrtnosti, kdy zvýšení míry úmrtnosti vede ke zvýšení hodnoty pojistných závazků. Klíčovým v této definici je efekt zvýšení úmrtnosti. Podíváme-li se na typické portfolio smluv životních pojišťoven, tak největší hrozbou rizika úmrtnosti jsou upsaná riziková pojištění pro případ smrti. Navýšení míry úmrtnosti v tomto případě vede k porušení principu ekvivalence, a to tak, že dochází k dřívějším úmrtím pojištěných osob, než bylo při uzavření smluv očekáváno, a tím pádem jsou dříve vyplácena pojistná plnění. Navíc u regulérně placených pojištění jsou pozitivní peněžní toky plynoucí z placení pojistného zastaveny dříve, než se očekávalo a ztráta z těchto smluv je ještě více signifikantní. Riziko úmrtnosti v souvislosti se stochastickými úmrtnostními modely je podhodnocení budoucí míry úmrtnosti oproti skutečným realizacím.

V kontextu *Solvency II* je třeba rozlišovat mezi rizikem úmrtnosti a životní katastrofy, kdy obě tyto rizika souvisí s dřívějším úmrtím oproti předchozímu očekávání, avšak katastrofické se pojí se šokovým skokem úmrtnosti v krátkém časovém horizontu v řádu jednotek let (epidemie, zemětřesení, tsunami,...), naopak pod ri-

zíkem úmrtnosti si můžeme představit například předčasné úmrtí spojené s kouřením, obezitou či nízkou fyzickou aktivitou. Katastrofické riziko je definováno ve standardním vzorci jako riziko ztráty nebo nepříznivé změny hodnoty pojistného závazku vyplývajícího ze značné neurčitosti předpokladů při tvorbě cen a stanovení rezerv v souvislosti s mimořádnými nebo zvláštními událostmi. Formálně dle *Solvency II* je riziko úmrtnosti definované jako okamžitý a trvalý nárůst úmrtnosti o 15 %, tedy $q_x \times 1.15$. Riziko katastrofy je definováno jako okamžité absolutní zvýšení úmrtnosti o 0.0015 po dobu jednoho roku, tedy $q_x + 0.0015$.

4.2 Riziko dlouhověkosti

Rizikem dlouhověkosti rozumíme ztrátu nebo nepříznivé změny hodnoty pojistného závazku vyplývající ze změn úrovně, vývoje nebo volatility měr úmrtnosti, kdy snížení míry úmrtnosti vede ke zvýšení hodnoty pojistných závazků. Povšimneme si podobnosti s rizikem úmrtnosti, kde jediný rozdíl v definici je snížení míry úmrtnosti vedoucí k zvýšení závazků. Budeme-li se opět zabývat typickým portfoliem smluv životní pojišťovny nebo penzijní společnosti, tak riziko dlouhověkosti má největší dopad na anuitní smlouvy a pojištění pro případ dožití. Toto riziko je esenciální pro správce penzijních fondů a současně způsobuje obrovské finanční ztráty, jak penzijním společností, tak státům starajícím se o starobní důchodce. Extrémní finanční dopady pojící se s tímto rizikem v penzijních systémech jsou spojeny se stále zlepšujícím se zdravotnictvím a životním stylem pojištěných a také obrovským pojistným kmenem, který se díky těmto faktorům stále zvětšuje. Podstatným vlivem v neudržitelnosti penzijních systému je také vývoj porodnosti kohort a demografická struktura obyvatelstva, těmito faktory se nebudeme dále zabývat, neboť se přímo nepojí s úmrtnostními modely. Riziko dlouhověkosti je jedním z nejvíce materiálních rizik pro životní pojišťovny a penzijní společnosti, proto se budeme v další kapitole zabývat přístupem k jeho zajištění. Formálně dle *Solvency II* je riziko dlouhověkosti definované jako okamžitý a trvalý pokles úmrtnosti o 20 %, tedy $q_x \times 0.80$.

4.3 Riziko úrovně

Rizikem úrovně nebo také rizikem špatného odhadu rozumíme riziko, jež vzniká, je-li současný pojistitelův odhad úmrtnosti nesprávný nebo odchýlený od skutečnosti. Pod toto riziko spadá i nepřesné určení demografických předpokladů, např. předpoklady různých kategorií morbidit. Pojistitel má k dispozici pouze konečný počet relevantních pozorování pro daná portfolia a na základě nich stanovuje svoje odhady a předpoklady. Příkladem může být různé chování kmenů jednotlivých druhů pojištění dané pojišťovny. Používá-li pojistitel svoje vlastní úmrtnostní tabulky založené na úmrtí v jeho pojistných kmenech, pak může vzniknout riziko úrovně. Takovéto tabulky jsou typicky založeny na relativně malém množství pozorování a tím pádem mohou být vychýleny od skutečných úmrtností. Dalším příkladem tohoto rizika může být nerefektování případné změny pravidel úpisu do svých úmrtnostních odhadů.

4.4 Riziko trendu

Rizikem trendu nebo také vývoje rozumíme riziko odchýlení budoucího vývoje úmrtnosti od nejlepšího předpokladu založeného na historických datech. Neboť jsou všechny odhady úmrtnosti vycházející ze stochastických úmrtnostních modelů založeny na principu modelování budoucnosti na základě minulosti, nemusí se v takovémto odhadu projevit trend, který není třeba nutně zachycen v historických datech nebo může být pouze mírně reflektován v nejnovější datech a tím pádem tento trend nebude zohledněn v nejlepším odhadu budoucího vývoje úmrtnosti. Riziko trendu můžeme rozdělit na riziko modelu a selekční riziko, které si představíme v další samostatné podkapitole. Rizikem modelu rozumíme rozdílnost nejlepších odhadů úmrtnosti založené na různých modelech. S jistotou nelze říct, který z uvažovaných modelů má ten správný budoucí odhad. Můžeme zavést určité diagnostiky a testy na posouzení správnosti daných modelů k redukcí tohoto rizika, ale nikdy ho nemůžeme plně mitigovat, proto pojistitel musí vzít v potaz existenci tohoto rizika a počítat s ním ve své kapitálové pozici. Diagnostické testy jsme si zavedli v podkapitole 3.4 *Diagnostika modelů* a budeme je aplikovat na data v podkapitole 6.2 *Empirická diagnostika modelů*.

4.5 Selekční riziko

Selekčním rizikem rozumíme ztrátu nebo nepříznivé změny hodnoty pojistného závazku vyplývající z faktu, že úmrtnosti pojistných kmenů se neřídí úmrtnostmi populace. Je-li tedy pojistný produkt, jehož kmen nebude následovat uvažované úmrtnosti populace, oceněn pomocí úmrtnostních tabulek populace, tak dochází k realizaci selekčního rizika. Pojistné kmeny mají typicky vyšší pravděpodobnosti dožití, než predikují úmrtnostní tabulky populace. Tato skutečnost je způsobena několika faktory. Jedním z nejvýznamnějších je selekce pojištěných osob na základě jejich zdravotní dokumentace, profese či zálib. Další příčina je, že životní pojištění si typicky pořizují lidé s vyššími příjmy a vyšším vzděláním, než je průměr populace, jak tvrdí práce Browne a Kim (1993). Tyto vlastnosti jsou spojené také s lepším životním stylem. Poslední možný faktor, je autorova domněnka, že dojde-li k zhoršení zdravotního stavu pojištěného a s tím například spojená ztráta zaměstnání, pak je-li daný pojištěný i pojistníkem, tak jeho nová finanční situace může napomoci k předčasnému ukončení placení pojistného a tedy i ukončení pojistné smlouvy bez nároku na případné plnění v případě smrti. Poznamenejme, že selekčnímu riziku se také v některých pramenech říká riziko základu.

5. Zajištění rizika dlouhověkosti

V této části se zaměříme na zajištění rizika, konkrétně na zajištění vůči riziku dlouhověkosti, které bude naším centrálním zájmem i v podkapitole 6.3 *Ocenění longevity dluhopisu*. Jak jsme již uvedli v předchozí kapitole, riziko dlouhověkosti považujeme jako jedno z nejméně materiálních rizik pro penzijní fondy a pojišťovny s významným portfoliem anuitantů. Nejprve si představíme longevity dluhopis, který představuje finanční nástroj pro pojistitele na zajištění proti riziku dlouhověkosti anuitních kmenů. Následně se zaměříme na teoretický přístup k ocenění takového kontraktu pomocí parametru tržního rizika skrze přístup zvaný jako Wangova transformace.

5.1 Longevity dluhopis

Jak již bylo uvedeno výše, jedním z nejzávažnějších pojistných rizik, kterým čelí životní pojišťovny a penzijní fondy, zejména v případě výplaty doživotních důchodů, je nepredikovaný nárůst dlouhověkosti. V této části si ukážeme, jak je možné se zajistit proti riziku dlouhověkosti v portfoliu anuit pomocí speciálního typu dluhopisu. Tato podkapitola je zpracovaná dle Lin a Cox (2005).

Předpokládejme, že pojistitel má portfolio, které obsahuje l_x životních anuit ve věku x v čase $t = 0$. Pro jednoduchost necht všechny smlouvy mají roční výplatu v jednotkové hodnotě. Pojistitel se chce zajistit proti riziku dlouhověkosti, neboť jeho očekávané peněžní toky v roce t jsou stanovené pomocí očekávaných počtu přeživších l_{x+t} v roce t . Naopak jeho skutečné výdaje jsou odvozené ze skutečně realizovaného počtu přeživších \hat{l}_{x+t} . Pokud platí, že $\hat{l}_{x+t} \gg l_{x+t}$ pak je porušen princip ekvivalence a pojistitel trpí neočekávanou ztrátou. Jinými slovy riziko odchylky jeho nejlepšího odhadu není reflektováno v ceně pojištění.

V rámci longevity dluhopisu uvažujeme tři subjekty, které mezi sebou uzavírají finanční smlouvu: pojistitel, zprostředkovatel a investor. Pojistitel se chce zajistit vůči riziku dlouhověkosti, investor chce zhodnotit svojí investici a zprostředkovatel slouží jako prostředník mezi oběma subjekty. Zprostředkovatele uvažujeme, abychom eliminovali kreditní riziko protistrany. V čase $t = 0$ pojistitel kupuje zajištění od zprostředkovatele za cenu P , zároveň investor nakupuje od zprostředkovatele dluhopis za cenu V . Prostředník pomocí kombinací peněžních toků od obou subjektů nakoupí bezrizikový (státní) dluhopis za cenu W . Zde je vidět úloha zprostředkovatele, který nákupem bezrizikového aktiva zredukuje kreditní riziko. Zprostředkovatel bude vyžadovat, aby platil vztah $P + V \geq W$. Pro zjednodušení předpokládejme, že platí $P + V = W$, tedy žádná provize pro správce smlouvy. Bezrizikový dluhopis nese roční kuponové platby ve výši C_t . Tyto kupony rozdělí zprostředkovatel v čase t následovně:

$$C_t = B_t + D_t, \tag{5.1}$$

kde B_t je kupon pro pojistitele a D_t je kupon pro investora. Pojistitelovy peněžní toky v roce $t > 0$ plynoucí ze zajištění jsou:

$$\begin{aligned} B_t &= C_t, & \hat{l}_{x+t} - l_{x+t} &> C_t \\ &= \hat{l}_{x+t} - l_{x+t}, & l_{x+t} &\leq \hat{l}_{x+t} \leq C_t \\ &= 0 & \hat{l}_{x+t} &< l_{x+t}. \end{aligned} \quad (5.2)$$

Celkovou náhodnou ztrátu v roce $t > 0$ pro pojistitele bez zajištění můžeme vyjádřit jako:

$$Z(t) = \max(\hat{l}_{x+t} - l_{x+t}, 0). \quad (5.3)$$

Zajistí-li se pojistitel longevity bondem, pak jeho ztrátu můžeme vyjádřit jako:

$$Z_H(t) = Z(t) - B_t,$$

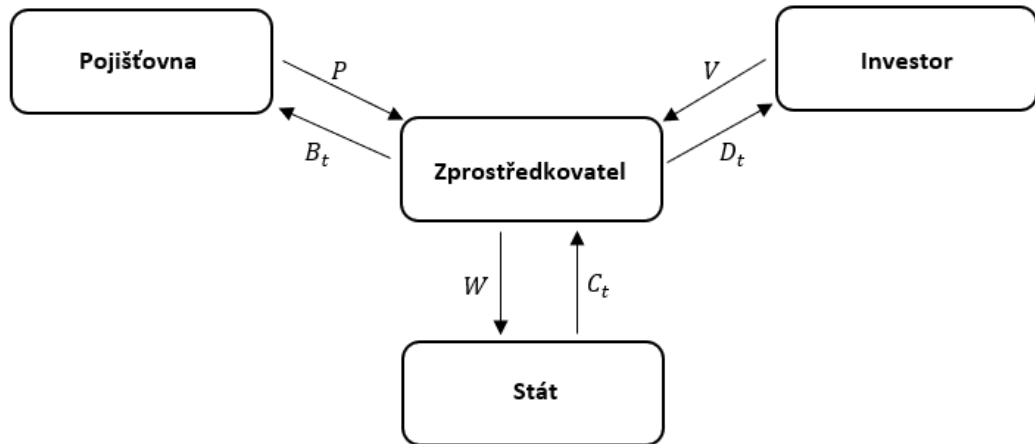
která je nenulová pouze v tom případě, že $\hat{l}_{x+t} - l_{x+t} > C_t$. Tedy překročí-li počet přeživších určitou vyšší mez stanovenou v podobě kuponu C_t . Kupon D_t pro investora určíme jednoduše pomocí rovnic (5.1) a (5.2):

$$\begin{aligned} D_t &= 0, & \hat{l}_{x+t} - l_{x+t} &> C_t \\ &= C_t + l_{x+t} - \hat{l}_{x+t}, & l_{x+t} &\leq \hat{l}_{x+t} \leq C_t \\ &= C_t & \hat{l}_{x+t} &< l_{x+t}. \end{aligned} \quad (5.4)$$

Rovnice zprostředkovatele:

$$P + V = W = Fd(0,T) + \sum_{t=1}^T C_t d(0,t),$$

kde F je nominální hodnota státního dluhopisu a $d(0,t)$ je diskontní faktor do času t . První rovnost plyne z předpokladu nulové provize zprostředkovatele. Druhá rovnost plyne z principu férového ocenění dluhopisu, který v časech t vyplácí kupóny C_t a v době splatnosti vyplácí nominál F . V obrázku 5.1 dodáváme náčrt schéma peněžních toků longevity dluhopisu.



Obrázek 5.1: Schéma peněžních toků longevity dluhopisu.

Celkově je tedy zprostředkovatelův peněžní tok roven nule, pojistitel je plně zajištěn vůči riziku dlouhověkosti, pokud nedojde k extrémnímu nárůstu realizovaných přeživších oproti předpokládané hodnotě a i v tom případě je jeho realizovaná ztráta ponížena o hodnotu kupónu C_t . Investor profituje z kontraktu, jsou-li realizovaní přeživší blízko očekávané hodnoty nebo dokonce pod očekávanou hodnotou. Následuje zásadní otázka: Jaká má být cena longevity dluhopisu při předpokladu bezarbitrážního principu?

Poznamenejme, že zajištění proti riziku dlouhověkosti se dá dosáhnout i bez zprostředkovatele a to formou swapu. Avšak tato forma bez přísunu bezrizikových kuponů v sobě nese riziko, že investor nebude moci splácet kuponové platby náležící pojišťovně, vzrostou-li významně realizované počty přeživších nad očekávané hodnoty. V tomto případě je nutné uvažovat v ocenění longevity dluhopisu i kreditní riziko protistrany.

5.2 Wangova transformace

Jedním z možných přístupů k ocenění longevity kontraktu je použití metody distorze, konkrétně v podobě Wangovy transformace, která je popsána například v Lin a Cox (2005). Princip tohoto přístupu je transformovat rozdělení distribuční funkce tak, aby zahrnovala i tržní hodnotu ceny rizika a to skrze neznámý parametr. Předpokladem je, že trh uvažuje nejistotu spojenou s cenou anuity, která vzniká odhadem z úmrtnostních tabulek a reflektuje jí ve svém ohodnocení daného kontraktu. V našem případě budeme Wangovu transformaci aplikovat na rozdělení pravděpodobnosti úmrtí. Postup bude následovný: tržní cenu anuity položíme rovnu Wangově transformaci na pravděpodobnost úmrtí, kterou odhadneme pomocí modelu uvedených v kapitole stochastických úmrtnostních modelů a spočteme hodnotu parametru tržního rizika. Nakonec určíme spravedlivé pojistné P pro zajistnou smlouvu i cenu V pro investora.

Necht $\Phi(x)$ reprezentuje distribuční funkci standardního normálního rozdělení. Poté můžeme definovat distorzní operátor distribuční funkce $\Phi(x)$ jako:

$$g_\lambda(u) = \Phi(\Phi^{-1}(u) - \lambda), \quad (5.5)$$

kde $0 < u < 1$ a parametr $\lambda \in \mathbb{R}$ představuje tržní cenu rizika. Pomocí operátoru (5.5) je možné pro libovolnou distribuční funkci $F(x)$ definovat její Wangovu transformaci jako:

$$F^*(x) = g_\lambda(F(x)), \quad x \in \mathbb{R}. \quad (5.6)$$

$F^*(x)$ reprezentuje skrze parametr λ rizikové rozšíření distribuční funkce $F(x)$. Férová cena jakéhokoliv závazku by pak měla být získaná pomocí diskontované střední hodnoty vycházející z rozdělení představující $F^*(x)$, které je již obohacené o systematické riziko trhu, protože všechny strany toto riziko zohledňují a tím pádem i reflektují v ceně daného závazku. Tímto typem rizika rozumíme riziko, které se nedá odstranit diverzifikací a je tedy přijato trhem jako nejnižší možné

riziko a proto je reflektované v každé tržní ceně. Při uvažování longevity bondu je příkladem takového rizika nejistota spojená s konstrukcí úmrtnostní tabulky, která je použita pro ocenění anuit, ale také nejistý vývoj budoucích smrtí. Tyto nejistoty musí být součástí ocenění v bezarbitrážním přístupu, neboť je nelze eliminovat. Parametr λ odhadneme z dat pomocí následující rovnice:

$$a_x^{market}(t_0) = \sum_{t=1}^T [1 - \phi(\phi^{-1}({}_t\hat{q}_x) - \lambda_x(t_0))]d(0,t), \quad (5.7)$$

kde a_x^{market} je tržní cena jednotkové anuity pro věk x a člen

$${}_t\hat{q}_x^* := \phi(\phi^{-1}({}_t\hat{q}_x) - \lambda_x(t_0)) \quad (5.8)$$

je Wangova transformace pravděpodobnosti úmrtnosti, reprezentující pravděpodobnost úmrtí obohacenou o rizikovou přírážku. Tato tržní cena zahrnuje i systematické riziko a tím pádem můžeme numericky určit parametr λ_x . Vztah (5.7) je odvozen z netto přístupu, kde jest $\lambda_x = 0$, a tedy rovnice:

$$a_x(t_0) = \sum_{t=1}^T [1 - \phi(\phi^{-1}({}_t\hat{q}_x) - 0)]d(0,t) = \sum_{t=1}^T [1 - {}_t\hat{q}_x]d(0,t),$$

představuje spravedlivou cenu jednotkové anuity, která vychází z principu ekvivalence. Formulí lze nalézt v Cipra (2006). Parametr tržního rizika λ_x v rovnici (5.7) reprezentuje úpravu netto přístupu o tržní riziko, které je obsaženo v tržní ceně anuity.

Známe-li tržní cenu rizika můžeme již přímočaře spočítat spravedlivé pojistné P pro zajistnou smlouvu i cenu V pro investora jako:

$$P = \sum_{t=1}^T E^*[B_t]d(0,t),$$

$$V = \sum_{t=1}^T E^*[D_t]d(0,t) + Fd(0,T),$$

$E^*[Z]$ je střední hodnota vztažena k ${}_t\hat{q}_x^*$, B_t a D_t jsou kupóny definované v (5.2) a (5.4).

6. Empirická část

V této části budeme odhadovat, diagnostikovat a aplikovat úmrtnostní modely definované v teoretické podkapitole 3.2 *Podstata modelů a jejich třídy*. Pro účely diagnostiky využijeme tři testy, které jsme si zavedli v podkapitole 3.4 *Diagnostika modelů*. Předlohou pro aplikaci pro nás bude kapitola 5. *Zajištění rizika*. Kapitola má následující strukturu: nejprve si zadefinujeme datovou sadu, ze které budeme vycházet. Poté v podkapitole 6.1 *Odhady modelů* provedeme odhady všech modelů na základě jejich definic. Budeme analyzovat odhadnuté parametry a residua. Následně v části 6.2 *Empirická diagnostika modelů* prozkoumáme jednotlivé modely všemi třemi teoretickými predikčními testy pro vybrané věky. V posledním úseku budeme aplikovat modely na odhad férové ceny longevity dluhopisu a s ním spojené tržní ceny rizika. Dodejme, že všechny výpočty byly provedeny v softwaru Rstudio.

Nejprve si zadefinujeme datovou sadu, ze které budeme vycházet v celé kapitole 6. Datovým vzorkem pro nás je jak sada počtu centrálních úmrtí, tak sada expozic v České republice. Obě databáze začínají rokem 1950 a končí v roce 2019. Naměřené počty jsou od věku 0 do věku 110, přičemž jsou uvedeni samostatně muži a ženy. Datové sady pocházejí z *The Human Mortality Database* HMD (2021), která obsahuje úmrtnostní a populační(expoziční) data určená pro vědce, analytiku, zpravodaje,... Jedná se o veřejně přístupnou databázi, která obsahuje data ze 41 zemí včetně České republiky. Česká data jsou přebírána přímo z *Českého statistického úřadu*. Databáze obsahuje různé průniky časových a věkových oken obsahující úmrtnosti, expozice, očekávané doby dožití, úmrtnostní tabulky a velikosti populací dělené pro obě pohlaví. V různých částí této kapitoly se budeme omezovat na specifické podmnožiny těchto dat. V tabulce 6.1 dodáváme statistiky vývoje průměrného věku úmrtí a průměrného věku populace po desetiletých oknech, zvláště pro muže a ženy. Pozorujeme rostoucí trend ve vývoji všech sledovaných hodnot. Povšimneme si i rozdílů v průměrném věku úmrtí mužů a žen. U starších roků vidíme rozdíl kolem pěti let ve středu intervalu dokonce i osm let.

Roky	Průměrný věk úmrtí		Průměrný věk expozice	
	Muži	Ženy	Muži	Ženy
1950-1959	59,5	64,3	32,0	34,3
1960-1969	63,3	69,2	33,1	35,9
1970-1979	64,6	71,1	33,4	36,7
1980-1989	66,5	73,6	33,4	36,8
1990-1999	67,1	75,1	34,9	38,2
2000-2009	68,8	76,6	37,6	40,7
2010-2019	71,3	78,5	39,7	42,6

Tabulka 6.1: Statistiky průměrných věků úmrtí a expozic pro ženy a muže z českých dat od roku 1950 do 2019 po desetiletých skupinách.

6.1 Odhady modelů

V této kapitole postupně odhadneme všechny modely z podkapitoly 3.2. *Podstata modelů a jejich třídy* a budeme zkoumat jejich vlastnosti a interpretaci jednotlivých parametrů. Datovou sadu pro účel odhadů budeme brát celou v dimenzi roku, tedy od roku 1950 do roku 2019, v dimenzi věku se omezíme na věky od 50 do 95. Takto zúžený výběr dat je motivován faktem velice nízkých měřených historických hodnot pro starší roky, který je způsoben ztlačení střední dobou dožití pro obě pohlaví hlavně pro roky po půlce 20. století. V souladu se značením používaným v práci máme: $n_x = 46$, $n_y = 70$ a $x_1 = 50, \dots, x_{n_x} = 95$, $y_1 = 1950, \dots, y_{n_y} = 2019$.

Budeme zkoumat odhad intenzity úmrtnosti, kterou aproximujeme diskrétní centrální úmrtnostní mírou, dle vztahu (2.3). Poznamenejme, že u všech modelů uvažujeme logaritmickou transformaci odezvy. Dále předpokládáme, že počet smrti pochází z Poissonova rozdělení, tak jako bylo uvedeno v (2.4). Pro každý model spočteme takzvaná *pracovní* residua, to jsou residua z poslední iterace algoritmu iterativních vážených nejmenších čtverců. Residua jsou tvaru:

$$res_k = \frac{m_k - \hat{m}_k}{\hat{m}_k}, \quad k = 1, \dots, n_x \times n_y.$$

Poznamenejme, že residua jsou řazeny dle vektoru odezvy, tedy od nejstarších roků a nejmladších věků. Složky vektoru residui odpovídají:

$$[(y_1, x_1), (y_1, x_2), \dots, (y_1, x_{n_x}), (y_2, x_1), \dots, (y_{n_y}, x_{n_x})].$$

Všechny odhady a analýzy jsou provedeny na mužské populaci. Ekvivalentní grafy pro ženskou populaci může čtenář nalézt v *Příloze A.1*.

6.1.1 Lee-Carterův model

Připomeňme, že (LC) model je definován následovně:

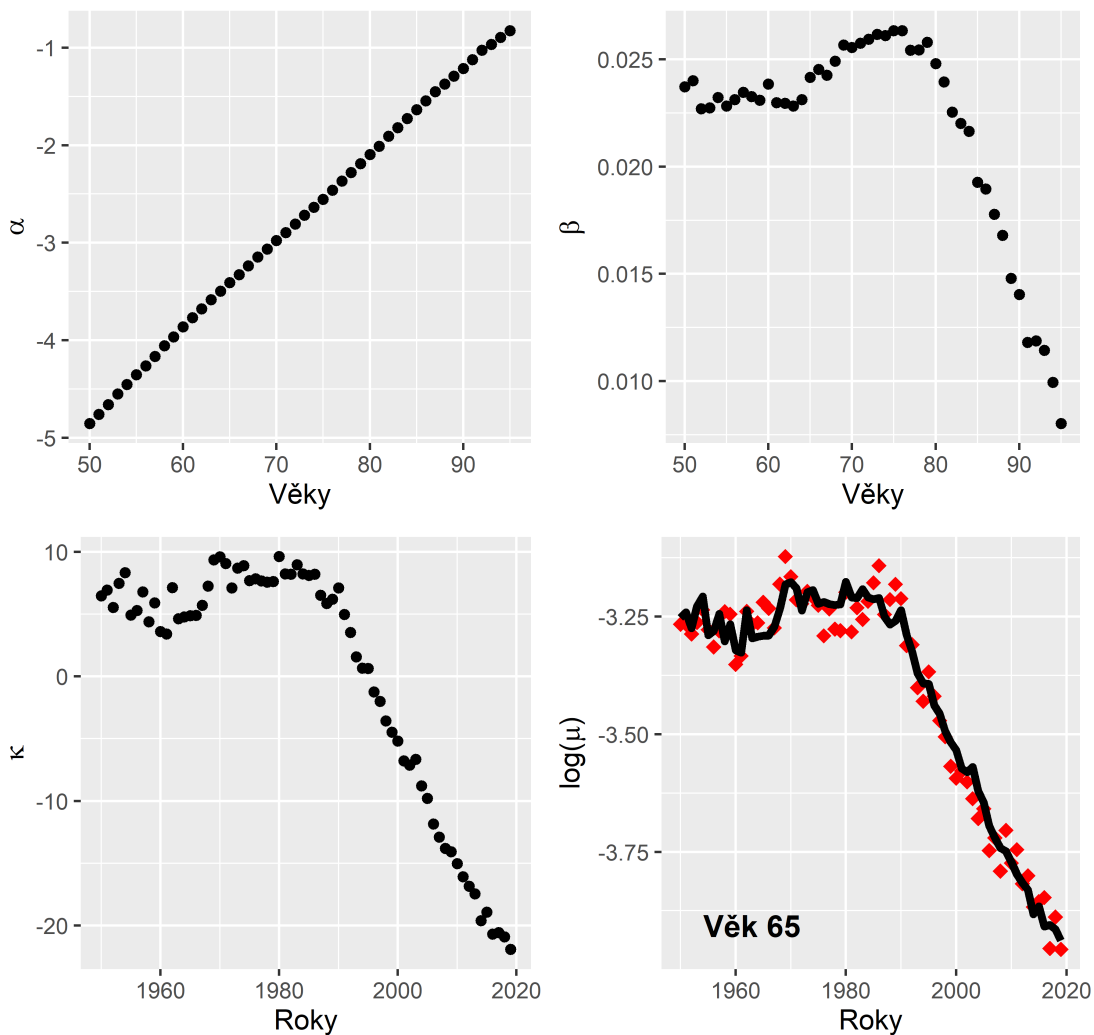
$$(LC) : \log(\mu_{x,y}) = \alpha_x + \beta_x \kappa_y, \quad x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y}.$$

Provedeme odhad pomocí metody (GNM) a poté transformujeme parametry dle (3.3) tak, aby platila jednoznačnost odhadnutých parametrů, a tedy i vztah (3.1). V grafu 6.1 nalezneme odhadnuté parametry modelu pro mužskou populaci a porovnání odhadu věku 65 a skutečně realizovaných úmrtností v jednotlivých letech.

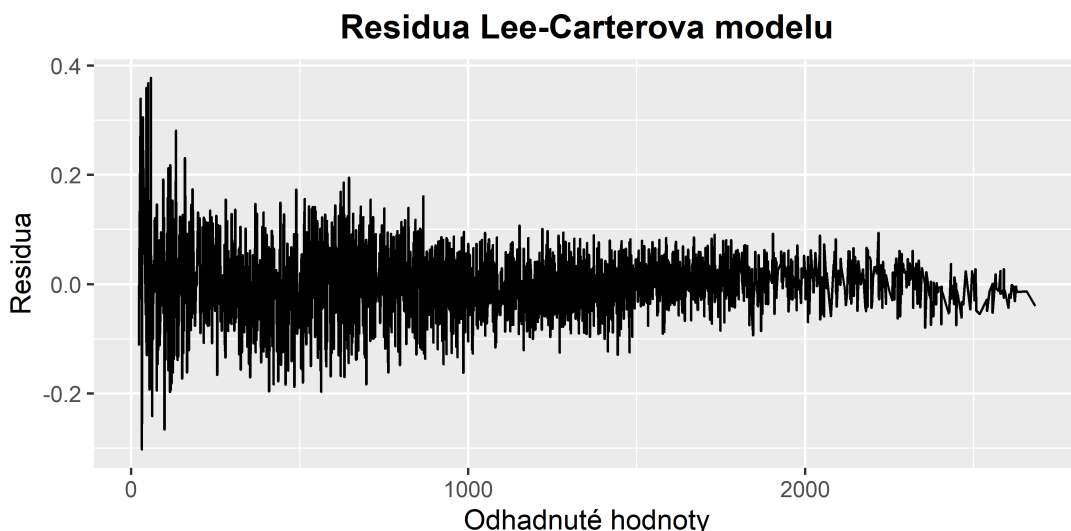
Zaměříme-li se na tento obrázek společně s podmínkami měřítka a lokace definovaných v (3.1), tak můžeme parametry modelu interpretovat následovně: α_x vysvětluje úmrtnost jako průměrný příspěvek daného věku přes všechny roky, β_x pro věk x škáluje parametr κ_y , který představuje změnu (zlepšení) úmrtnosti mezi jednotlivými roky. Zaměříme se na obrázky κ a odhadu úmrtnosti pro věk 65. Můžeme si povšimnout, že obě trajektorie jsou stejné. Toto tvrzení platí pro libovolně zvolený věk a vychází z podstaty modelu a podmínek jednoznačnosti.

V grafu 6.2 pozorujeme residua (LC) modelu pro mužskou populaci. Analyzujeme-li průběh residuí, tak docházíme k závěru, že residua mají nulovou střední hodnotu. Nicméně je patrný trend rozptylu, který postupně klesá s tím, jak model odhaduje novější data. Hodnota Durbin-Watsonovy statistiky jest 1.10, a zamítáme tedy nulovou hypotézu o nekorelovanosti residuí na hladině 0.05.

Odhad Lee-Carterova modelu



Obrázek 6.1: Odhad (LC) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu α_x , β_x a κ_y . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.



Obrázek 6.2: Residua (LC) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.

6.1.2 Lee-Carterův model s kohortním členem

(LC-C) model jsme si zadefinovaly jako:

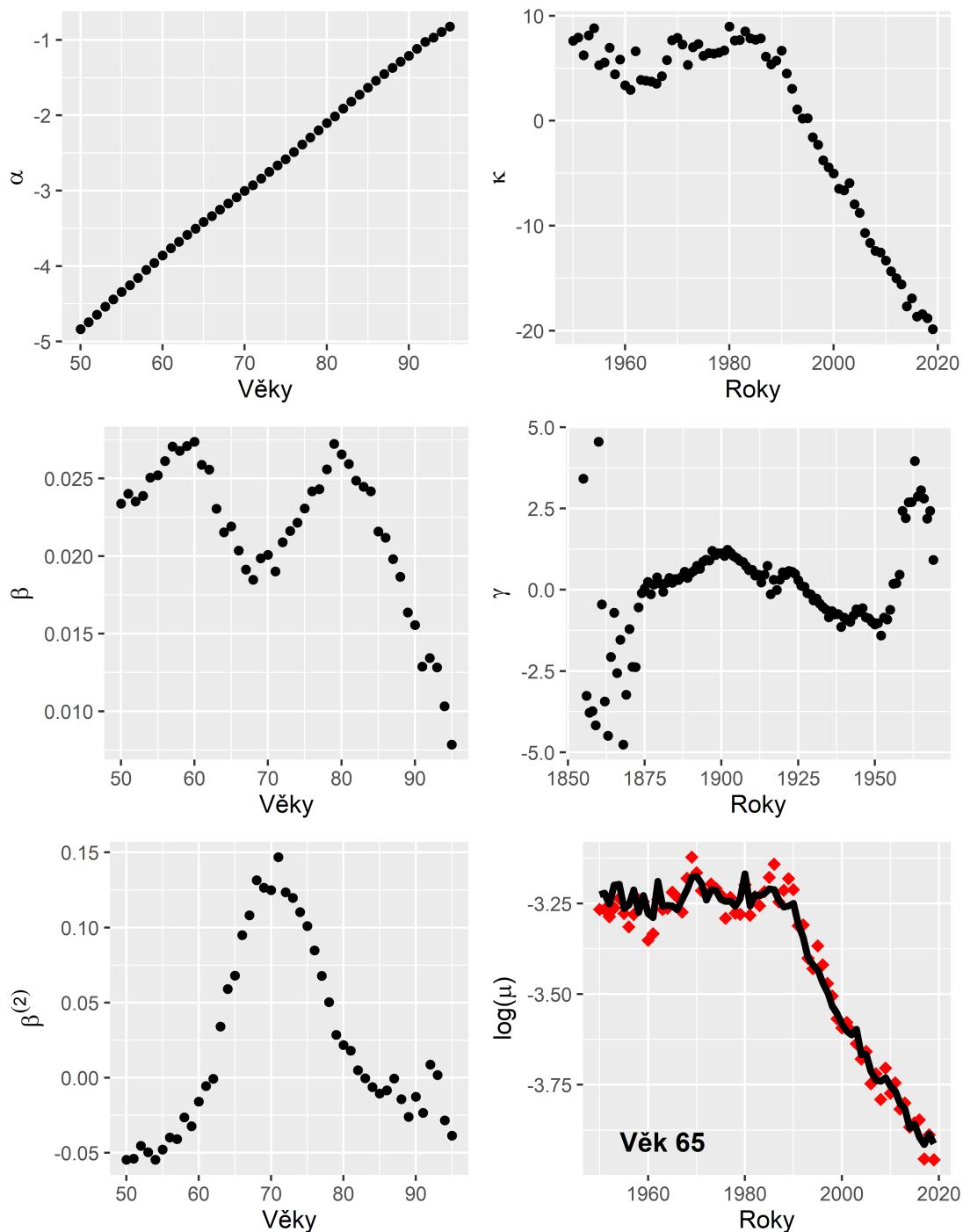
$$(LC - C) : \log(\mu_{x,y}) = \alpha_x + \beta_x \kappa_y + \beta_x^{(2)} \gamma_c, \quad x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y}.$$

Opět provedeme odhad pomocí metody (GNM) a poté transformujeme parametry dle (3.6) tak, aby platila jednoznačnost odhadnutých parametrů, tedy i vztahy (3.4). Poznamenejme, že množství odhadovaných parametrů společně s nelineárním vztahem parametrů zapříčiní, že metoda (GNM) má problémy dokonvergovat k řešení. Z tohoto důvodu první pokus odhadu je vždy spuštěn se startovacími hodnotami, které jsme obdrželi z odhadu (LC) modelu. Pokud takto definovaný model nekonverguje, tak je následně několikrát spuštěn s náhodnými startovacími hodnotami. Časová náročnost a počet iterací je znatelný obzvláště pro velký rozsah roků i věků. Dodejme, že odhad (LC-C) modelu pro ženskou populaci měl velký problém konvergovat kvůli rozsáhlosti a proměnlivosti dat. V grafu 6.3 nalezneme odhadnuté parametry modelu pro mužskou populaci a porovnání odhadu věku 65 a skutečně realizovaných úmrtností v jednotlivých letech.

Podíváme-li se na tento obrázek společně s rozšířenými podmínkami měřítka a lokace definovaných v (3.4), tak můžeme parametry modelu interpretovat následovně: α_x stejně jako u (LC) modelu vysvětluje úmrtnost jako průměrný příspěvek daného věku přes všechny roky, β_x pro věk x škáluje parametr κ_y , který opět představuje změnu (zlepšení) úmrtnosti mezi jednotlivými roky, $\beta_x^{(2)}$ pro daný věk x škáluje kohortní člen γ_c , jež představuje specifické úmrtnostní chování jednotlivých kohort. Poznamenejme, že opět platí fakt shodnosti trajektorií pro κ_y a odhad modelu. Z obrázků příslušných pro γ , $\beta^{(2)}$ a β je patrné, že kohortní člen je významný a vhodný pro modelování úmrtnosti, avšak pouze pro kohorty, které nejsou na kraji intervalu. Ve středu obrázku pozorujeme stabilitu kohort

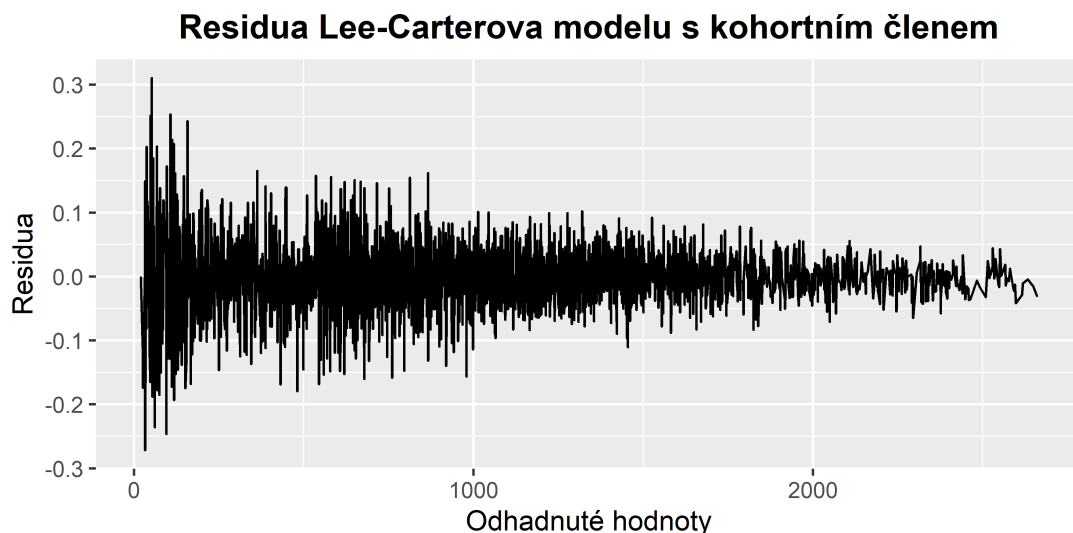
s tím i rostoucí příslušný škálovací člen $\beta^{(2)}$ a tím pádem i klesající škálovací člen β . Naopak na kraji vidíme nestabilitu kohort, způsobenou malým vzorkem dat, který byl použit na jejich odhad.

Odhad Lee-Carterova modelu s kohortním členem



Obrázek 6.3: Odhad (LC-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu α_x , κ_y , β_x , γ_c a $\beta_x^{(2)}$. V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

V grafu 6.4 dodáváme residua (LC-C) modelu pro mužskou populaci. Zaměřili se na průběh residuí, tak docházíme ke stejnému závěru jako pro (LC) model a to, že residua mají nulovou střední hodnotu a rozptyl postupně klesá s tím jak model odhaduje novější data. Hodnota Durbin-Watsonovy statistiky jest 1.43, opět zamítáme nulovou hypotézu o nekorelovanosti residui na hladině 0.05.



Obrázek 6.4: Residua (LC-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.

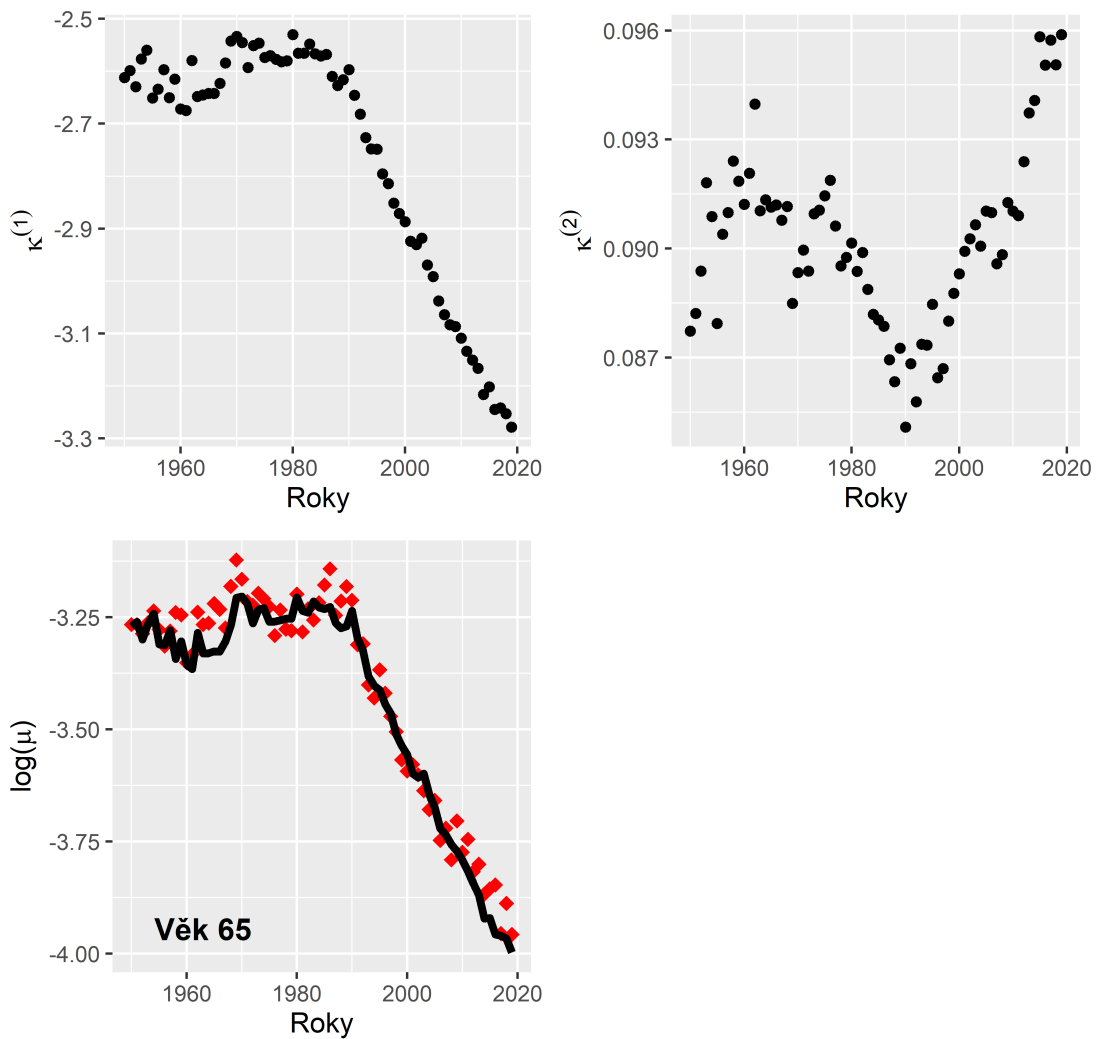
6.1.3 Cairns–Blake–Dowdův model

V teoretické části jsme si (CBD) model definovali následovně:

$$(CBD) : \log(\mu_{x,y}) = \kappa_y^{(1)} + \kappa_y^{(2)}(x - \bar{x}), \quad x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y}.$$

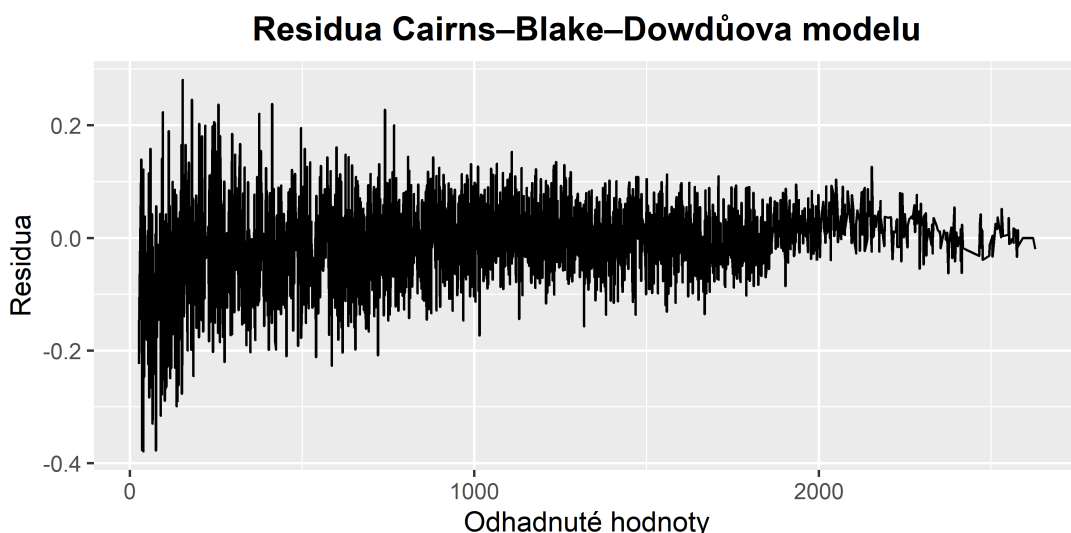
Provedeme odhad pomocí metody (GLM). V grafu 6.5 nalezneme odhadnuté parametry modelu pro mužskou populaci a také porovnání odhadu věku 65 a skutečně realizovaných úmrtností v jednotlivých letech. U konstantního členu $\kappa_y^{(1)}$ je opět vidět podobnost trajektorie s odhadem (CBD) modelu. Povšimneme si opačného chování dvou odhadovaných parametrů. Oba parametry jsou u starších roků bez jednoznačného trendu, poté kolem roku 1990 začne parametr sklonu $\kappa_y^{(2)}$ růst, a proti tomu parametr $\kappa_y^{(1)}$ začne klesat.

Odhad Cairns–Blake–Dowdova modelu



Obrázek 6.5: Odhad (CBD) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají odhadu $\kappa_y^{(1)}$ a $\kappa_y^{(2)}$. V posledním grafu je černou křivkou vyobrazen odhad (CBD) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

V grafu 6.6 dodáváme residua (CBD) modelu pro mužskou populaci. Střední hodnota se zdá u starší roků mírně vychýlena do záporných hodnot, následně se blíží k nule. Trend rozptylu není tak významný jako u (LC) modelů. Nicméně je stále patrný klesající směr. Spočteme-li Durbin-Watsonovu statistiku, tak dostaneme hodnotu 0.82, a opět zamítáme nulovou hypotézu o nekorelovanosti residui na hladině 0.05. Pozitivní autokorelovanost residuí je tedy mnohem silnější než u modelů (LC) a (LC-C).



Obrázek 6.6: Residua (CBD) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.

6.1.4 Cairns–Blake–Dowdův model s kohortním členem

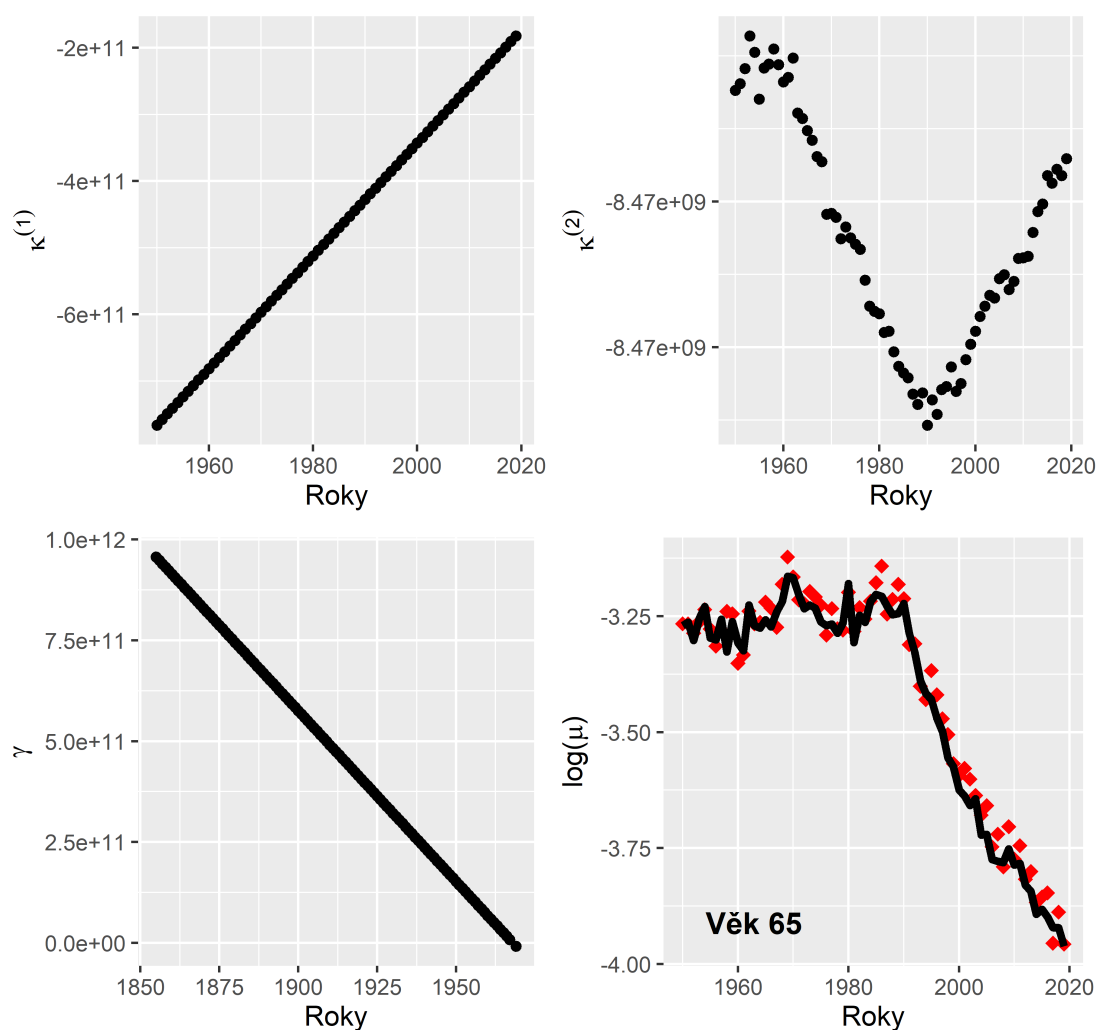
Připomeňme, že (CBD-C) má tvar:

$$(CBD - C) : \log(\mu_{x,y}) = \kappa_y^{(1)} + \kappa_y^{(2)}(x - \bar{x}) + \gamma_c,$$

$$x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y}, c = y_1 - x_{n_x}, \dots, y_{n_y} - x_1.$$

Odhad modelu je dosažen skrze (GLM) metodu. V grafu 6.7 nalezneme odhadnuté parametry modelu pro mužskou populaci, a také porovnání odhadu věku 65 a skutečně realizovaných úmrtností v jednotlivých letech. Pozorujeme, že trend obou konstantních parametrů $\kappa_y^{(1)}$ a γ_c je opačný. U odhadu ženské populace jsou sklony konstantních parametrů prohozené, tedy $\kappa_y^{(1)}$ je klesající a γ_c je rostoucí. U lineárního parametru $\kappa_y^{(2)}$ jsme dostali podobnou trajektorii jako u (CBD) modelu. Přesto, že hodnoty parametrů jsou enormní, tak model věrohodně prokládá skutečné realizace o čemž svědčí poslední obrázek.

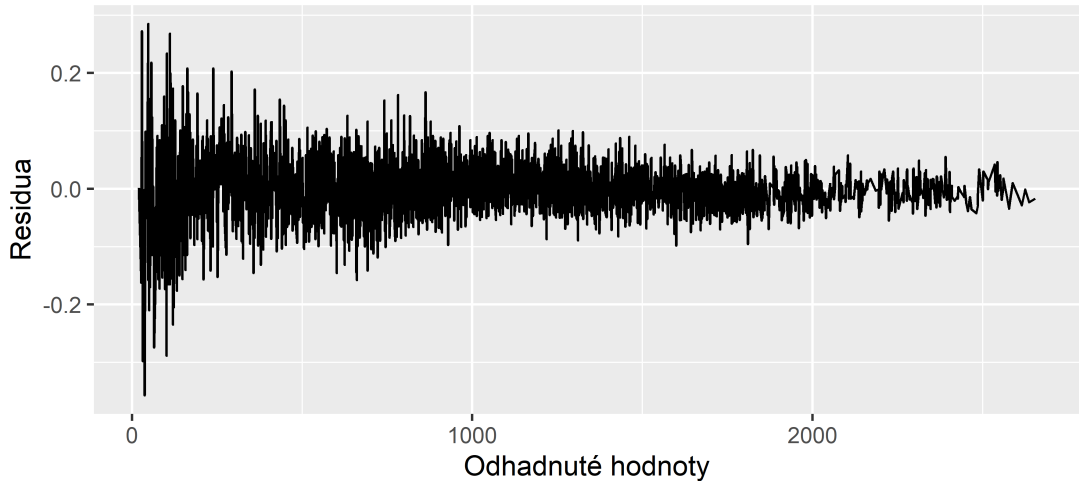
Odhad Cairns–Blake–Dowdůva modelu s kohortním členem



Obrázek 6.7: Odhad (CBD-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}$, $\kappa_y^{(2)}$ a γ_c . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (CBD-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

V grafu 6.8 dodáváme residua (CBD-C) modelu pro mužskou populaci. Střední hodnota je u starších roků opět mírně vychýlena do záporných hodnot, následně se blíží k nule. Trend rozptylu má opět nekonstantní průběh, konkrétně se jedná o klesající směr. Spočteme-li Durbin-Watsonovu statistiku, tak dostaneme hodnotu 1.46, a opět zamítáme nulovou hypotézu o nekorelovanosti residui na hladině 0.05. Pozitivní autokorelovanost residuí je patrná pro mladší roky.

Residua Cairns–Blake–Dowdůva modelu s kohortním členem



Obrázek 6.8: Residua (CBD-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.

6.1.5 Cairns–Blake–Dowdův model s variačním členem

Připomeňme, že (CBD-V) má tvar:

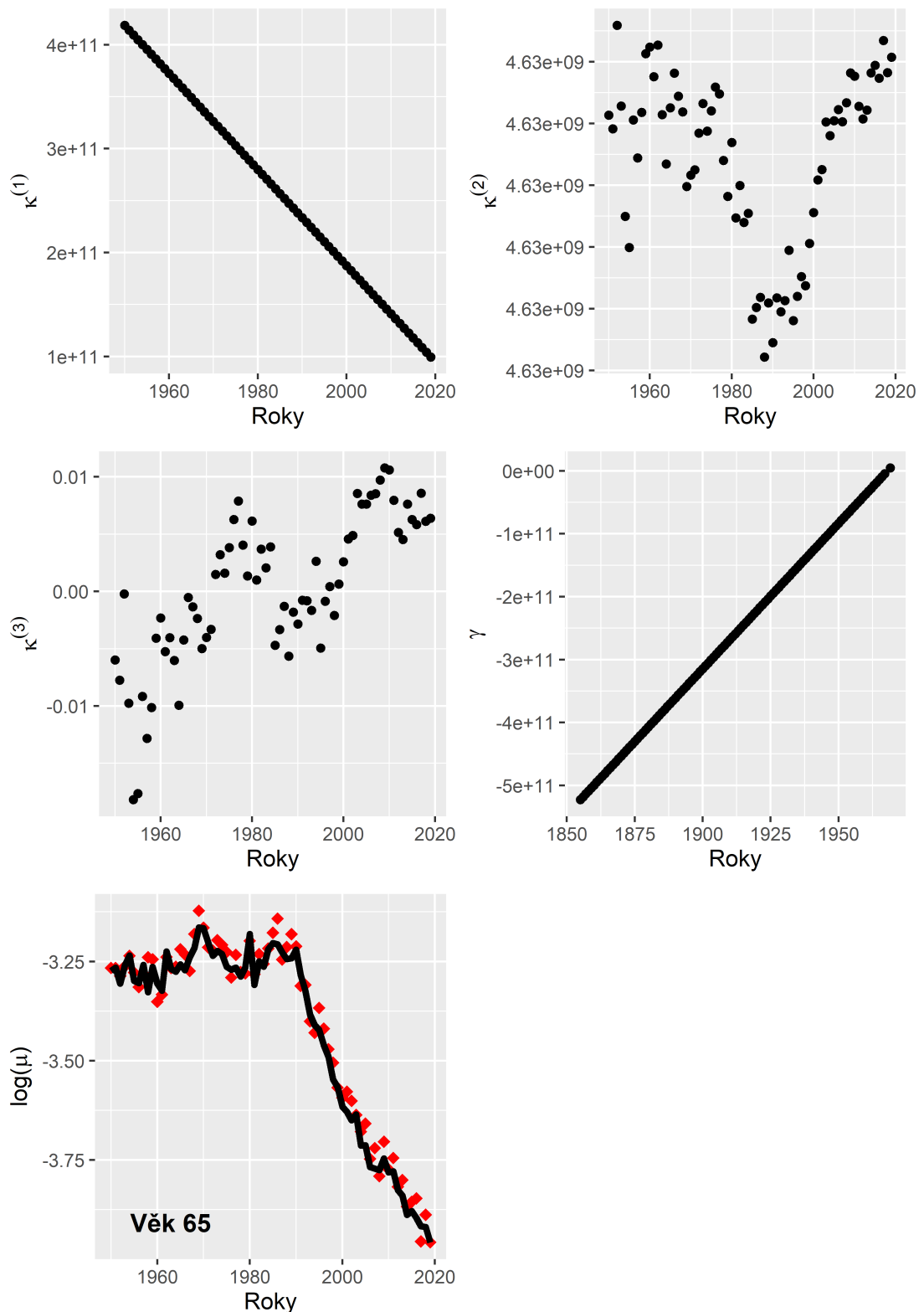
$$(CBD - V) : \log(\mu_{x,y}) = \kappa_y^{(1)} + \kappa_y^{(2)}(x - \bar{x}) + \kappa_y^{(3)}|(x - \bar{x}) - \varsigma_x| + \gamma_c,$$

$$x = x_1, \dots, x_{n_x}, y = y_1, \dots, y_{n_y}, c = y_1 - x_{n_x}, \dots, y_{n_y} - x_1.$$

Odhad modelu je stejně jako u předchozích variant (CBD) modelu odhadnut metodou (GLM). V grafu 6.9 nalezneme odhadnuté parametry modelu pro mužskou populaci a také porovnání odhadu věku 65 a skutečně realizovaných úmrtností v jednotlivých letech. Totožně jako u (CBD-C) modelu platí, že trend obou konstantních parametrů $\kappa_y^{(1)}$ a γ_c je opačný. U lineárního parametru $\kappa_y^{(2)}$ můžeme opět vidět tvar V, který byl charakteristický pro obě předchozí varianty. Rozšiřující parametr $\kappa_y^{(3)}$ má rostoucí konkávní sekvenci s propadem v roce 1990.

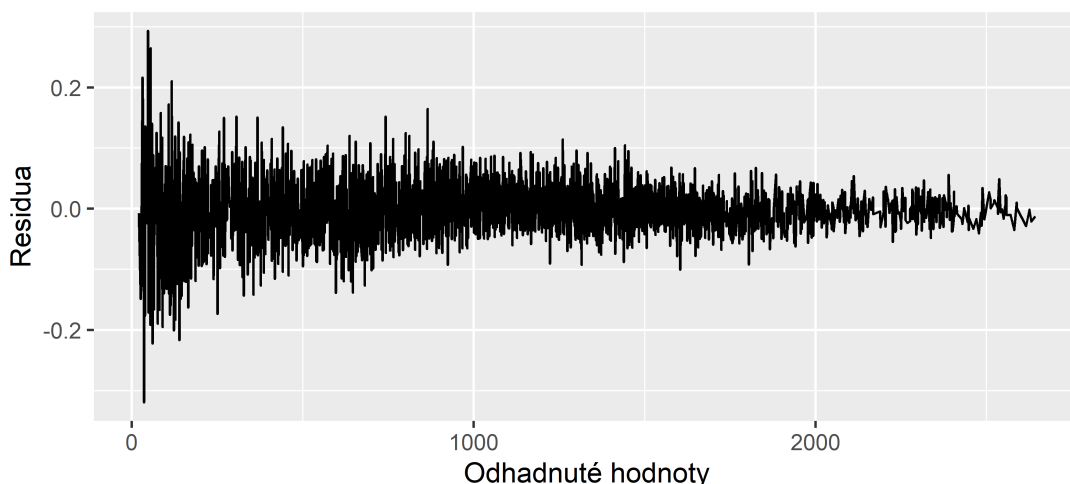
V grafu 6.10 dodáváme residua (CBD-V) modelu pro mužskou populaci. Vlastnosti střední hodnoty a rozptylu residuí jsou totožné k zbylým dvěma variantám. Durbin-Watsonova statistika má hodnotu 1.55, a opět zamítáme nulovou hypotézu o nekorelovanosti residuí na hladině 0.05. Pozitivní autokorelovanost residuí je patrná pro mladší roky.

Odhad Cairns–Blake–Dowdůva modelu s variačním členem



Obrázek 6.9: Odhad (CBD-V) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}$, $\kappa_y^{(2)}$, $\kappa_y^{(3)}$ a γ_c . V posledním grafu je černou křivkou vyobrazen odhad (CBD-V) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

Residua Cairns–Blake–Dowduova modelu s variačním členem



Obrázek 6.10: Residua (CBD-V) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.

6.2 Empirická diagnostika modelů

V této kapitole budeme zkoumat predikční vlastnosti modelů definovaných v podkapitole 3.2 *Podstata modelů a jejich třídy*. Využijeme k tomu testy definované v podkapitole 3.4 *Diagnostika modelů*.

Nejdříve uvedme, že pro vyhodnocování jednotlivých testů se budeme omezovat jen na určitou část aktuálnějších dat, typicky blíže k přelomu tisíciletí. Testy budeme provádět zvlášť pro obě pohlaví, neboť chování úmrtnosti v pozdějších stádiích života má diametrální rozdíly pro muže a ženy. Naše hlavní zaměření cílí právě na tyto pozdější věky.

Nyní určíme věky, na které se zaměříme při testování modelů. Jinými slovy, jaké věky jsou pro naše účely nejvíce podstatné. Jak již bylo zmíněno v kapitole 4. *Pojištění matematická rizika v kontextu úmrtnosti*, tak jedním z nejvíce materiálních rizik pro pojišťovnu je riziko dlouhověkosti. V další podkapitole 6.3 *Ocenění longevity dluhopisu* budeme počítat cenu zajistných kontraktů pro portfolio anuitantů, kde věrohodnost a správnost modelů v predikci úmrtnosti v pozdějším věku je esenciální. Z těchto důvodů budeme zkoumat vlastnosti modelů pro věk 65, ke kterému se momentálně blíží věk odchodu do starobního důchodu v České republice. Dodejme, že věk odchodu do starobního důchodu se liší jak přes pohlaví, tak u žen dle počtu vychovaných dětí. Nicméně věk 65 se dá považovat jako vhodný pro účely stanovení začátku starobního důchodu v České republice. Poznamenejme, že zbývající střední doba dožití věků 65 je pro pojišťovny, které mají velké pojistné kmeny důchodových anuit, esenciální parametr.

V několika posledních dekádách byla signifikantně zvýšena střední délka života pro muže a ženy, avšak tento růst byl zastaven pandemií Covid-19. V tabulce 6.2 nalezneme střední doby dožití novorozenců v České republice pro vybrané roky mezi 1993 a 2020. Data jsou získaná z CSU (2021). Pozorujeme rychle rostoucí trend do roku 2019 po kterém následuje pokles v roce 2020 kvůli pandemii. Naším dalším zkoumaným věkem bude 85 let. Tento výběr je motivován faktem, že střední doba dožití pro obě pohlaví při věku 85 je v jednotkách let, a tím pádem očekávání jest, že budoucí predikovaní přežívší věku 85 budou rychle klesat. Pokud bude však predikce klesat s pomalejší dynamikou, než pojišťovny předpokládaly při stanovení pojistného důchodových kontraktů, tak to pro ně nutně znamená velké budoucí finanční ztráty.

pohlaví	1993	2000	2010	2019	2020
Muži	69.3	71.6	74.4	76.3	75.3
Ženy	76.5	78.4	80.6	82.1	81.4

Tabulka 6.2: Vývoj střední doby dožití novorozenců v České republice pro obě pohlaví v uvažovaném horizontu od roku 1993 do 2020.

V následujících podkapitolách provedeme jednotlivé testy. Měřenou náhodnou veličinou je již dříve zmíněná intenzita úmrtnosti, která je ovšem aproximovaná centrální úmrtnostní měrou, dle vztahu (2.3). Pro všechny porovnání budeme uvažovat délku historického okénka v rozsahu $h = 20$ let. Jinými slovy, veškeré naše odhady v této podkapitole budou založeny na historických datových oknech o délce 20-ti let. Pro každý test, model, věk a pohlaví budeme hodnotit, zda daný model prošel (✓) nebo ne (✗). Kategorii věku budeme hodnotit jako celek, neboť od modelu požadujeme, aby byl úspěšných u obou testovaných věků. Nepřipouštím tedy možnost použití vícero typů modelů pro různé věky, ovšem připouštíme použití jiných modelů pro konkrétní pohlaví. Typicky vyjdeme z grafického porovnávání a naměřených statistik. Poznamenejme, že u všech testů volíme 90% predikční interval.

6.2.1 Test konvergence zužujícího horizontu

Tento test sleduje konvergenci predikce ke skutečné hodnotě. Zvolíme si rok $Y_1 := 2015$ a $n_1 := 15$. Provedeme tedy 15-krát odhad modelů od roků 1981, ..., 1995 do roků 2000, ..., 2014 a budeme predikovat rok 2015. Jinými slovy, délky budoucích predikcí budou 15, 14, ..., 1 rok. Vhodný model by měl postupně konzistentně konvergovat ke skutečné hodnotě roku 2015. Zobrazíme výsledky mužů do grafů 6.11 a 6.12, žen do grafů 6.13 a 6.14 pro věk 65 respektive 85, společně s predikčními intervaly.

- Pro (LC) model pozorujeme konzistentní konvergenci ke skutečné hodnotě pro obě pohlaví. Poslední jednoroční předpověď je mírně vychýlena, to způsobí, že dvouletá předpověď z roku 2013 je přesnější než jednoletá předpověď z roku 2014. Nicméně tento fakt se týká většiny testovaných modelů a můžeme tedy (LC) model považovat za úspěšných pro obě pohlaví.

- U (LC-C) modelu pozorujeme signifikantní skoky v predikci. Více jsou patrné u žen, kde nemůžeme hovořit o konzistentní konvergenci ke skutečné hodnotě. Pro oba věky žen pozorujeme, že predikce z roku 2008 je významně vychýlena. Poté se předpověď postupně vrací k skutečné hodnotě, ale nedostatečně rychle. (LC-C) model můžeme doporučit jako vhodný pro mužské predikce neb i přes znatelné vychýlení se model rychle vrací ke skutečné hodnotě, toto tvrzení ovšem neplatí u žen, proto ho nepovažujeme za úspěšný.
- Základní (CBD) model dosahuje dobrých výsledků pro ženy s mírným podhodnocujícím trendem v kratších predikcích. Tento fakt platí i pro muže, kde u věku 85 je ještě akceptovatelný, avšak u věku 65 toto tvrdit nelze. Z tohoto důvodu nemůžeme (CBD) model považovat jako vhodný pro muže.
- Model (CBD-C) projevuje opět znaky podhodnocení budoucí realizace stejně jako drtivá většina modelů v tomto testu. U žen ve věku 65 v grafu 6.13 se tento trend překvapivě otočil a model nadhodnocuje predikci oproti realizaci v roce 2015. Model však konzistentně konverguje ke skutečné hodnotě ve všech případech, a proto ho považujeme za úspěšný pro obě pohlaví.
- Poslední (CBD-V) model dosahl nejlepších výsledků v tomto testu. Je třeba zkoumat obrázky podrobně, neboť množství predikovaných parametrů je u tohoto modelu vyšší než u ostatních, a tím pádem je i predikční interval společně s měřítkem rozšířeny. Nicméně i tak je patrná nejlepší stabilní predikce a to dokonce i u dlouhých oken předpovědi.

Výsledky prvního testu nalezneme v tabulce 6.3

model	LC	LC-C	CBD	CBD-C	CBD-V
Muži	✓	✓	✗	✓	✓
Ženy	✓	✗	✓	✓	✓

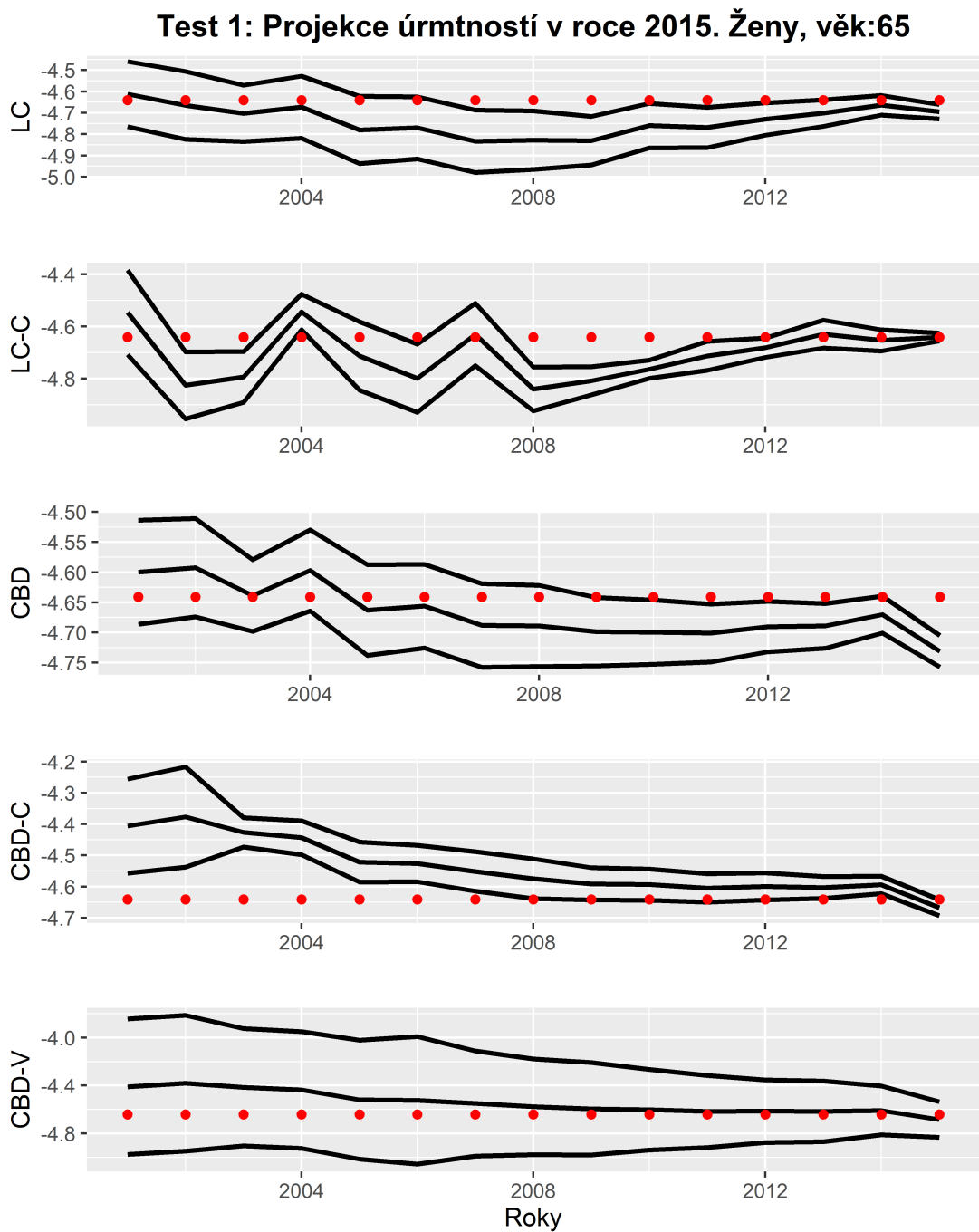
Tabulka 6.3: Výsledky testu konvergence zužujícího horizontu.



Obrázek 6.11: Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1981, ..., 1995 do roků 2000, ..., 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, ..., 1-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.



Obrázek 6.12: Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1981, ..., 1995 do roků 2000, ..., 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, ..., 1-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.



Obrázek 6.13: Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1981, ..., 1995 do roků 2000, ..., 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, ..., 1-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.



Obrázek 6.14: Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1981, ..., 1995 do roků 2000, ..., 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, ..., 1-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.

6.2.2 Test dlouhé predikce

V dalším testu se zaměříme na to, jak jednotlivé modely věrohodně predikují delší predikce do budoucnosti. To je jedna z klíčových vlastností dobrých modelů. Zvolíme si rok $Y_2 := 2000$ a $n_2 := 15$. Odhadneme daný model z dat od roku 1981 do 2000 a poté predikujeme jednoroční, dvouroční, ..., patnáctiroční předpověď budoucích hodnot, které porovnáme se skutečnými realizacemi. Zaměříme se na hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$, kterou jsme si definovali v teoretické části textu. Poznamenejme, že ideální model by měl prokládat realizované hodnoty středem, tedy $x_{C_L} \cong x_{C_U}$ a počet hodnot mimo 90% predikční interval by měl být do $n_2 * \alpha = 15 * 0.1 = 1.5$ pro jednotlivý věk, tedy pro oba věky by počet překročení měl být do 3. Dosažené výsledky mužů nalezneme v grafech 6.15 a 6.16, žen v grafech 6.17 a 6.18 pro věk 65 respektive 85, společně s predikčními intervaly. Hodnoty statistik $[x_U, x_L, x_{C_U}, x_{C_L}]$ nalezneme v tabulkách 6.4 a 6.5 pro muže respektive ženy.

model	LC	LC-C	CBD	CBD-C	CBD-V
x_U	3	2	0	2	0
x_L	0	10	7	5	0
x_{C_U}	19	4	12	14	14
x_{C_L}	11	26	18	16	16

Tabulka 6.4: Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test dlouhé predikce mužů, oba věky.

model	LC	LC-C	CBD	CBD-C	CBD-V
x_U	1	2	1	2	0
x_L	5	15	7	14	0
x_{C_U}	3	5	13	16	6
x_{C_L}	27	25	17	14	24

Tabulka 6.5: Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test dlouhé predikce žen, oba věky.

- (LC) model dobře predikuje v případě věku 65. Mužská předpověď je proložena středem realizovaných hodnot, u žen je v první polovině intervalu predikce nadhodnocena oproti skutečně realizovaným hodnotám, v druhé části je patrný návrat ke středu hodnot. Pro věk 85 je mužská predikce vychýlena pod realizované hodnoty, a model tedy podhodnocuje budoucí realizace. U žen je ke konci patrné slabé nadhodnocení. U ženské predikce jsme naměřili celkem šest hodnot mimo predikční interval, nicméně z grafu 6.17 je patrné, že hodnoty leží téměř na spodním rameni intervalu. U mužského odhadu jsme dostali celkem tři hodnoty, které leží mimo predikční pás, tedy predikce se pohybuje v rámci tolerance. Z výše uvedeného považujeme (LC) model za úspěšný pro obě pohlaví.
- (LC-C) model dosahl dobrých výsledků pouze pro muže ve věku 65. Pro ostatní predikce jsou výsledky nedostatečné o čem svědčí jak grafické zobrazení, tak i statistiky x_U a x_L z tabulek 6.4 a 6.5, které celkem naměřily 29

překročení predikčních intervalů. Tím pádem nelze (LC-C) model považovat za úspěšný ani pro jedno pohlaví.

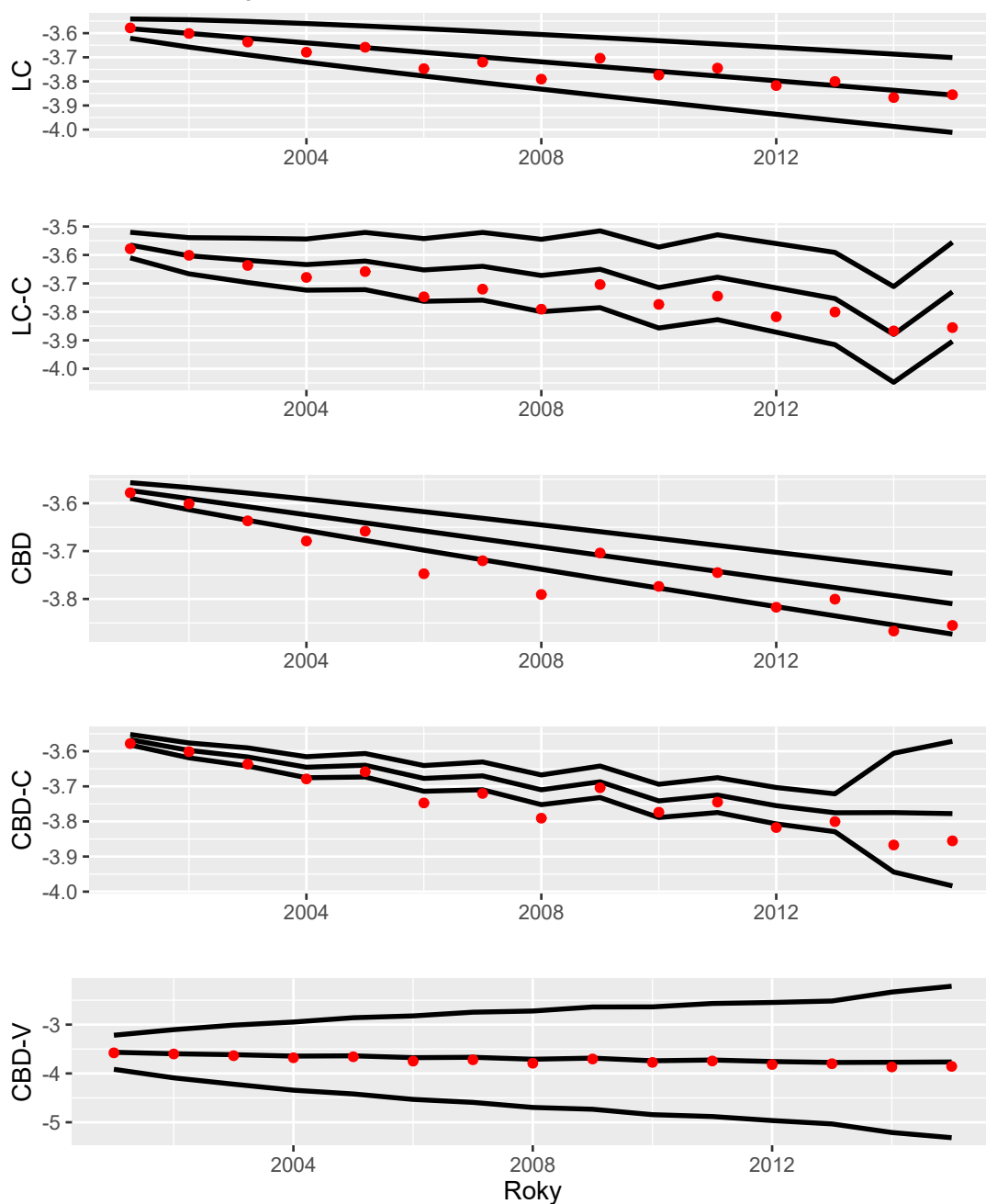
- (CBD) model pro věk 65 v první polovině intervalu značně nadhodnocuje predikci pro obě pohlaví. Ve druhé části se predikce narovnává na skutečné hodnoty. Výsledky pro věk 85 jsou uspokojivě jak pro ženy, tak pro muže. Uzavíráme s relativně uspokojivým výsledkem pro muže, neboť odlehlé realizace věku 65 jsou poměrně blízko dolnímu ramenu intervalu. Nicméně pro ženy nelze dospět ke stejnému závěru.
- Výsledky predikce modelu (CBD-C) pro věk 65 u mužů jsou nadhodnoceny, kde čtyři hodnoty leží mimo predikční interval. Pro ženy je predikce naprosto vychýlena a pouze jedna hodnota leží v intervalu. Naopak pro věk 85 jsou dosažené výsledky uspokojivé pro obě pohlaví s mírným podhodnocením. Diametrální rozdíl mezi predikcí pro jednotlivé věky je způsoben kohortním členem, který je pro starší věky odhadnout na větším vzorku dat než pro mladší věky. Pro věk 65 můžeme pozorovat, že v roce 2014 již dochází k predikci kohortního členu a tím pádem i k rozšíření ramen predikčního intervalu. (CBD-C) model považujeme za úspěšný pouze pro muže, neboť hodnoty pod predikčním intervalem jsou poměrně blízko hranice pásu.
- Predikce modelu (CBD-V) je nejlepší ze všech uvedených modelů. Žádná z hodnot nepřekročila predikční pas. To je ovšem z části způsobeno poměrně širokým intervalem, neboť nejistota obsažena v rozpětí predikčního intervalu je daná odhadem tři neznámých parametrů a pro věk 65 ke konci dokonce čtyřmi parametry. Nicméně tabulka 6.4 ukazuje, že predikce pro muže je proložena středem realizovaných hodnot. Pro ženy bylo dosaženo nadhodnocení skutečným hodnot, o čemž svědčí hodnota x_{CL} z tabulky 6.5. Model (CBD-V) považujeme za úspěšný pro obě pohlaví.

Výsledky druhého testu nalezneme v tabulce 6.6

model	LC	LC-C	CBD	CBD-C	CBD-V
Muži	✓	✗	✓	✓	✓
Ženy	✓	✗	✗	✗	✓

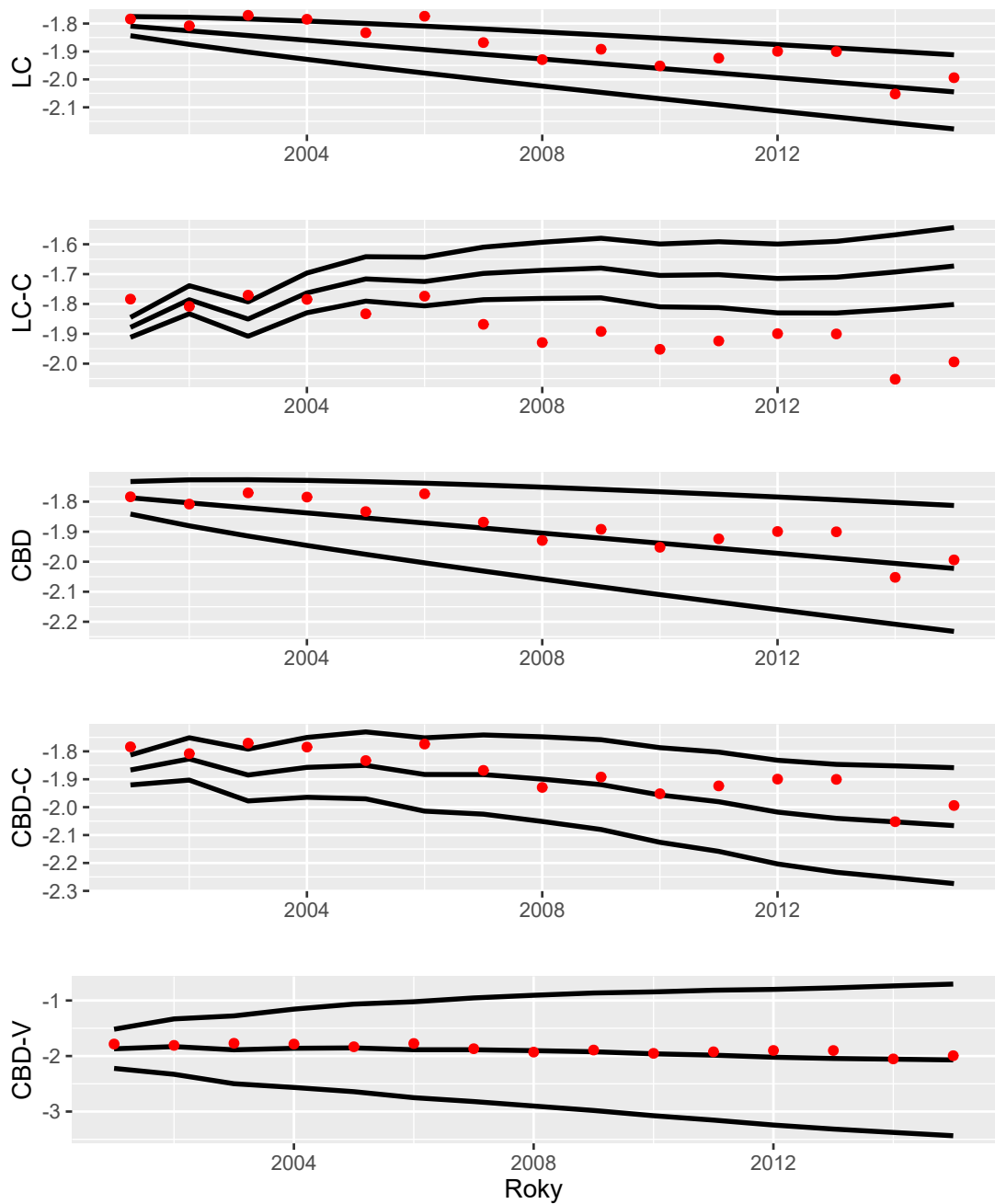
Tabulka 6.6: Výsledky testu dlouhé predikce.

Test 2: Projekce úmrtností od roku 2001 do 2015. Muži, věk:65



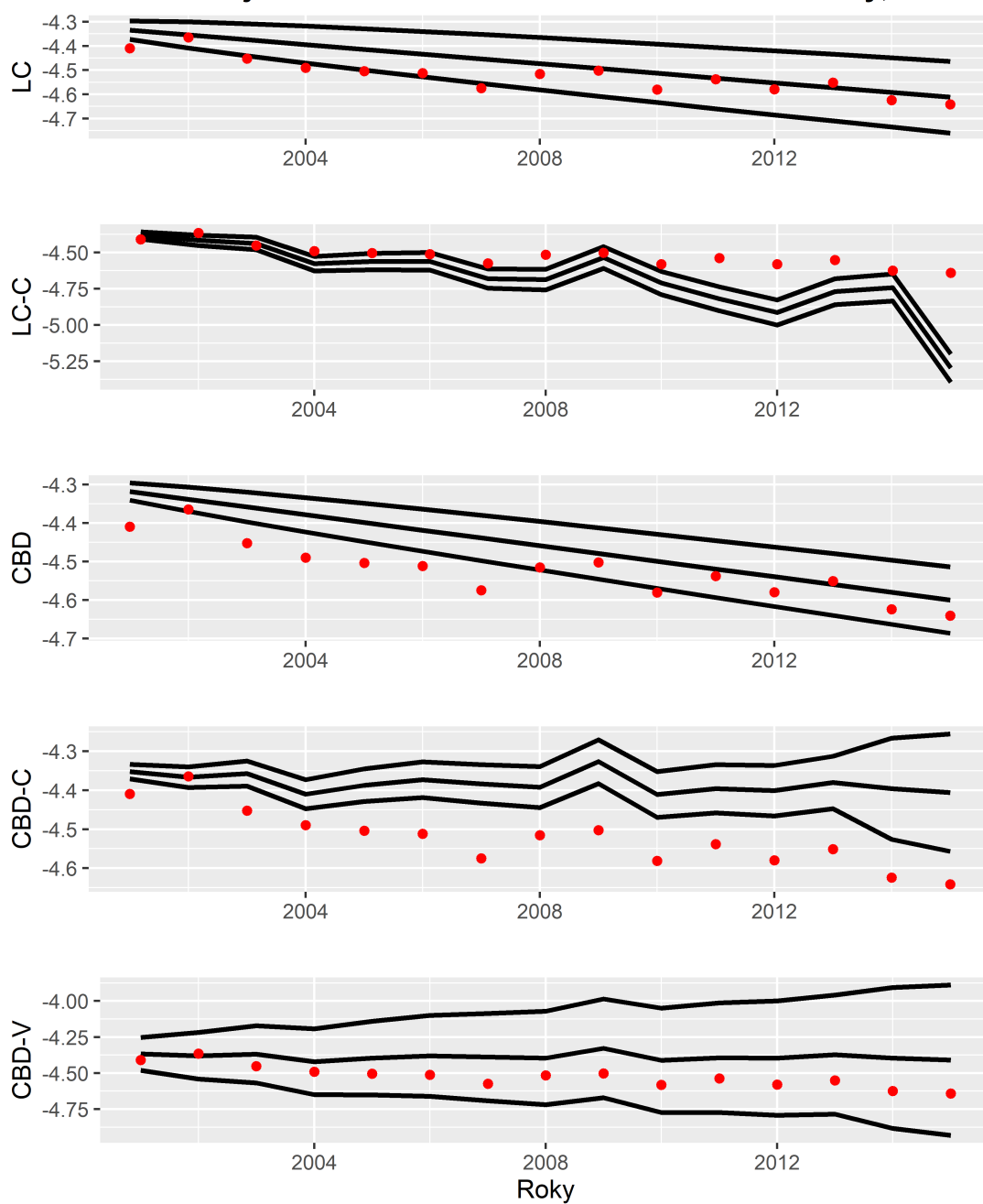
Obrázek 6.15: Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.

Test 2: Projekce úmrtností od roku 2001 do 2015. Muži, věk:85



Obrázek 6.16: Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.

Test 2: Projekce úmrtností od roku 2001 do 2015. Ženy, věk:65



Obrázek 6.17: Predikce založena na českých datech úmrtnosti žen pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.

Test 2: Projekce úmrtností od roku 2001 do 2015. Ženy, věk:85



Obrázek 6.18: Predikce založena na českých datech úmrtnosti žen pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.

6.2.3 Test posouvajícího se okna predikce

Zde budeme zkoumat vývoj a stabilitu predikce, při pohyblivém okně predikce. Zvolíme si rok $Y_3 := 1986$, $n_{3_s} := 20$ a $n_{3_e} := 30$. Celkem budeme odhadovat $n_{3_e} - n_{3_s} = 10$ -krát modely založené na datech od roků 1967, ..., 1976 do roků 1986, ..., 1995 a poté provedeme pro každý odhad dvacetiroční predikci do budoucnosti, tedy do roků 2006, ..., 2015. Porovnáme se skutečně realizovanými hodnotami v daných letech. Opět budeme sledovat čtveřici statistik $[x_U, x_L, x_{C_U}, x_{C_L}]$, kterou jsme si definovali v teoretické části textu. Dosažené výsledky mužů nalezneme v grafech 6.19 a 6.20, žen v grafech 6.21 a 6.22 pro věk 65 respektive 85, společně s predikčními intervaly. V tabulkách 6.7 a 6.8 dodáváme hodnoty statistik $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro oba věky mužů respektive žen.

model	LC	LC-C	CBD	CBD-C	CBD-V
x_U	0	x	0	0	0
x_L	18	x	15	6	0
x_{C_U}	0	x	0	2	2
x_{C_L}	20	x	20	18	18

Tabulka 6.7: Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test posouvajícího se okna predikce mužů, oba věky.

model	LC	LC-C	CBD	CBD-C	CBD-V
x_U	0	x	0	0	0
x_L	16	x	11	9	0
x_{C_U}	0	x	0	3	0
x_{C_L}	20	x	20	17	20

Tabulka 6.8: Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test posouvajícího se okna predikce žen, oba věky.

- (LC) model silně nadhodnocuje oba věky v predikci pro muže. Pro věk 65 dokonce ani jedna realizovaná hodnota neleží v predikčním intervalu, jak je vidět na grafu 6.19. Predikce pro věk 85 na konci horizontu konvergovaly ke skutečným hodnotám, ale i tak je patrné markantní nadhodnocení. V případě žen vidíme slabší nadhodnocení, které je více znatelné u věku 65. (LC) model má očividně problém věrohodně predikovat úmrtnosti mužů, u žen není patrná signifikantní trendová odchylka od skutečných hodnot proto tuto predikci lze považovat za vhodnou.
- (LC-C) model byl schopen úspěšně konvergovat pouze u žen pro věk 85. U ostatních odhadů, obzvláště u věků 65 měl problém s konvergencí. Spousta hodnot nebyla predikována, z tohoto důvodu neuvádíme čtveřici hodnotících statistik, protože by nebyla konzistentně porovnatelná s ostatními modely. Neschopnost odhadnutí daných roků a problém extrémně širokých ramen modelu je způsobeno nelineárním vztahem parametrů a velkým datovým vzorkem, a to hlavně v dimenzi věků. Tím pádem (LC-C) model není úspěšný v tomto testu.

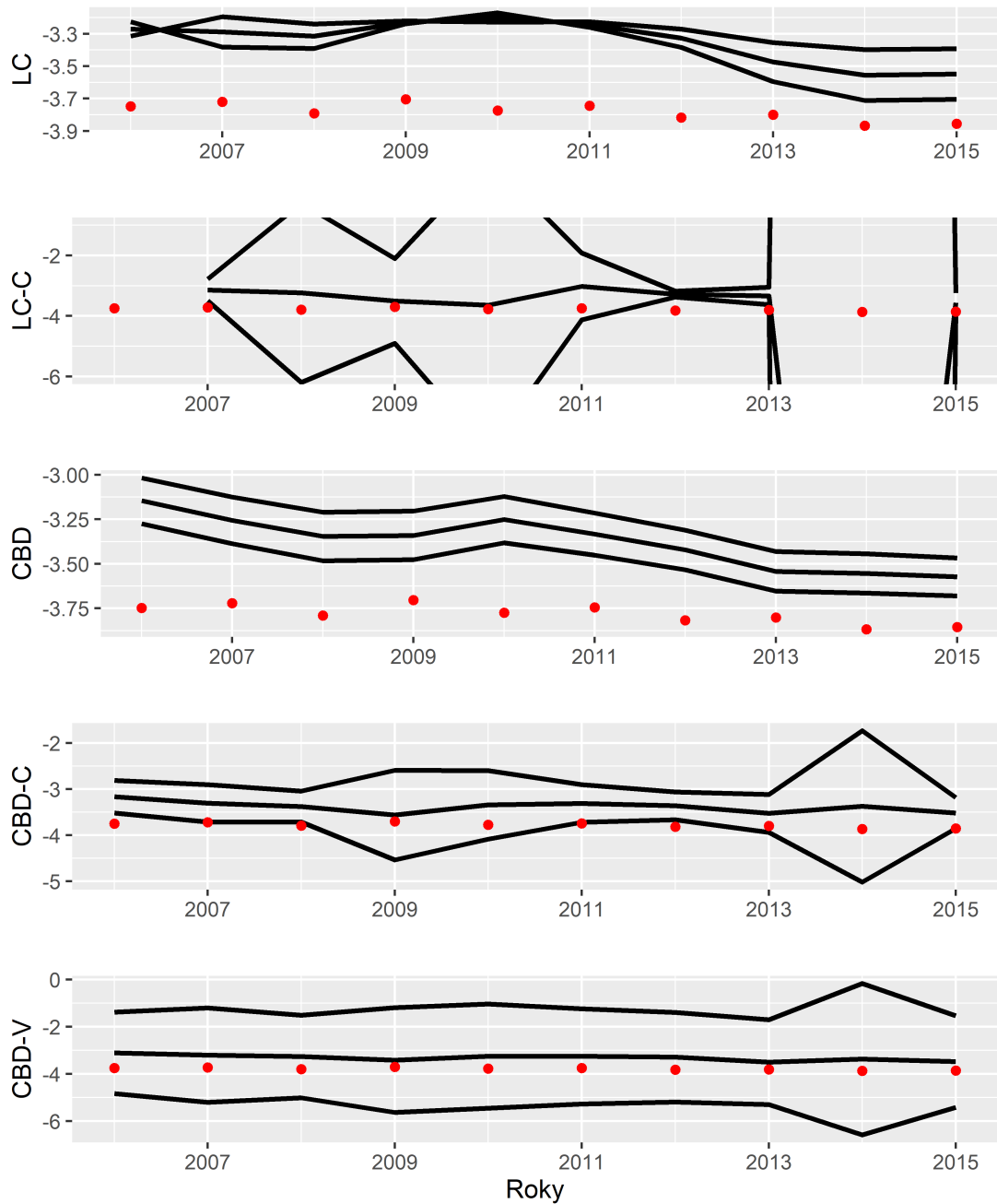
- (CBD) model silně nadhodnocuje pro věky 65. Zaměříme-li se na věky 85, tak stále pozorujeme nadhodnocení pro oba věky, avšak většina hodnot již leží v predikčních intervalech. Nevhodnost předpovědi pro věk 65 nás vede k hodnocení (CBD) modelu jako nevhodný pro obě pohlaví.
- Model(CBD-C) má podobnou trajektorii jako předcházející (CBD) model. Avšak centrální predikce mnohem lépe prokládá skutečné realizace a také mnohem víc bodů leží v predikčních intervalech o čem svědčí i rozdíl ve statistikách x_L a x_U , který je poměrně znatelný u obou pohlaví, přičemž u mužů významněji. (CBD-C) považujeme za úspěšný pro obě pohlaví.
- Predikční interval (CBD-V) modelu obsahuje všechny skutečné realizace, což je důsledkem širších ramen spojených s větší náhodou obsaženou v predikci. Centrální predikce poměrně kopíruje centrální predikci (CBD-C) modelu, což je patrné ze statistik x_{CL} a x_{CU} . Model považujeme za vhodný pro obě pohlaví.

Výsledky třetího testu nalezneme v tabulce 6.9

model	LC	LC-C	CBD	CBD-C	CBD-V
Muži	✓	✗	✗	✓	✓
Ženy	✗	✗	✗	✓	✓

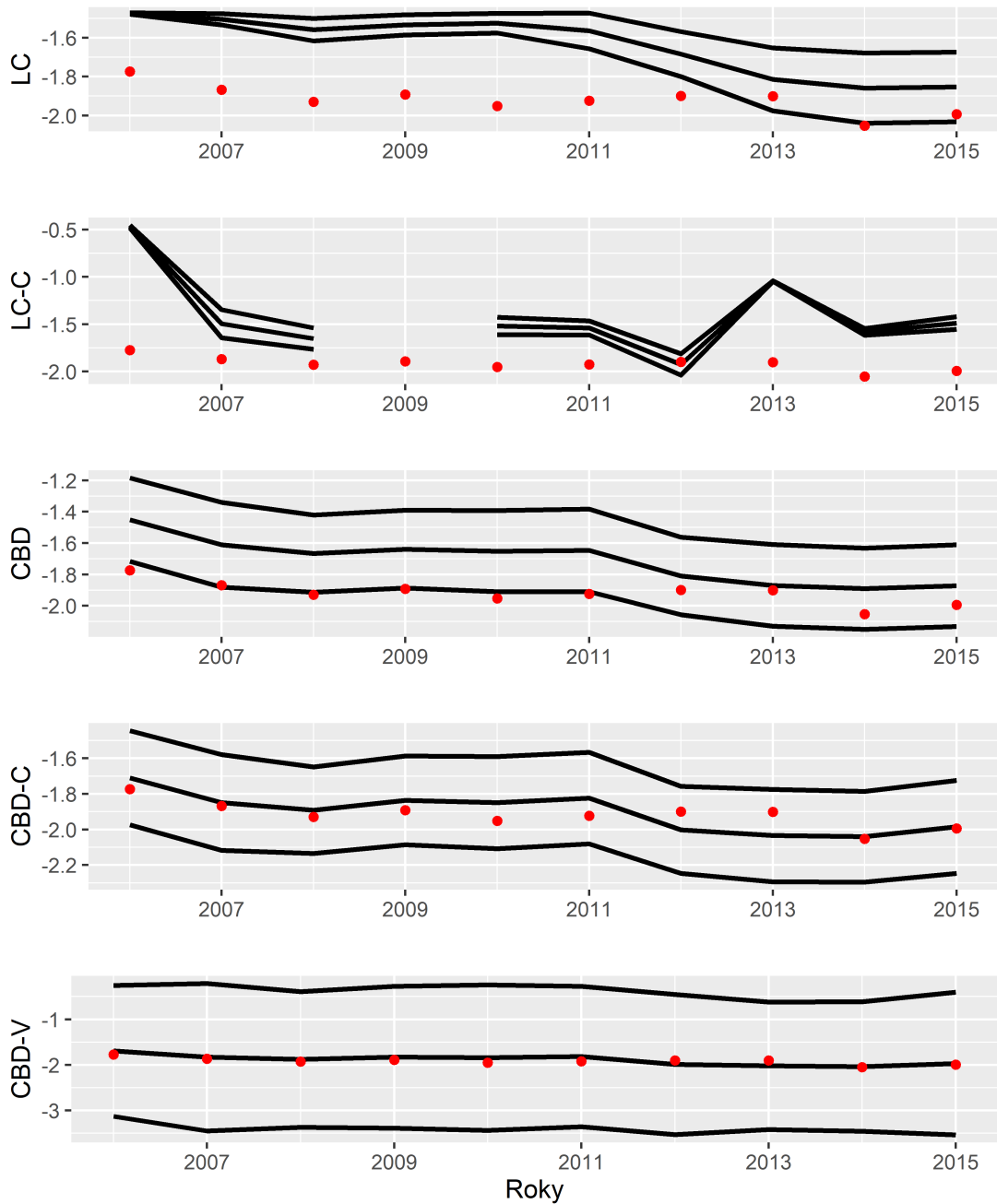
Tabulka 6.9: Výsledky testu posouvajícího se okna predikce.

Test 3: Projekce hodnot od roku 2006 do 2015. Muži, věk:65

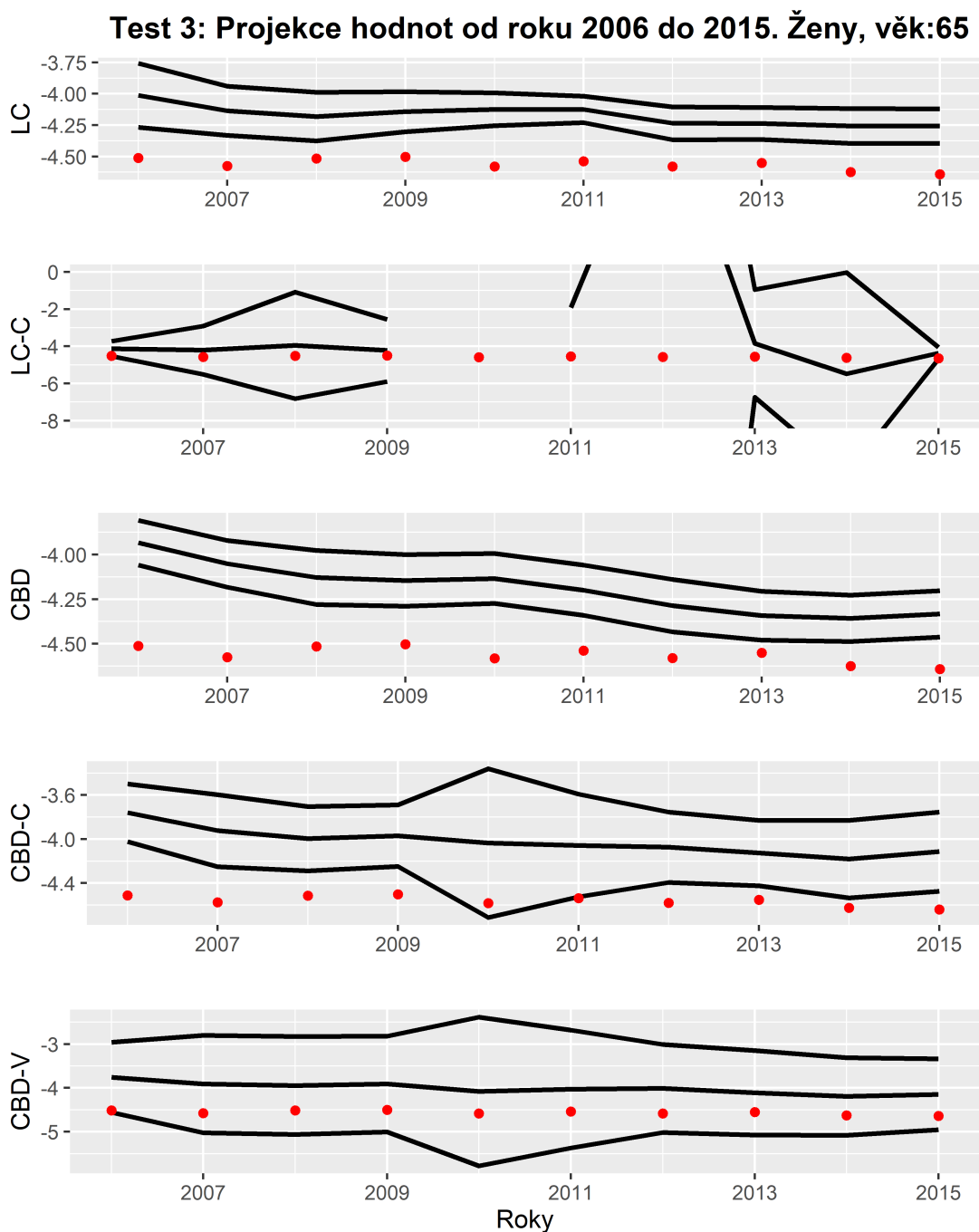


Obrázek 6.19: Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.

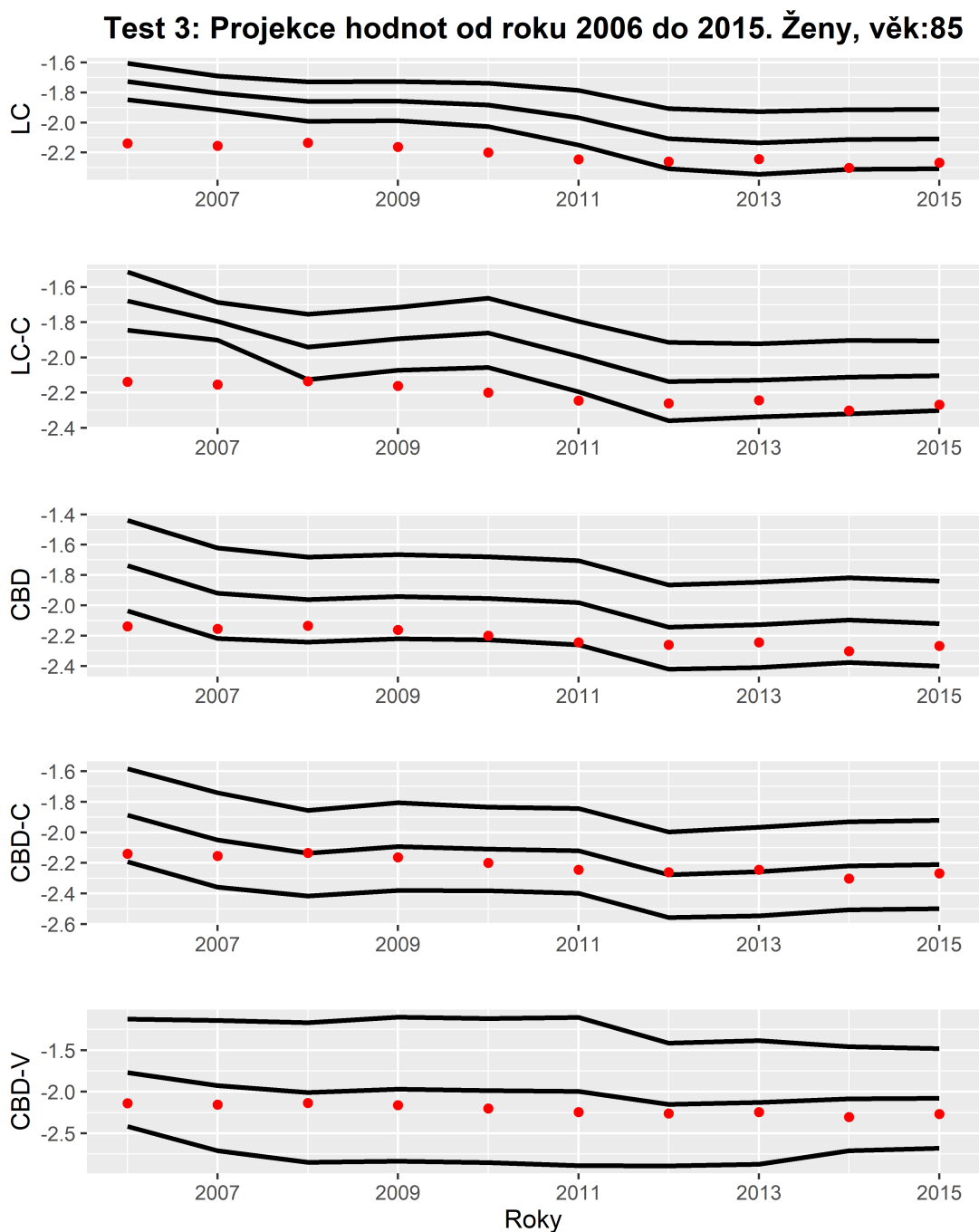
Test 3: Projekce hodnot od roku 2006 do 2015. Muži, věk:85



Obrázek 6.20: Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.



Obrázek 6.21: Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1967, ..., 1976 do roků 1986, ..., 1995, které jsou predikované do roků 2006, ..., 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.



Obrázek 6.22: Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.

6.2.4 Shrnutí

Z výše uvedené veskrze grafické evaluace modelů v jednotlivých testech můžeme provést celkové hodnocení. Nejlepších výsledků dosáhl model (CBD-V), který

uspěl u všech testů. Poznamenejme, že šířka predikčních ramen intervalu, která pomohla ke splnění některých testů je pro reálné aplikace problematická, neboť čím širší predikční interval zvolíme, tím je aplikace rizikovější. Rozdíl v ocenění mezi hodnotou horního a spodního predikčního intervalu může být signifikantní. Modely (LC) a (CBD-C) splnily všechny testy pro mužskou populaci, avšak oba neprošly v jednom testu pro ženskou populaci. Můžeme je tedy jednoznačně doporučit pro modelování mužské úmrtnosti a s drobnými výhradami i pro ženskou populaci. (CBD) model měl problémy s nepřesnou predikcí, a tím pádem ho nemůžeme doporučit jako vhodný. Nicméně poznamenejme, že byl numericky stabilní. To se ovšem nedá tvrdit o modelu (LC-C), který uspěl pouze v mužské predikci prvního testu. Obecně měl problém se stabilitou, konvergencí i časovou náročností odhadů. Z těchto důvodů ho považujeme za krajně nevhodný pro modelování a predikci úmrtností.

6.3 Ocenění longevity dluhopisu

V této sekci budeme počítat spravedlivou cenu longevity dluhopisu, tak jak jsme si ji definovali v teoretické kapitole 5. *Zajištění rizika dlouhověkosti*. Postup bude následovný: pro portfolio anuitantů l_{x_0} věku x_0 určíme tržní cenu annuity k roku t_0 . Poté pomocí úmrtnostních modelů odhadneme vývoj úmrtností po roce t_0 . Následně odhadneme tržní cenu rizika λ , kterou použijeme pro Wangovu transformaci. Máme-li odhadnutý parametr λ , můžeme ho použít na výpočet jednotlivých kuponů a spravedlivé ceny longevity dluhopisu. Výsledkem této analýzy by měl být přehled tržních cen rizika a cen longevity dluhopisů pro obě pohlaví a všech pět úmrtnostních modelů.

Nejdříve je třeba si zvolit několik parametrů. Referenční rok zvolíme jako $t_0 = 2019$. Tato volba je motivována tím, že v první sekci empirické kapitoly máme již všechny úmrtnostní modely odhadnuty k tomuto roku. Věk, který nás bude zajímat, jest $x_0 = 65$, ten používáme v této práci jako referenční rok odchodu do starobního důchodu v České republice, viz diskuze na začátku podkapitoly 6.2 *Empirická diagnostika modelu*. Pro tento rok máme jednotlivé modely otestované. Nicméně analýzu provedeme pro všechny uvedené modely, tedy i pro ty, které nesplnily všechny testy v předchozí sekci. Předpokládáme, že velikost našeho anuitního kmene je $l_{x_0} = 10000$.

Výnosovou křivku, budeme pro jednoduchost předpokládat plochou, tedy, že výnos v jednotlivých letech je konstantní a to konkrétně $i = 0.03$ p.a.. Diskontní faktor určíme pomocí spojitého úročení:

$$d(0,t) = e^{-\int_0^t i(v)dv} = e^{-\int_0^t i dv} = e^{-0.03t}.$$

Takto definovaný diskontní faktor použijeme jak při stanovení ceny annuity, tak pro diskontování kuponů longevity dluhopisu. Maturitu dluhopisu volíme jako $T = 35$ let. Tedy portfolio anuitantů bude v době maturity dluhopisu ve věku 100.

V článku Lin a Cox (2005) použili autoři jako odhad tržní cenu anuity $a_x^{market}(t_0)$ kvótovanou cenu z sekundárního U.S. anuitního trhu. Tento to trh v České republice neexistuje, proto je třeba vhodně aproximovat tuto cenu. Náš postup bude takový, že k určení $a_x^{market}(t_0)$ použijeme České úmrtnostní tabulky založené na datech získaných z HMD (2021) k roku 2019. Takto spočtené úmrtnostní pravděpodobnosti musíme přeskálovat selekčními faktory, které škálují populační úmrtnosti na úmrtnosti pojistných kmenů. Tato úprava nám zajistí mitigaci takzvaného selekčního rizika a výslednou cenu anuity můžeme považovat za proxy k tržní ceně anuity. Poznamenejme, že se autor pokoušel získat tyto selekční faktory českých pojišťoven jak u České společnosti aktuárů, tak i u největších životních pojišťoven v České republice. Bohužel Česká společnost aktuárů taková data neneviduje a soukromé pojišťovny nebyly vstřícné ke sdílení těchto citlivých dat. Z těchto důvodů budeme v ocenění používat selekční faktory z anglického pojistného trhu, které nalezneme ve článku CMI (2005) od anglického institutu aktuárů (*The Institute and Faculty of Actuaries*). Tato práce analyzuje selekční faktory pro mužskou a ženskou populaci zvláště přes jednotlivé druhy pojištění. My použijeme selekční faktory z celkového kmene uvažovaných pojistných smluv. Tento postup je inspirován článkem Levantesi a kol. (2009), který zvolil obdobný přístup pro italskou populaci. Selekční faktory můžeme nalézt v tabulce 6.10. Transformované pravděpodobnosti úmrtí získáme jako:

$${}_t\hat{q}_x = {}_tq_x^{(LT)} / SF_{i,j},$$

kde ${}_tq_x^{(LT)}$ je populační pravděpodobnost úmrtí obdržena z HMD (2021) a $SF_{i,j}$ je selekční faktor pro věkovou skupinu i a pohlaví j .

Věková skupina	Muži	Ženy
65-69	1,26	1,34
70-74	1,32	1,35
75-79	1,25	1,27
80-84	1,15	1,17
85-89	1,10	1,11
90-94	1,14	1,17
95-99	1,12	1,17
100-104	0,85	1,02

Tabulka 6.10: Selekční faktory pro ženy a muže v jednotlivých věkových kategoriích pro agregované pojistné smlouvy. List č.2 (muži), list č.6 (ženy), průměry za roky 2000-2003 získané z CMI (2005).

Máme-li transformované pravděpodobnosti úmrtí ${}_t\hat{q}_x$, pak je můžeme použít ke stanovení tržní ceny anuity pro muže a ženy dle vztahu:

$$a_{65}^{market}(2019) = \sum_{t=1}^{35} [1 - {}_t\hat{q}_{65}] d(0,t).$$

Odhadnuté hodnoty tržní ceny anuity s jednotkovou roční výplatou jsou 12.51 a 14.61 pro muže, respektive ženy. Obě tyto ceny použijeme pro stanovení hodnoty

λ_{65} pomocí následující rovnice:

$$a_{65}^{market}(2019) = \sum_{t=1}^{35} [1 - \phi(\phi^{-1}(tq_{65}) - \lambda_{65}(2019))]d(0,t),$$

kde $\phi(x)$ reprezentuje distribuční funkci standardního normálního rozdělení a tq_{65} jsou pravděpodobnosti úmrtí získané z predikce úmrtnostního modelu. Provedeme pro všech pět modelů i obě pohlaví a dostáváme hodnoty tržního rizika v tabulce 6.11. V grafech 6.23, 6.24 a 6.25–6.26 dodáváme porovnání Wangových transformací úmrtností dle (5.8) a škálovaných úmrtností $t\hat{q}_{65}$ vybraných věků pro muže, respektive ženy. Z tabulky si můžeme všimnout znatelného vychýlení tržního rizika ženské populace pro model (CBD-V) to i odpovídá grafům 6.25 a 6.26, kde pozorujeme opačný trend oproti zbytku modelů. To jest, že pro věky 74–76 je transformované úmrtí pod tq_{65} a pro věky 89–91 je naopak nad. Ostatní modely mají tento jev opačný. Dané pozitivní vychýlení bude mít za následek vyšší cenu P, při stanovení férové ceny dluhopisu. Poznamenejme, že tržní riziko je přes jednotlivé modely poměrně stabilní pro mužskou populaci, u žen je patrný rostoucí trend λ_{65} .

model	LC	LC-C	CBD	CBD-C	CBD-V
Muži	0.0699	0.1224	0.0908	0.0937	0.1070
Ženy	0.0203	0.0349	0.0476	0.1453	0.3628

Tabulka 6.11: Hodnoty tržního rizika věku 65 získané z Wangovy transformace pro obě pohlaví.

Nyní již máme vše potřebné pro stanovení spravedlivé ceny longevity dluhopisu pro naše portfolio 65–letých anuitantů, $l_{65} = 10000$. Dodejme, že kupon C_t volíme fixní pro všechny časové okamžiky, a to jako hodnotu $C_t = C = i * F = 0.03 * 10000 = 300$, kde $F = 10000$ je nominální hodnota státního dluhopisu. Dodejme, že cena dluhopisu je taktéž rovna $W = 10000$, dluhopis je tedy upsán *at par*. Každý anuitant dostává jednotkovou výplatu za rok. Upisovatel anuity je tedy plně zajištěn proti případnému neočekávanému nárůstu přeživších anuitantů v každém roce až do výše 300 (včetně), dojde-li v daném roce k překročení této hranice, pak je upisovatelova ztráta ponížena o 300, ale zůstává již kladná.

Výsledky nalezneme v tabulkách 6.12 a 6.13 pro muže, respektive ženy. Poznamenejme, že z konstrukce pravděpodobností úmrtí a Wangovy transformace daná cena P reflektuje jak riziko dlouhověkosti, tak selekční riziko. U mužů vidíme, že cena P se pohybuje kolem hodnoty 3640, vyjma modelu (CBD), kde je pouze 3578 naopak největší hodnotu odhadl (LC-C) model 3702. Dodejme, že právě modely (LC-C) a (CBD) dosáhly nejhorších diagnostických výsledků pro muže a tím pádem je i patrné odchýlení ceny P od ostatních modelů. U žen jsou výsledky ceny P významně odlišné. Nejmenší hodnoty bylo dosaženo pro (LC) model 929, nejvyšší pro (CBD-V) 3781. Poznamenejme, že modely (CBD-C) a (CBD-V), které dosahovaly dobrých diagnostických predikčních výsledků pro ženskou populaci, odhadují cenu P relativně konzistentně. Tím pádem považujeme jejich odhady jako vhodné pro určení ceny P u ženské populace. Pozorujeme, že rostoucí tržní

hodnota rizika je konzistentní s rostoucí hodnotou P . Tento fakt je intuitivní, za předpokladu, že selekční riziko, které je také reflektované v P , je v každém úmrtnostním modelu zohledněnou stejnou měrou.

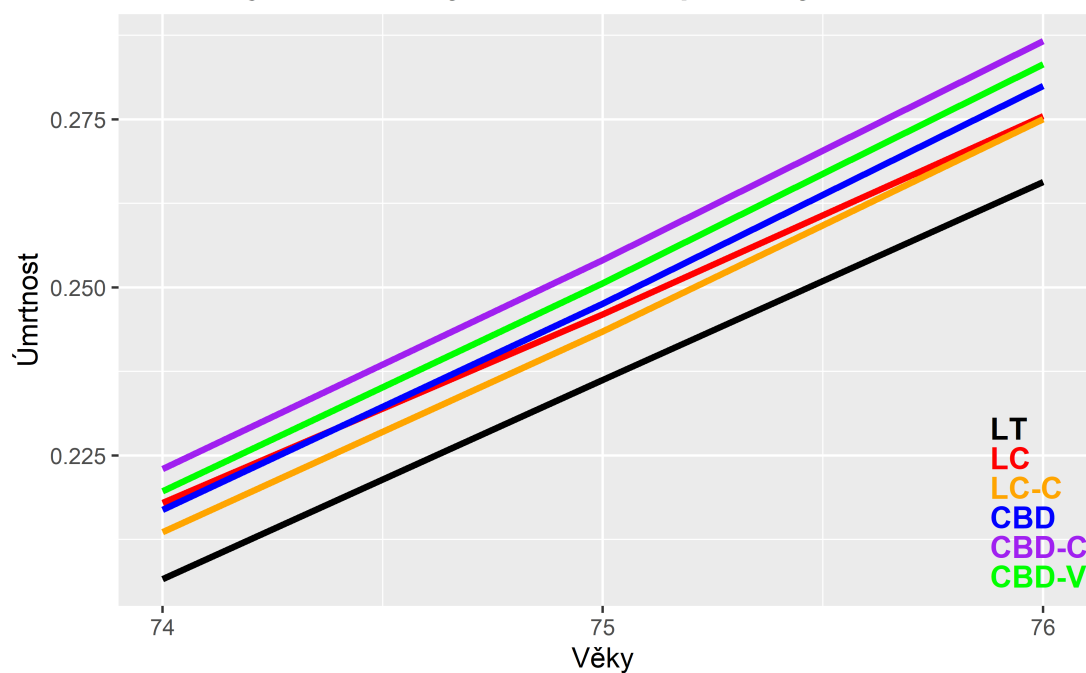
model	LC	LC-C	CBD	CBD-C	CBD-V
λ	0.0699	0.1224	0.0908	0.0937	0.1070
P	3608	3702	3578	3681	3630
V	6392	6298	6422	6319	6370
W	10000	10000	10000	10000	10000

Tabulka 6.12: Výsledky ocenění longevity dluhopisu P, V pro mužskou populaci věku 65, s portfoliem velikosti 10000 k roku 2019 a úrokovou mírou 3 % *p.a.* společně s tržní hodnotou rizika přes jednotlivé modely.

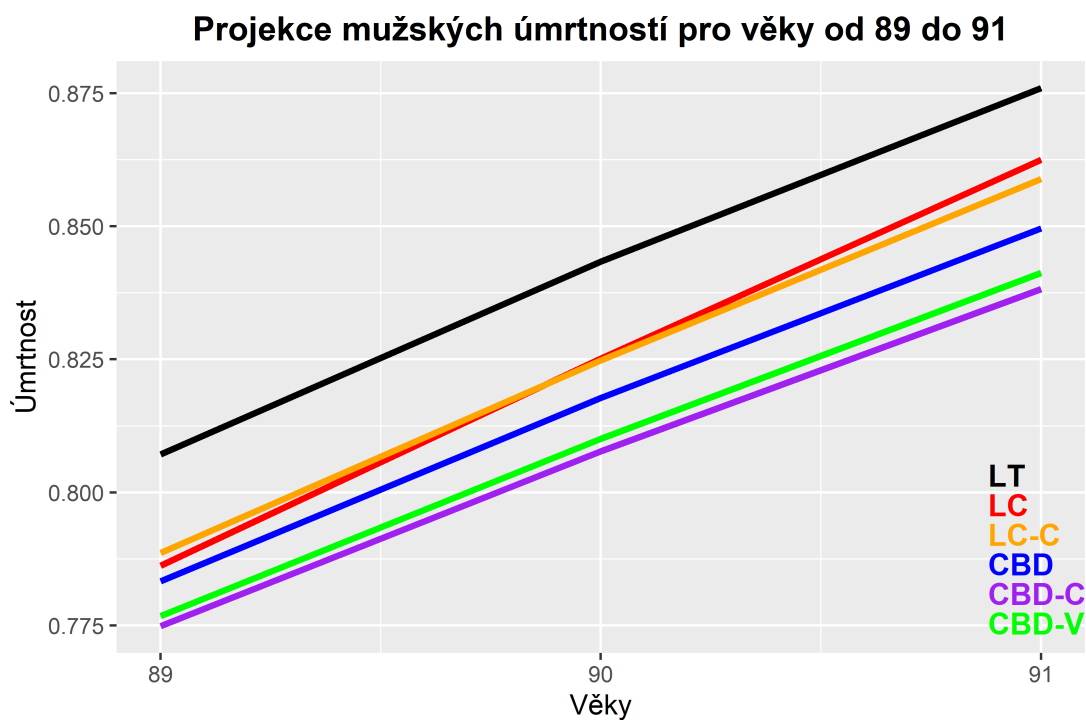
model	LC	LC-C	CBD	CBD-C	CBD-V
λ	0.0203	0.0349	0.0476	0.1453	0.3628
P	929	1599	2274	3453	3781
V	9071	8401	7726	6547	6219
W	10000	10000	10000	10000	10000

Tabulka 6.13: Výsledky ocenění longevity dluhopisu P, V pro ženskou populaci věku 65, s portfoliem velikosti 10000 k roku 2019 a úrokovou mírou 3 % *p.a.* společně s tržní hodnotou rizika přes jednotlivé modely.

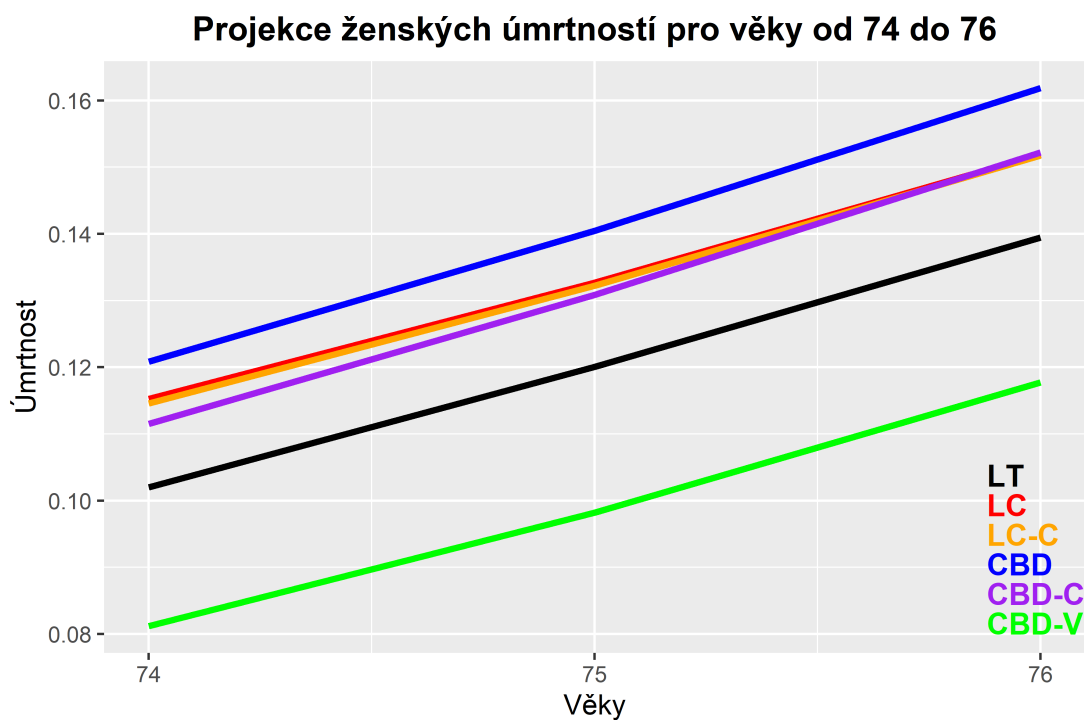
Projekce mužských úmrtností pro věky od 74 do 76



Obrázek 6.23: Predikce mužských Wangových transformací úmrtnostní ${}_tq_{65}^*$ jednotlivých modelů pro roky 74–76 společně s rozšířenými pravděpodobnostmi ${}_t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti mužů pro věky 50-101 od roku 1950 do 2019.

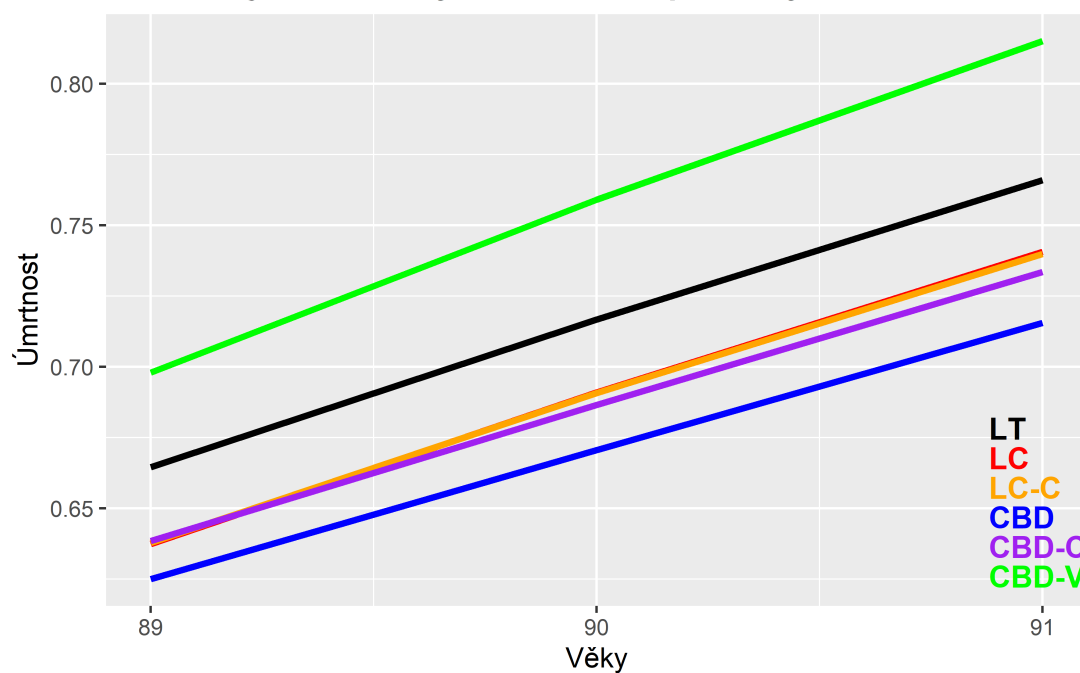


Obrázek 6.24: Predikce mužských Wangových transformací úmrtnostní tq_{65}^* jednotlivých modelů pro roky 89–91 společně s rozšířenými pravděpodobnostmi $t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti mužů pro věky 50-101 od roku 1950 do 2019.



Obrázek 6.25: Predikce ženských Wangových transformací úmrtnostní ${}_tq_{65}^*$ jednotlivých modelů pro roky 74 – 76 společně s rozšířenými pravděpodobnostmi ${}_t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti žen pro věky 50-101 od roku 1950 do 2019.

Projekce ženských úmrtností pro věky od 89 do 91



Obrázek 6.26: Predikce ženských Wangových transformací úmrtnostní ${}_tq_{65}^*$ jednotlivých modelů pro roky 89 – 91 společně s rozšířenými pravděpodobnostmi ${}_t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti žen pro věky 50-101 od roku 1950 do 2019.

Závěr

Tato práce se zabývala úmrtnostními stochastickými modely v kontextu kvantifikace vybraných pojistněmatematických rizik. Nejprve jsme uvedli stručný historický vývoj úmrtnosti a motivaci k zkoumání tohoto tématu. Následoval krátký přehled definic a tvrzení, o které se práce opírala v jednotlivých sekcích.

Ve třetí kapitole jsme si ujasnili základní myšlenky ztvárňující naši představu o úmrtnosti. Poté jsme zavedli pět stochastických úmrtnostních modelů: Lee-Carterův model a jeho rozšíření o kohortní člen, Cairns-Blake-Dowdův model s jeho dvěma rozšířeními o kohortní a variační člen. Dále jsme se zaměřili na přístupy k modelování predikcí. Tato sekce byla zakončena definicemi diagnostických testů, které nám posloužily jako hodnotící kritéria v empirické části práce.

Čtvrtá kapitola se zabývala pojistněmatematickými riziky v kontextu úmrtnosti. Nejdříve jsme si představili rizika vycházející z takzvaného standardního vzorce, který slouží pro výpočet solventnostního rizikového kapitálu pojišťoven. Poté jsme dodali rizika spojená s odhadem modelů.

Pátá kapitola se věnovala zajištění rizika dlouhověkosti, které považujeme za nejvíce materiální pro životní pojišťovny, penzijní společnosti a penzijní státní fondy. Sekci jsme zahájili s definicí longevity dluhopisu a vysvětlením jeho principu a peněžních toků. Poté jsme se zabývali přístupem ke stanovení spravedlivé ceny tohoto finančního kontraktu. K tomu jsme použili Wangovu transformaci, která se pojí s takzvaným parametrem ceny tržního rizika.

Poslední šestá kapitola byla empirickou částí práce. Aplikovali jsme zde vybrané teoretické sekce práce na českých datech úmrtnosti. Začali jsme s odhadem všech stochastických úmrtnostních modelů, kde jsme popisovali jednotlivé odhady parametrů a proložení skutečných realizací pro mužskou populaci. Věnovali jsme se také základní diagnostice residuí. Odhady všech modelů věrohodně prokládaly skutečně realizované úmrtnosti věku 65 a splňovaly naši teoretickou představu o chování jednotlivých parametrů. Odhad (LC-C) modelu byl poměrně náročný, neboť nelineární vztah odhadovaných parametrů na velké datové sadě ne vždy konvergoval k řešení. Několik iterací s náhodnými startovacími body bylo třeba použít jak pro muže, tak i pro ženy. Analýza residuální složka vykazovala známky autokorelace pro všechny modely, nejvíc patrná byla u mladší roků. Taktéž klesající rozptyl byl znatelný u všech modelů.

Ve druhé části empirické kapitoly jsme provedli predikční testy na jednotlivé modely. Vždy pro obě pohlaví a pro věky 65 a 85. Pro každý test a pohlaví jsme uvedli tabulku úspěšnosti daných modelů. Z dosažených výsledků jsme došli k závěru, že nejlepší model jest (CBD-V), který uspěl u všech testů. Modely (LC) a (CBD-C) splnily všechny testy pro mužskou populaci, avšak oba neprošly v jednom testu pro ženskou populaci. Můžeme je tedy jednoznačně doporučit pro modelování mužské úmrtnosti a s drobnými výhradami i pro ženskou populaci.

(CBD) model měl problémy s nepřesnou predikcí, a tím pádem ho nemůžeme doporučit jako vhodný. Nicméně poznamenejme, že byl numericky stabilní. To se ovšem nedá tvrdit o modelu (LC-C), který uspěl pouze v mužské predikci prvního testu. Obecně měl problém se stabilitou, konvergencí i časovou náročností odhadů. Z těchto důvodů nepovažujeme (LC-C) za vhodný pro modelování a predikování úmrtností.

V posledním empirickém úseku jsme ocenili longevity dluhopis pomocí Wangovy transformace a tržní ceny rizika. Dodali jsme rozpady spravedlivé ceny kontraktu pro pojišťovnu a investora společně s odhadnutou cenou tržního rizika. U výsledku pro mužskou populaci pozorujeme podobné hodnoty ceny P pro všechny modely. Toto tvrzení ovšem neplatí pro ženskou populaci, kde výsledky ceny P jsou významně odlišné. Nejmenší hodnoty bylo dosaženo pro (LC) model 929, nejvyšší pro (CBD-V) 3781. Tento rozptyl je výsledkem nevhodnosti některých modelů, kterou jsme detekovali v podkapitole *6.2 Empirická diagnostika modelů*, pro ženskou populaci. Uzavíráme s faktem, že rostoucí tržní hodnota rizika je konzistentní s rostoucí hodnotou P, což je v souladu s naší teoretickou představou.

Seznam použité literatury

- BLAKE, D., CAIRNS, A. J. G. a DOWD, K. (2006). *A two-factor model for stochastic mortality with parametr uncertainty*. The Journal of Risk and Insurance, 73, 687-718.
- BROUHNS, N., DENUIT, M. a VERMUNT, J. K. (2002). *A Poisson log-bilinear approach to the construction of projected lifetables*. Insurance: Mathematics and Economics, 31(3), 373-393.
- BROWNE, M. J. a KIM, K. (1993). *An International Analysis of Life Insurance Demand*. Journal of Risk and Insurance, 60, 4, 676-634.
- CIPRA, T. (2006). *Pojistná matematika: teorie a praxe*. Ekopress, Praha. ISBN 80-86929-11-6.
- CMI (2002). *Working Paper 1: An interim basis for adjusting the '92 series' mortality projections for cohort effects*. Institute and Faculty of Actuaries.
- CMI (2005). *Working Paper 17: Report on the preliminary results of an analysis into the mortality experience of pensioners of self-administered pension schemes for the period 2000 to 2003*. Institute and Faculty of Actuaries.
- CSU (2021). *Naděže dožití ve vybraných věcích v České republice v letech 1920-2020*. Český statistický úřad, Dostupné na https://www.czso.cz/csu/czso/obyut_cr.
- DOWD, K., CAIRNS, A. J. G., BLAKE, D., COUGHLAN, G. D., EPSTEIN, D. a KHALAF-ALLAH, M. (2010). *Backtesting stochastic mortality models: An ex post evaluation of multiperiod-ahead density forecasts*. North American Actuarial Journal, 14, 281-298.
- EU (2009). *Směrnice Evropského parlamentu a Rady 2009/138/ES*. Evropský parlament a rada evropské unie.
- HMD (2021). *Human mortality database*. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), Available at www.mortality.org or www.humanmortality.de.
- KULICH, M. (2021). *Extended Course Notes: NMST432 Advanced Regression Model*. Department of Probability and Mathematical Statistics Faculty of Mathematics and Physics, Charles University.
- LEE, R. D. a CARTER, L. (1992). *Modeling and forecasting US mortality*. Journal of the American Statistical Association, 87, 659-671.
- LEVANTESI, S., MENZIETTI, M. a TORRI, T. (2009). *Longevity bond pricing models: an application to the Italian annuity market and pension schemes*. 72. Giornale dell'Istituto Italiano degli Attuari, 125-147.
- LIN, Y. a COX, S. H. (2005). *Securitization of mortality risks in life annuities*. The Journal of Risk and Insurance, 72, 227-252.

- MACDONALD, A., RICHARDS, S. a CURRIE, I. (2018). *Modelling Mortality with Actuarial Applications*. Cambridge: Cambridge University Press.
- NELDER, J. a WEDDERBURN, R. (1972). *Generalized linear models*. Journal of the Royal Statistical Society, A135(3), 370-384.
- POLLARD, J. H. (1972). *Mathematical Models for the Growth of Human Populations*. Journal of the Institute of Actuaries, 99(3), 307-308.
- PRÁŠKOVÁ, Z. a LACHOUT, P. (1998). *Základy náhodných procesů*. 1.vydání. Praha : Karolinum. ISBN 80-7184-688-0.
- SII (2009). *Zákon č. 277/2009 Sb. Zákon o pojišťovnictví*. Sbírka zákonů 31.08.2009.

Seznam obrázků

5.1	Schéma peněžních toků longevity dluhopisu.	23
6.1	Odhad (LC) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu α_x , β_x a κ_y . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	28
6.2	Residua (LC) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.	29
6.3	Odhad (LC-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu α_x , κ_y , β_x , γ_c a $\beta_x^{(2)}$. V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	30
6.4	Residua (LC-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.	31
6.5	Odhad (CBD) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají odhadu $\kappa_y^{(1)}$ a $\kappa_y^{(2)}$. V posledním grafu je černou křivkou vyobrazen odhad (CBD) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	32
6.6	Residua (CBD) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.	33
6.7	Odhad (CBD-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}$, $\kappa_y^{(2)}$ a γ_c . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (CBD-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	34
6.8	Residua (CBD-C) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.	35
6.9	Odhad (CBD-V) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}$, $\kappa_y^{(2)}$, $\kappa_y^{(3)}$ a γ_c . V posledním grafu je černou křivkou vyobrazen odhad (CBD-V) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	36
6.10	Residua (CBD-V) modelu na českých datech úmrtnosti mužů pro věky 50-95 od roku 1950 do 2019.	37
6.11	Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1981, . . . , 1995 do roků 2000, . . . , 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, . . . , 1-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.	40

6.12	Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1981, . . . , 1995 do roků 2000, . . . , 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, . . . , 1-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.	41
6.13	Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1981, . . . , 1995 do roků 2000, . . . , 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, . . . , 1-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.	42
6.14	Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1981, . . . , 1995 do roků 2000, . . . , 2014, které jsou predikované do roků 2015. Černé křivky představují 15-roční, 14-roční, . . . , 1-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují realizovanou hodnotu úmrtnosti v roce 2015.	43
6.15	Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	46
6.16	Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	47
6.17	Predikce založena na českých datech úmrtnosti žen pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	48
6.18	Predikce založena na českých datech úmrtnosti žen pro věky 50-100 v letech 1981-2000. Černé křivky představují x-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	49
6.19	Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	52
6.20	Predikce založena na českých datech úmrtnosti mužů pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	53

6.21	Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 65. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	54
6.22	Predikce založena na českých datech úmrtnosti žen pro věky 50-100 od roků 1967, . . . , 1976 do roků 1986, . . . , 1995, které jsou predikované do roků 2006, . . . , 2015. Černé křivky představují 20-roční centrální predikci a příslušný 90% predikční interval pro věk 85. Červené body reprezentují skutečně realizované hodnoty v jednotlivých letech.	55
6.23	Predikce mužských Wangových transformací úmrtnostní tq_{65}^* jednotlivých modelů pro roky 74 – 76 společně s rozšířenými pravděpodobnostmi $t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti mužů pro věky 50-101 od roku 1950 do 2019.	60
6.24	Predikce mužských Wangových transformací úmrtnostní tq_{65}^* jednotlivých modelů pro roky 89 – 91 společně s rozšířenými pravděpodobnostmi $t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti mužů pro věky 50-101 od roku 1950 do 2019.	61
6.25	Predikce ženských Wangových transformací úmrtnostní tq_{65}^* jednotlivých modelů pro roky 74 – 76 společně s rozšířenými pravděpodobnostmi $t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti žen pro věky 50-101 od roku 1950 do 2019.	62
6.26	Predikce ženských Wangových transformací úmrtnostní tq_{65}^* jednotlivých modelů pro roky 89 – 91 společně s rozšířenými pravděpodobnostmi $t\hat{q}_{65}$ založených na úmrtnostních tabulkách z roku 2019 škálované selekčními faktory (černá). Odhady modelu konstruovaný z českých dat úmrtnosti žen pro věky 50-101 od roku 1950 do 2019.	63
A.1	Odhad (LC) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu α_x , β_x a κ_y . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	73
A.2	Residua (LC) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019	74
A.4	Residua (LC-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.	74

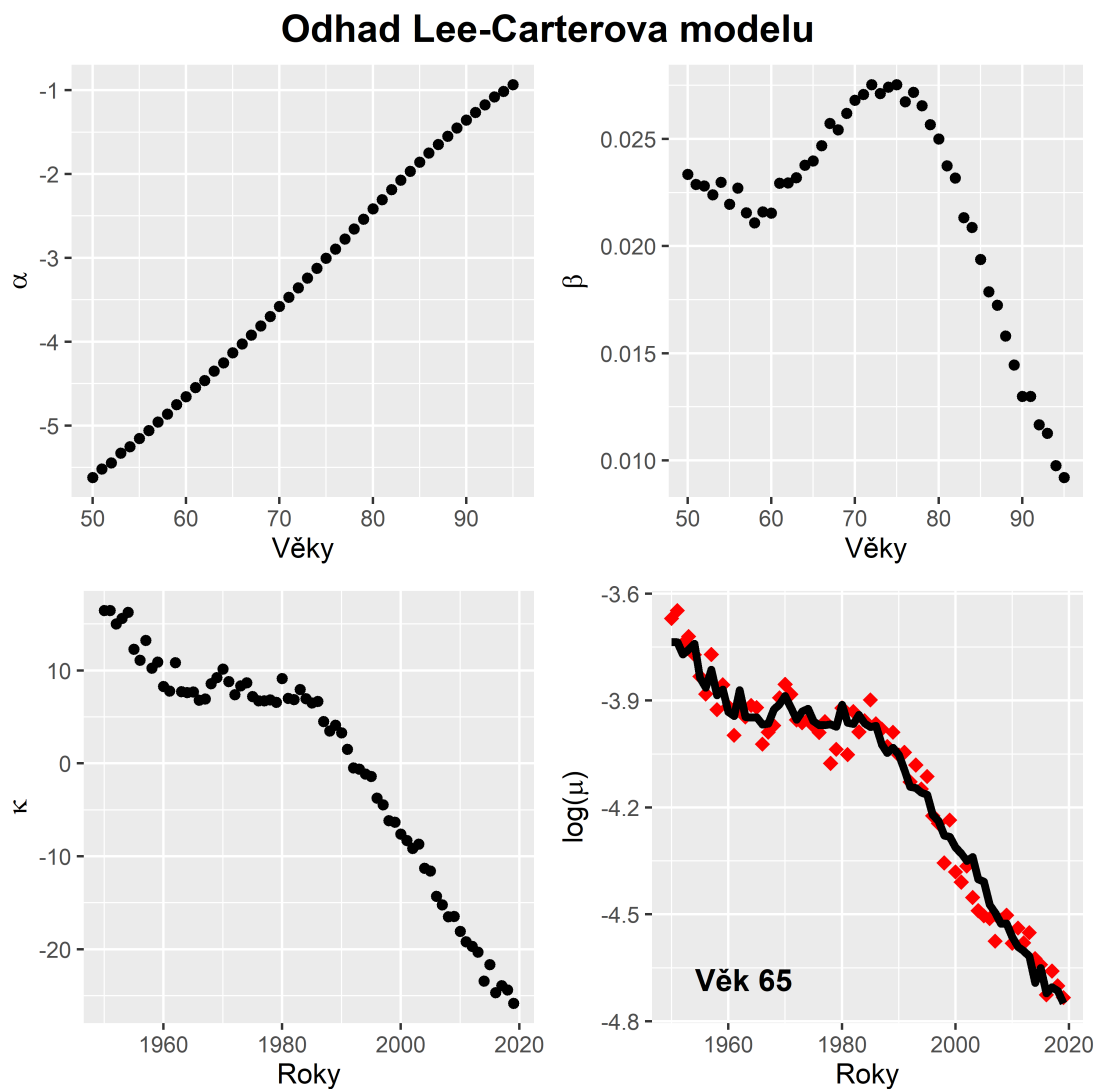
A.3	Odhad (LC-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\alpha_x, \kappa_y, \beta_x, \gamma_c$ a $\beta_x^{(2)}$. V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	75
A.5	Odhad (CBD) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají odhadu $\kappa_y^{(1)}$ a $\kappa_y^{(2)}$. V posledním grafu je černou křivkou vyobrazen odhad (CBD) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	76
A.6	Residua (CBD) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.	77
A.7	Odhad (CBD-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}, \kappa_y^{(2)}$ a γ_c . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (CBD-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.	78
A.8	Residua (CBD-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.	79
A.9	Residua (CBD-V) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.	79
A.10	Odhad (CBD-V) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}, \kappa_y^{(2)}, \kappa_y^{(3)}$ a γ_c . V posledním grafu je černou křivkou vyobrazen odhad (CBD-V) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti. . .	80

Seznam tabulek

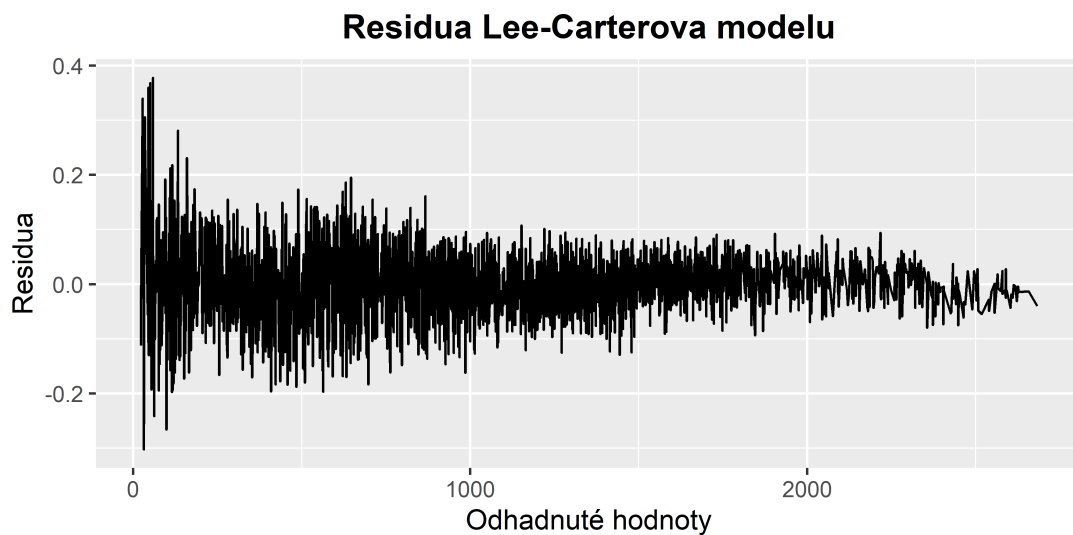
6.1	Statistiky průměrných věků úmrtí a expozic pro ženy a muže z českých dat od roku 1950 do 2019 po desetiletých skupinách.	26
6.2	Vývoj střední doby dožiti novorozenců v České republice pro obě pohlaví v uvažovaném horizontu od roku 1993 do 2020.	38
6.3	Výsledky testu konvergence zužujícího horizontu.	39
6.4	Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test dlouhé predikce mužů, oba věky.	44
6.5	Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test dlouhé predikce žen, oba věky.	44
6.6	Výsledky testu dlouhé predikce.	45
6.7	Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test posouvajícího se okna predikce mužů, oba věky.	50
6.8	Hodnoty čtveřice $[x_U, x_L, x_{C_U}, x_{C_L}]$ pro test posouvajícího se okna predikce žen, oba věky.	50
6.9	Výsledky testu posouvajícího se okna predikce.	51
6.10	Selekční faktory pro ženy a muže v jednotlivých věkových kategoriích pro agregované pojistné smlouvy. List č.2 (muži), list č.6 (ženy), průměry za roky 2000-2003 získané z CMI (2005).	57
6.11	Hodnoty tržního rizika věku 65 získané z Wangovy transformace pro obě pohlaví.	58
6.12	Výsledky ocenění longevity dluhopisu P, V pro mužskou populaci věku 65, s portfoliem velikosti 10000 k roku 2019 a úrokovou mírou 3 % <i>p.a.</i> společně s tržní hodnotou rizika přes jednotlivé modely. .	59
6.13	Výsledky ocenění longevity dluhopisu P, V pro ženskou populaci věku 65, s portfoliem velikosti 10000 k roku 2019 a úrokovou mírou 3 % <i>p.a.</i> společně s tržní hodnotou rizika přes jednotlivé modely. .	59

A. Přílohy

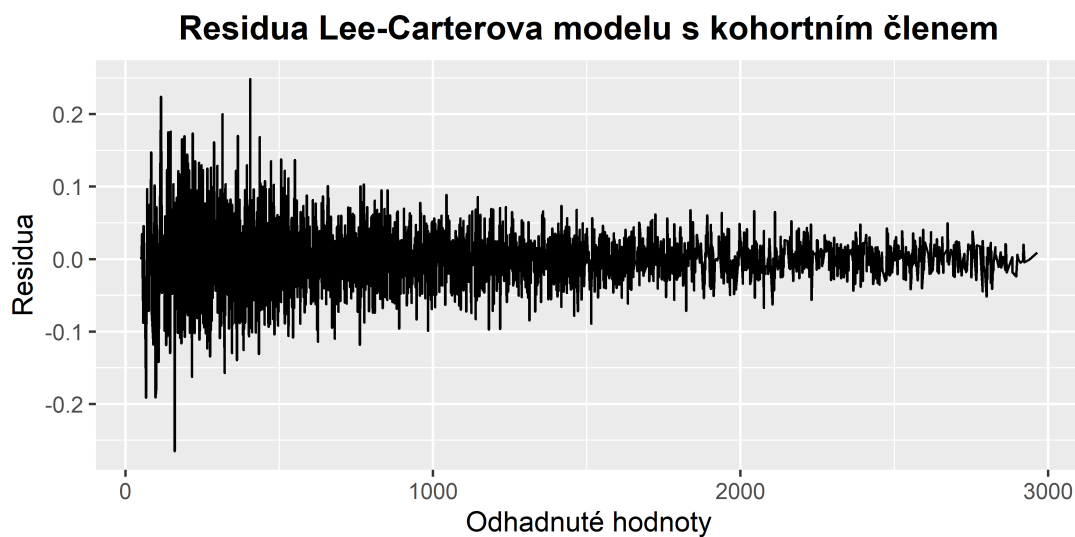
A.1 Odhady modelů a residuí pro ženskou populaci



Obrázek A.1: Odhad (LC) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu α_x , β_x a κ_y . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

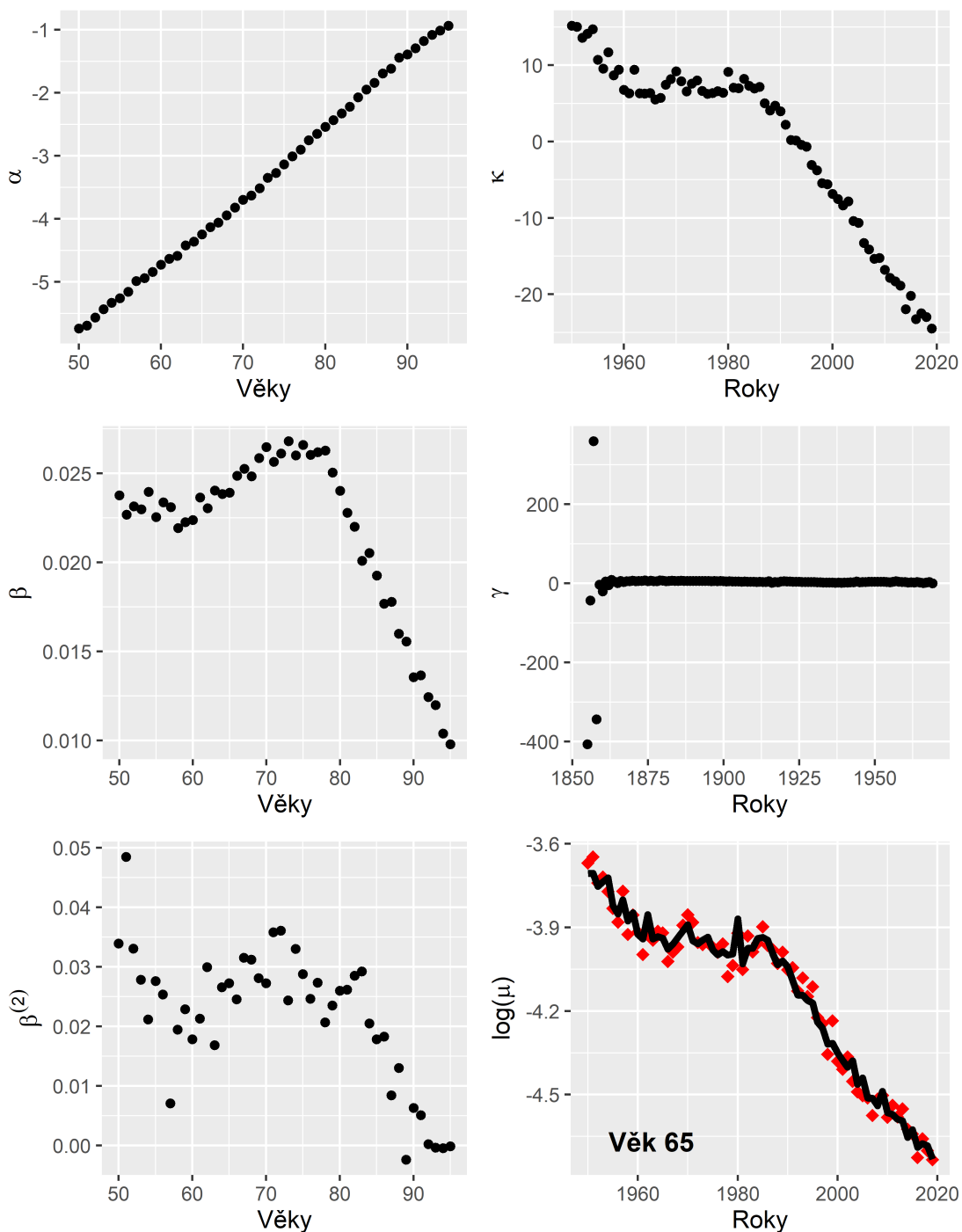


Obrázek A.2: Residua (LC) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019



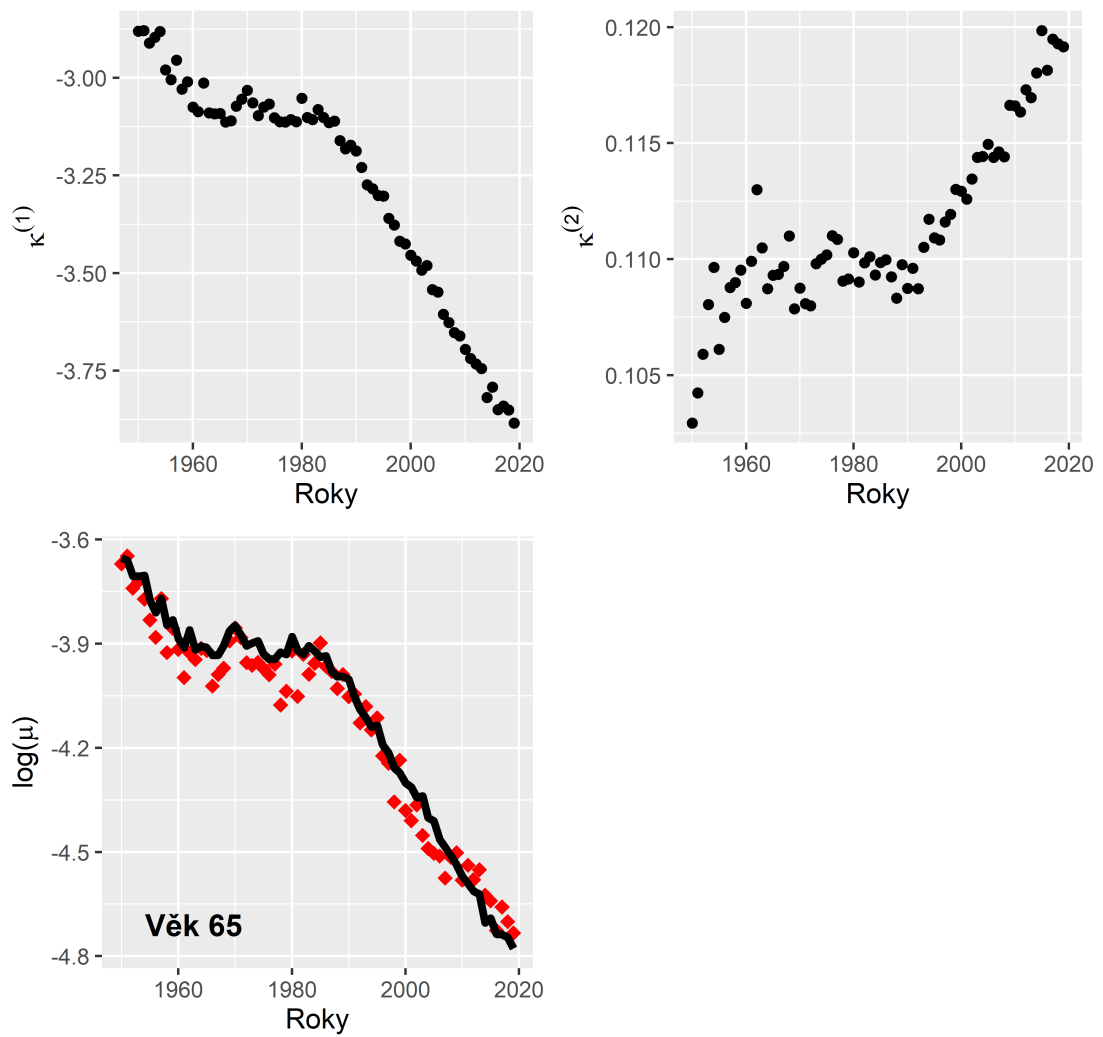
Obrázek A.4: Residua (LC-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.

Odhad Lee-Carterova modelu s kohortním členem

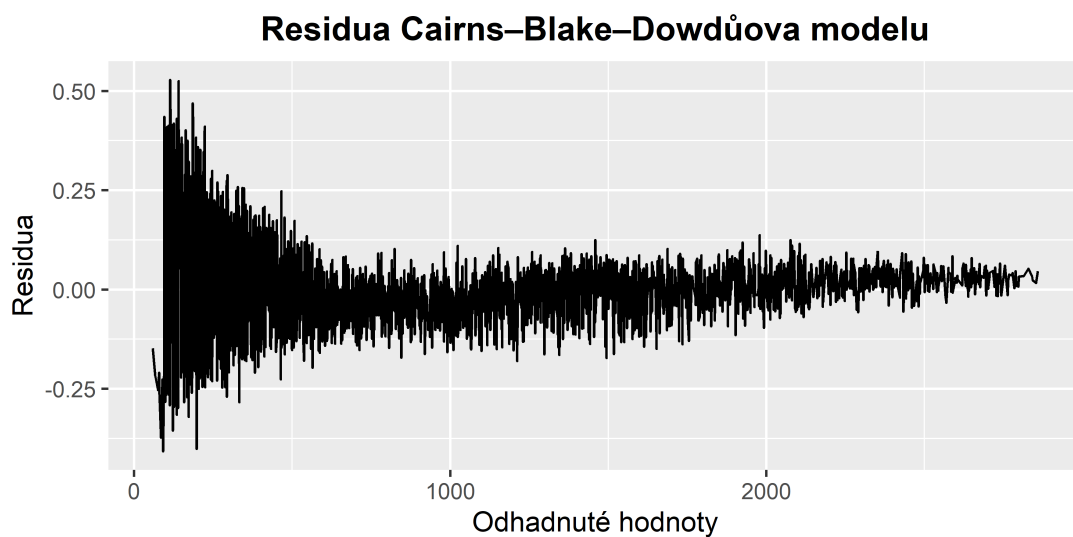


Obrázek A.3: Odhad (LC-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\alpha_x, \kappa_y, \beta_x, \gamma_c$ a $\beta_x^{(2)}$. V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (LC-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

Odhad Cairns–Blake–Dowdůva modelu

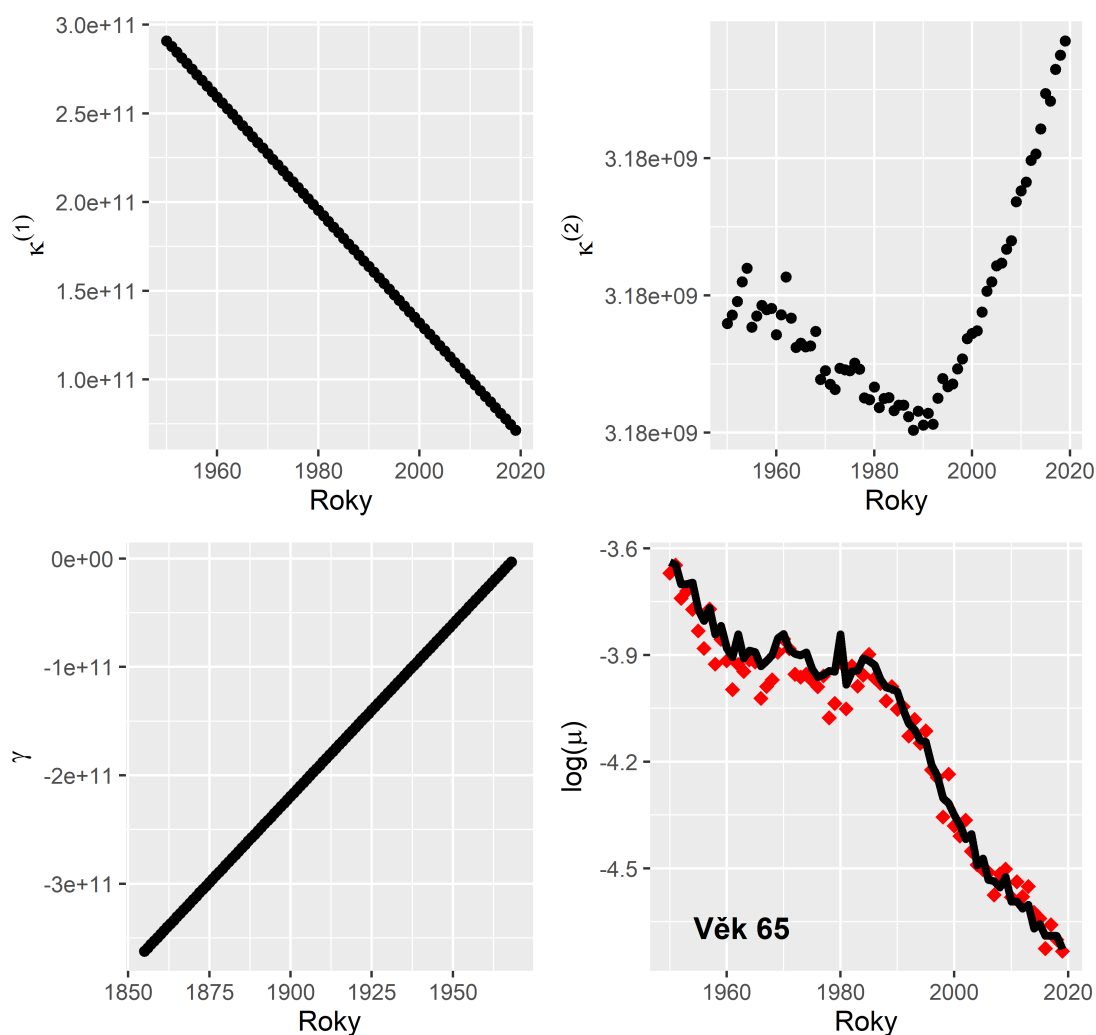


Obrázek A.5: Odhad (CBD) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají odhadu $\kappa_y^{(1)}$ a $\kappa_y^{(2)}$. V posledním grafu je černou křivkou vyobrazen odhad (CBD) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.



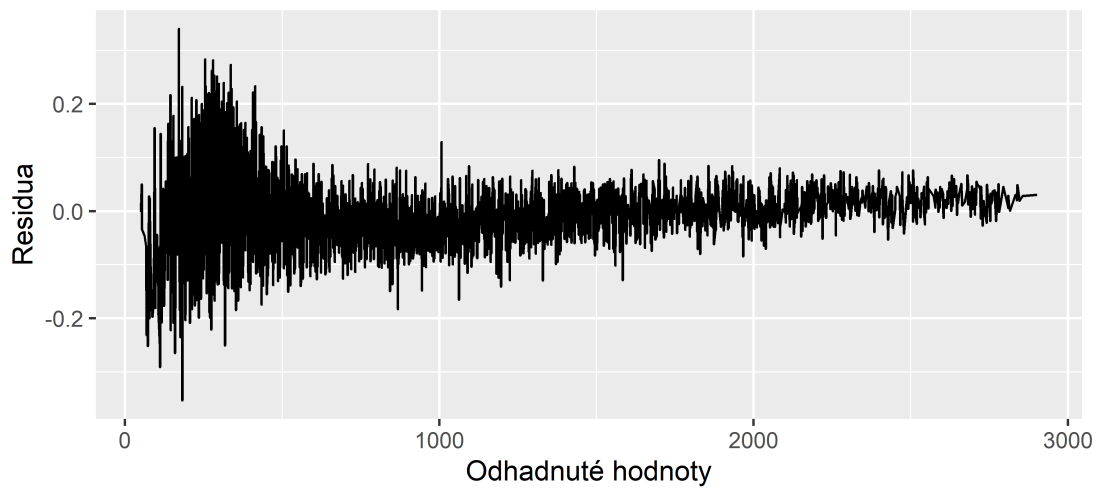
Obrázek A.6: Residua (CBD) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.

Odhad Cairns–Blake–Dowdůva modelu s kohortním členem



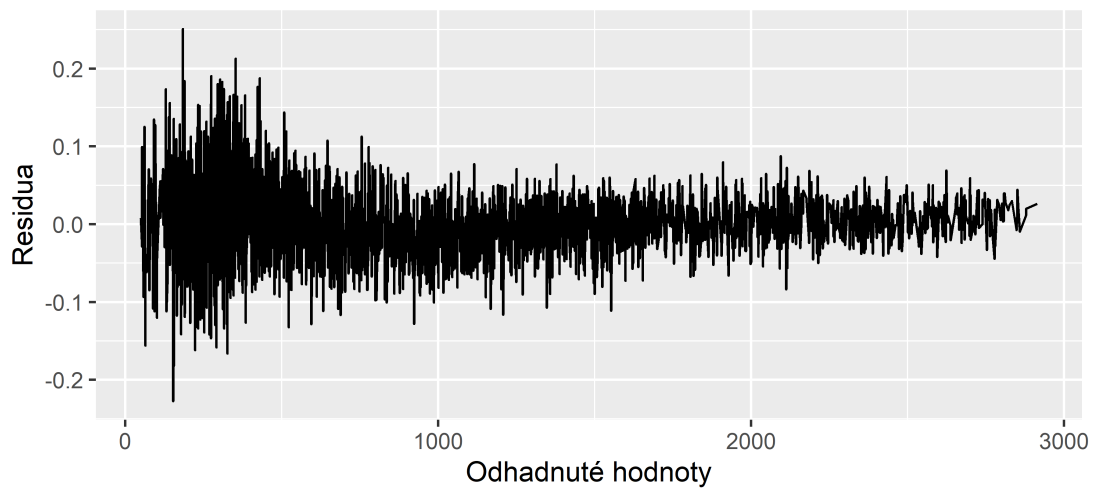
Obrázek A.7: Odhad (CBD-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}$, $\kappa_y^{(2)}$ a γ_c . V posledním grafu (vpravo dole) je černou křivkou vyobrazen odhad (CBD-C) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.

Residua Cairns–Blake–Dowdůva modelu s kohortním členem



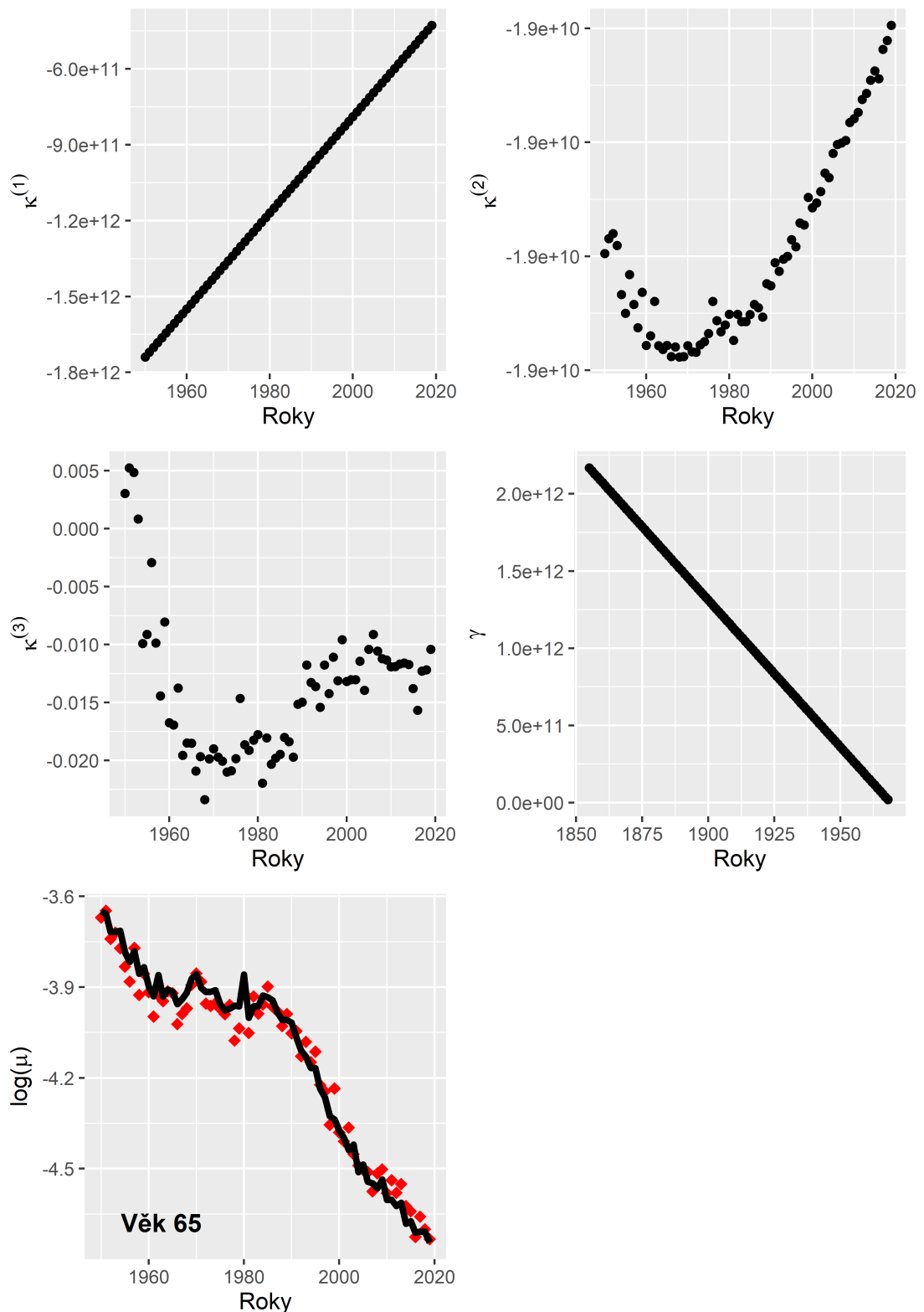
Obrázek A.8: Residua (CBD-C) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.

Residua Cairns–Blake–Dowdůva modelu s variačním členem



Obrázek A.9: Residua (CBD-V) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019.

Odhad Cairns–Blake–Dowdůva modelu s variačním členem



Obrázek A.10: Odhad (CBD-V) modelu na českých datech úmrtnosti žen pro věky 50-95 od roku 1950 do 2019. Černé body v grafech odpovídají postupně odhadu $\kappa_y^{(1)}$, $\kappa_y^{(2)}$, $\kappa_y^{(3)}$ a γ_c . V posledním grafu je černou křivkou vyobrazen odhad (CBD-V) modelu pro věk 65 a červené body (kosočtverce) představují skutečně realizované úmrtnosti.