

# Model izgladivanja empirijskih podataka

---

**Galović, Matea**

**Master's thesis / Diplomski rad**

**2016**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:217:514755>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-07-28**



*Repository / Repozitorij:*

[Repository of the Faculty of Science - University of Zagreb](#)



**SVEUČILIŠTE U ZAGREBU**  
**PRIRODOSLOVNO–MATEMATIČKI FAKULTET**  
**MATEMATIČKI ODSJEK**

Matea Galović

**MODEL IZGLAĐIVANJA**  
**EMPIRIJSKIH PODATAKA**

Diplomski rad

Voditelj rada:  
prof. dr. sc. Damir Bakić

Zagreb, 2016.

Ovaj diplomski rad obranjen je dana \_\_\_\_\_ pred ispitnim povjerenstvom u sastavu:

1. \_\_\_\_\_, predsjednik
2. \_\_\_\_\_, član
3. \_\_\_\_\_, član

Povjerenstvo je rad ocijenilo ocjenom \_\_\_\_\_.

Potpisi članova povjerenstva:

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

# Sadržaj

<b>Sadržaj</b>	<b>iii</b>
<b>Uvod</b>	<b>1</b>
<b>1 Modeli doživljenja</b>	<b>2</b>
1.1 Slučajna varijabla budućeg životnog vijeka . . . . .	2
1.2 Intenzitet smrtnosti . . . . .	5
1.3 Aktuarska notacija . . . . .	9
1.4 Očekivanje i standardna devijacija od $T_x$ . . . . .	11
1.5 Cjelobrojno buduće trajanje života . . . . .	14
1.6 Završne napomene . . . . .	16
<b>2 Tablice smrtnosti</b>	<b>17</b>
2.1 Osnovno o tablicama smrtnosti . . . . .	17
2.2 Pretpostavke o razlomljenoj dobi . . . . .	20
2.3 Nacionalne tablice smrtnosti . . . . .	25
2.4 Dodatne opaske vezane uz tablice smrtnosti . . . . .	27
2.5 Karupov interpolacijski polinom . . . . .	31
2.6 Spragueov interpolacijski polinom . . . . .	34
<b>3 Primjena rezultata uz Karupovu interpolaciju</b>	<b>36</b>
3.1 Računanje koeficijenata . . . . .	36
3.2 Grafički prikaz . . . . .	41
<b>Bibliografija</b>	<b>45</b>

# Uvod

U ovome radu predstaviti ćemo neke osnovne aktuarske matematike, s naglaskom na konstrukciju glatke funkcije  $l_x$ , koja predstavlja broj živih osoba u dobi  $x$ . Potreba za prvim aktuarima pojavila se u 18. stoljeću, kada su ih zapošljavala društva za životna osiguranja kako bi stvorila znanstveni temelj za upravljanje imovinom i obvezama društva. Obveze su ovisile o broju smrti koje bi se dogodile među osiguranim životima tijekom svake godine. Modeliranje smrtnosti postalo je tema od komercijalnog, ali i znanstvenog interesa, što je privlačilo mnoge značajne znanstvenike i matematičare aktuarskoj problematici. To je za rezultat imalo, među ostalim, i da su mnoga rana djela na polju vjerojatnosti bila usko povezana s razvojem rješenja aktuarskih problema.

U prvome poglavlju rada predstaviti ćemo modele doživljenja, uz definicije osnovnih pojmova kao što su intenzitet smrtnosti i cjelobrojno buduće trajanje života. Predstaviti ćemo budući životni vijek kao slučajnu varijablu i pokazati kako se unutar tog okvira mogu izračunati vjerojatnosti smrtnosti i doživljenja. Također uz mnoštvo primjera, uvodimo i donekle specifičnu aktuarsku notaciju.

U drugom poglavlju definiramo tablicu smrtnosti. Za tablicu tabeliranu samo po cjelobrojnim dobima, pokazujemo, koristeći pretpostavke o razlomljenoj dobi, kako izračunati vjerojatnosti doživljenja za sve dobi. Zatim diskutiramo neka svojstva nacionalnih tablica smrtnosti iz Australije, Engleske i SAD-a. Nakon poopćenja nekih rezultata, predstavljamo nekoliko opcija za aproksimaciju intenziteta smrtnosti, a zatim u posljednja dva potpoglavlja predstavljamo Karupov i Spragueov interpolacijski polinom, koji bi se mogli iskoristiti za izgladivanje empirijskih podataka  $l_x$  danih samo za cjelobrojne dobi  $x$ .

Konačno, u trećem poglavlju donosimo račun na temelju Karupove interpolacije, pri čemu koristimo stvarne podatke iz aktualnih Tablica aktuarske matematike za Republiku Hrvatsku (vidi [4]). Pronalazimo eksplicitnu formulu za  $l_x$  na segmentu  $[0, 101]$  izraženu po kraćim segmentima  $[i, i + 1]$ ,  $i = 0, 1, \dots, 100$  te donosimo grafički prikaz tako dobivene funkcije  $l_x$ , ali i na temelju nje izvedene aproksimacije intenziteta smrtnosti  $\mu_x$ .

# Poglavlje 1

## Modeli doživljenja

### 1.1 Slučajna varijabla budućeg životnog vijeka

Mnoge police osiguranja osiguravaju benefit prilikom smrti nositelja police. Kad osiguravajuće društvo izdaje takvu policu, datum smrti nositelja police nepoznat je, pa osiguravatelj ne zna točno kada će morati isplatiti benefit. Kako bi procijenio vrijeme isplate, osiguravatelju je potreban model smrtnosti ljudi, iz kojeg se mogu izračunati vjerojatnosti smrti u pojedinoj životnoj dobi.

Počnimo uvođenjem notacije. Neka  $(x)$  označava život starosti  $x$ , pri čemu je  $x \geq 0$ . Smrt od  $(x)$  može nastupiti u svakoj dobi većoj od  $(x)$  te preostalo trajanje života od  $(x)$  modeliramo neprekidnom slučajnom varijablom koju označavamo s  $T_x$ . Dakle,  $x + T_x$  predstavlja slučajnu varijablu ukupnog trajanja života, tj. starosti prilikom smrti za  $(x)$ .

Neka je  $F_x$  funkcija distribucije slučajne varijable  $T_x$ , odnosno

$$F_x(t) = \mathbb{P}(T_x \leq t).$$

Tada  $F_x(t)$  predstavlja vjerojatnost da  $(x)$  ne doživi dob  $x + t$ , te  $F_x$  nazivamo funkcijom distribucije preostalog trajanja života u dobi  $x$ . U mnogim primjenama vezanima uz životno osiguranje više nas zanima vjerojatnost doživljenja nego smrti te stoga definiramo  $S_x$  kao

$$S_x(t) = 1 - F_x(t) = \mathbb{P}(T_x > t).$$

Vidimo da  $S_x(t)$  predstavlja vjerojatnost da  $(x)$  doživi još barem  $t$  godina te  $S_x$  nazivamo funkcijom doživljenja.

Uz danu interpretaciju skupa slučajnih varijabli  $\{T_x\}_{x \geq 0}$  kao svih mogućih preostalih trajanja života pojedinaca, potrebna nam je veza između bilo koja dva predstavnika te kolekcije. Kako bismo došli do toga, uzmimo u obzir  $T_0$  i  $T_x$  za nekog

pojedinaca trenutno starog  $x$  godina. Slučajna varijabla  $T_0$  predstavlja budući životni vijek tog pojedinca pri rođenju, odnosno pri rođenju je dob prilikom smrti tog pojedinca predstavljena s  $T_0$ . Taj pojedinac mogao je umrijeti i prije nego navršši  $x$  godina – vjerojatnost tog događaja bila je  $\mathbb{P}(T_0 < x)$ , ali uzimamo da je ipak preživio. Budući da je preživio do dobi  $x$ , pa je  $T_0 > x$ , njegov budući životni vijek predstavljen je s  $T_x$  te je sada odgovarajuća dob prilikom smrti  $x + T_x$ . Ako pojedinac umre unutar sljedećih  $t$  godina, tada je  $T_x \leq t$  i  $T_0 \leq x + t$ . Neformalno, zahtijevamo da događaji  $\{T_x \leq t\}$  i  $\{T_0 \leq x + t\}$  budu ekvivalentni, uzimajući u obzir da je pojedinac doživio dob  $x$ . To zapisujemo formuliranjem sljedeće pretpostavke za sve  $x \geq 0$  i za sve  $t > 0$ :

$$\mathbb{P}(T_x \leq t) = \mathbb{P}(T_0 \leq x + t | T_0 > x). \quad (1.1)$$

Prisjetimo se iz teorije vjerojatnosti (npr. vidi [3]) da za dva događaja  $A$  i  $B$  vrijedi

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

pa ako interpretiramo događaj  $\{T_0 \leq x + t\}$  kao događaj  $A$  i događaj  $\{T_0 > x\}$  kao događaj  $B$ , možemo desnu stranu jednadžbe (1.1) zapisati kao u gornjoj jednakosti, odakle imamo

$$\mathbb{P}(T_x \leq t) = \frac{\mathbb{P}(x < T_0 \leq x + t)}{\mathbb{P}(T_0 > x)},$$

odnosno,

$$F_x(t) = \frac{F_0(x + t) - F_0(x)}{S_0(x)}. \quad (1.2)$$

Također, koristeći  $S_x(t) = 1 - F_x(t)$ , imamo

$$S_x(t) = \frac{S_0(x + t)}{S_0(x)}, \quad (1.3)$$

što se može zapisati i kao

$$S_0(x + t) = S_0(x)S_x(t). \quad (1.4)$$

Ovo je vrlo važan rezultat. On pokazuje da vjerojatnost doživljenja od dobi  $x$  do dobi  $x + t$  možemo interpretirati kao produkt

1. vjerojatnosti doživljenja od rođenja do dobi  $x$ , i

2. vjerojatnosti doživljenja dobi  $x + t$  kad je osoba već doživjela dob  $x$ .

Primijetimo kako  $S_x(t)$  možemo shvatiti kao vjerojatnost da (0) doživi barem dob  $x + t$  uz uvjet da (0) doživi dob  $x$ , pa ovaj rezultat možemo izvesti iz standardne vjerojatnosne jednadžbe koju smo već zapisali u nešto drukčijem obliku

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B),$$

gdje su događaji  $A = \{T_0 > x + t\}$  i  $B = \{T_0 > x\}$ , pa je

$$\mathbb{P}(A|B) = \mathbb{P}(T_0 > x + t | T_0 > x),$$

što je po jednadžbi (1.1) jednako vjerojatnosti  $\mathbb{P}(T_x > t)$ .

Slično, bilo koja vjerojatnost doživljenja za  $(x)$ , primjerice doživljenja još  $t + u$  godina može se podijeliti na vjerojatnost doživljenja prvih  $t$  godina, a potom i narednih  $u$  godina, uz uvjet doživljenja dobi  $x + t$ . Odnosno,

$$\begin{aligned} S_x(t + u) &= \frac{S_0(x+t+u)}{S_0(x)} \\ \implies S_x(t) &= \frac{S_0(x+t)}{S_0(x)} \frac{S_0(x+t+u)}{S_0(x+t)} \\ \implies S_x(t + u) &= S_x(t) S_{x+t}(u) \end{aligned} \tag{1.5}$$

Već smo vidjeli da ukoliko znamo vjerojatnosti doživljenja u trenutku rođenja, tada koristeći formulu (1.4) također znamo i vjerojatnosti doživljenja za istog pojedinca u bilo kojoj budućoj dobi. Formula (1.5) dodatno poopćava taj rezultat. Ona pokazuje da ukoliko znamo vjerojatnosti doživljenja u bilo kojoj dobi  $x (\geq 0)$ , znamo i vjerojatnosti doživljenja u bilo kojoj budućoj dobi  $x + t (\geq x)$ .

Kako bi bila valjana, bilo koja funkcija doživljenja za distribuciju životnog vijeka mora zadovoljavati sljedeće uvjete.

**Uvjet 1.**  $S_x(0) = 1$ ; odnosno, vjerojatnost da  $(x)$  preživi 0 godina je 1.

**Uvjet 2.**  $\lim_{t \rightarrow \infty} S_x(t) = 0$ ; odnosno, svaki život mora završiti u nekom trenutku.

**Uvjet 3.** Funkcija doživljenja mora biti nerastuća funkcija od  $t$ ; ne može biti vjerojatnije da  $(x)$  preživi još npr. 10.5 godina nego još 10 godina, budući da se prvo mora dogoditi da preživi još 10 godina.

Ovi uvjeti su nužni i dovoljni, pa stoga svaka funkcija  $S_x$  koja je funkcija od  $t (\geq 0)$  za fiksiran  $x (\geq 0)$  i koja zadovoljava ova tri uvjeta definira distribuciju životnog vijeka od dobi  $x$ , pa posljedično, koristeći formulu (1.5) i za sve kasnije životne dobi, veće od  $x$ .

Za sve takve distribucije, uvodimo još tri dodatne pretpostavke:



**Pretpostavka 1.**  $S_x(t)$  je diferencijabilna za sve  $t > 0$ . Primijetimo da zajedno s uvjetom 3 odozgo to znači  $\frac{d}{dx}S_x(t) \leq 0$  za sve  $t > 0$ .

**Pretpostavka 2.**  $\lim_{t \rightarrow \infty} tS_x(t) = 0$

**Pretpostavka 3.**  $\lim_{t \rightarrow \infty} t^2S_x(t) = 0$

Posljednje dvije pretpostavke osiguravaju postojanje očekivanja i varijance slučajne varijable  $T_x$ . To nisu posebno restriktivna ograničenja – ne trebamo se brinuti o distribucijama s beskonačnim očekivanjem ili varijancom u kontekstu pojedinčevog budućeg životnog vijeka. Ove tri dodatne pretpostavke vrijede za sve distribucije koje su prikladne za modeliranje ljudskog životnog vijeka.

**Primjer 1.1.** Neka je  $F_0(t) = 1 - (1 - \frac{t}{120})^{\frac{1}{6}}$  za  $0 \leq t \leq 120$ . Izračunajmo vjerojatnost da

(a) novorođenče doživi dob 30 godina,

(b) osoba stara 30 godina umre prije dobi od 50 godina,

(c) osoba stara 40 godina doživi dob 65 godina.

Rješenje. (a)  $S_0(30) = 1 - F_0(30) = (1 - \frac{30}{120})^{\frac{1}{6}} = 0.9532$

(b) Iz formule (1.2) imamo  $F_{30}(20) = \frac{F_0(50) - F_0(30)}{1 - F_0(30)} = 0.0410$ .

(c) Iz formule (1.3) imamo  $S_{40}(25) = \frac{S_0(65)}{S_0(40)} = 0.9395$ .

□

Primijetimo da je u gornjem primjeru  $S_0(120) = 0$ , što znači da u ovom modelu doživljenje iznad 120 godina nije moguće. U ovom slučaju prema dobi 120 se odnosimo kao graničnoj dobi modela. Općenito, ako postoji granična dob, označujemo je grčkim slovom  $\omega$ . U modelima u kojima ne postoji granična dob, često je ipak praktično uvesti ju u izračunima.

## 1.2 Intenzitet smrtnosti

Intenzitet smrtnosti (mortaliteta) je važan i fundamentalan koncept u modeliranju budućeg životnog vijeka. Označavamo intenzitet smrtnosti u dobi  $x$  sa

$$\mu_x = \lim_{dx \rightarrow 0^+} \frac{1}{dx} \mathbb{P}(T_0 \leq x + dx | T_0 > x). \quad (1.6)$$

Iz jednadžbe (1.1) vidimo da je ekvivalentan način definiranja  $\mu_x$

$$\mu_x = \lim_{dx \rightarrow 0^+} \frac{1}{dx} \mathbb{P}(T_x \leq dx),$$

što se može zapisati u terminima funkcije doživljenja  $S_x$  kao

$$\mu_x = \lim_{dx \rightarrow 0^+} \frac{1}{dx} (1 - S_x(dx)). \quad (1.7)$$

Primijetimo da intenzitet smrtnosti ovisi o jedinici vremena – ako vrijeme mjerimo u godinama, tada se i  $\mu_x$  mjeri u godinama.

Intenzitet smrtnosti najbolje možemo shvatiti ako uočimo da za vrlo mali  $dx$  formula (1.6) daje aproksimaciju

$$\mu_x dx \approx \mathbb{P}(T_0 \leq x + dx | T_0 > x). \quad (1.8)$$

Dakle, za vrlo mali  $dx$  možemo  $\mu_x dx$  interpretirati kao vjerojatnost da život koji je dosegao dob  $x$  završi prije dostizanja dobi  $x + dx$ . Npr. pretpostavimo da imamo muškarca starog točno 50 godina i da je intenzitet smrtnosti u toj dobi 0.0044 po godini. Mala vrijednost od  $dx$  može biti jedan dan, odnosno 0.00274 godina. Tada je približna vjerojatnost da će ta osoba umrijeti na dan svog pedesetog rođendana jednaka  $0.0044 \times 0.00274 = 1.2 \times 10^{-5}$ .

Intenzitet smrtnosti možemo povezati s funkcijom doživljenja u trenutku rođenja,  $S_0$ . Budući da je

$$S_x(dx) = \frac{S_0(x + dx)}{S_0(x)},$$

formula (1.7) daje

$$\begin{aligned} \mu_x &= \frac{1}{S_0(x)} \lim_{dx \rightarrow 0^+} \frac{S_0(x) - S_0(x + dx)}{dx} \\ &= \frac{1}{S_0(x)} \left( - \frac{d}{dx} S_0(x) \right). \end{aligned}$$

Dakle, imamo

$$\mu_x = \frac{-1}{S_0(x)} \frac{d}{dx} S_0(x). \quad (1.9)$$

Iz standardnih rezultata teorije vjerojatnosti (npr. vidi [3]) znamo da je vjerojatnosna funkcija gustoće za slučajnu varijablu  $T_x$ , koju označavamo s  $f_x$  povezana s funkcijom distribucije  $F_x$  i funkcijom doživljenja  $S_x$  preko

$$f_x(t) = \frac{d}{dt}F_x(t) = -\frac{d}{dt}S_x(t),$$

odakle, uz jednakost (1.9) slijedi

$$\mu_x = \frac{f_0(x)}{S_0(x)}.$$

Također možemo povezati intenzitet smrtnosti u bilo kojoj dobi  $x + t, t > 0$ , s distribucijom od  $T_x$ . Pretpostavimo da je  $x$  fiksiran i  $t$  varijabilan. Tada je  $d(x+t) = dt$  pa je

$$\begin{aligned} \mu_{x+t} &= -\frac{1}{S_0(x+t)} \frac{d}{d(x+t)} S_0(x+t) \\ &= -\frac{1}{S_0(x+t)} \frac{d}{dt} S_0(x+t) \\ &= -\frac{1}{S_0(x+t)} \frac{d}{dt} (S_0(x)S_x(t)) \\ &= -\frac{S_0(x)}{S_0(x+t)} \frac{d}{dt} S_x(t) \\ &= \frac{-1}{S_x(t)} \frac{d}{dt} S_x(t). \end{aligned}$$

Dakle

$$\mu_{x+t} = \frac{f_x(t)}{S_x(t)}. \quad (1.10)$$

Ova veza daje način pronalaska  $\mu_{x+t}$  uz dani  $S_x(t)$ . Također možemo koristiti jednakost (1.9) kako bismo pronašli formulu za  $S_x(t)$  u terminima funkcije intenziteta smrtnosti. Koristimo činjenicu da za derivabilnu funkciju  $h$  vrijedi

$$\frac{d}{dx} \ln h(x) = \frac{1}{h(x)} \frac{d}{dx} h(x),$$

pa iz jednadžbe (1.9) imamo

$$\mu_x = -\frac{d}{dx} \ln S_0(x),$$

odakle integriranjem na  $(0, y)$  dobivamo

$$\int_0^y \mu_x dx = -(\ln S_0(y) - \ln S_0(0)).$$

Budući da je  $\ln \mathbb{P}(T_0 > 0) = \ln 1 = 0$ , imamo

$$S_0(y) = \exp \left\{ - \int_0^y \mu_x dx \right\},$$

odakle slijedi

$$\begin{aligned} S_x(t) &= \frac{S_0(x+t)}{S_0(x)} \\ &= \frac{\exp \left\{ - \int_0^{x+t} \mu_r dr \right\}}{\exp \left\{ - \int_0^x \mu_r dr \right\}} = \exp \left\{ - \int_0^{x+t} \mu_r dr + \int_0^x \mu_r dr \right\} \\ &= \exp \left\{ - \int_x^{x+t} \mu_r dr \right\} \stackrel{\text{supstitucija } s=r-x}{=} \exp \left\{ - \int_0^t \mu_{x+s} ds \right\}. \end{aligned} \quad (1.11)$$

Ovo znači da ukoliko znamo  $\mu_x$  za sve  $x \geq 0$ , možemo izračunati i sve vjerojatnosti doživljenja  $S_x(t)$  za proizvoljne  $x$  i  $t$ . Drugim riječima, intenzitet smrtnosti u potpunosti opisuje distribuciju životnog vijeka, jednako kao što to čini i funkcija  $S_0$ . U biti, često je prikladnije opisivati distribuciju životnog vijeka pomoću intenziteta smrtnosti nego pomoću funkcije doživljenja.

**Primjer 1.2.** *Kao u primjeru 1.1, neka je  $F_0(x) = 1 - (1 - \frac{x}{120})^{\frac{1}{6}}$  za  $0 \leq x \leq 120$ . Izvedimo izraz za  $\mu_x$ .*

*Rješenje.* Budući da je  $S_0(x) = (1 - \frac{x}{120})^{\frac{1}{6}}$ , slijedi

$$\frac{d}{dx} S_0(x) = \frac{1}{6} \left(1 - \frac{x}{120}\right)^{-\frac{5}{6}} \left(-\frac{1}{120}\right),$$

pa je

$$\mu_x = \frac{-1}{S_0(x)} \frac{d}{dx} S_0(x) = \frac{1}{720} \left(1 - \frac{x}{120}\right)^{-1} = \frac{1}{720 - 6x}.$$

Alternativno, mogli bismo koristiti

$$\mu_x = -\frac{d}{dx} \ln S_0(x) = -\frac{d}{dx} \left( \frac{1}{6} \ln \left(1 - \frac{x}{120}\right) \right) = \frac{1}{720 \left(1 - \frac{x}{120}\right)} = \frac{1}{720 - 6x}.$$

□

**Primjer 1.3.** Neka je  $\mu_x = Bc^x$ ,  $x > 0$ , gdje su  $B$  i  $c$  konstante takve da je  $0 < B < 1$  i  $c > 1$ . Ovaj model zove se **Gompertzov zakon smrtnosti**. Izvedimo izraz za  $S_x(t)$ .

*Rješenje.* Iz jednakosti (1.11) imamo

$$S_x(t) = \exp \left\{ - \int_x^{x+t} Bc^r dr \right\}.$$

Ako  $c^r$  izrazimo kao  $\exp\{r \ln c\}$ , imamo

$$\int_x^{x+t} Bc^r dr = B \int_x^{x+t} \exp\{r \ln c\} dr = \frac{B}{\ln c} \exp\{r \ln c\} \Big|_x^{x+t} = \frac{B}{\ln c} (c^{x+t} - c^x),$$

odakle slijedi

$$S_x(t) = \exp \left\{ \frac{-B}{\ln c} c^x (c^t - 1) \right\}.$$

□

### 1.3 Aktuarska notacija

Notacija korištena u prethodnom tekstu,  $S_x(t)$ ,  $F_x(t)$  i  $f_x(t)$ , standardna je u statistici. Aktuarska znanost razvila je svoju vlastitu notaciju, Međunarodnu aktuarsku notaciju. Ona sažima vjerojatnosti i funkcije koje su od najvećeg interesa i koristi aktuarima. Oznaka za intenzitet mortaliteta,  $\mu_x$ , dolazi upravo iz Međunarodne aktuarske notacije. U ovom potpoglavlju donosimo sažetak relevantne aktuarske notacije i prepisujemo dosad spomenute bitne rezultate u terminima te notacije. Aktuarska notacija za vjerojatnosti doživljenja i smrtnosti je

$${}_t p_x = \mathbb{P}(T_x > t) = S_x(t), \quad (1.12)$$

$${}_t q_x = \mathbb{P}(T_x \leq t) = 1 - S_x(t) = F_x(t), \quad (1.13)$$

$${}_u | {}_t q_x = \mathbb{P}(u < T_x \leq u + t) = S_x(u) - S_x(u + t). \quad (1.14)$$

Dakle,

- ${}_t p_x$  je vjerojatnost da  $(x)$  doživi barem dob  $x + t$

- ${}_tq_x$  je vjerojatnost da  $(x)$  umre prije navršene dobi  $x + t$
- ${}_u|_tq_x$  je vjerojatnost da  $(x)$  doživi još  $u$  godina i potom umre u sljedećih  $t$  godina, odnosno u starosti između  $x + u$  i  $x + u + t$  godina. Ova vjerojatnost zove se **odgođena vjerojatnost smrtnosti**, budući da je to vjerojatnost da smrt nastupi u nekom odgođenom intervalu.

U slučaju da je vrijednost indeksa  $t$  jednaka 1, možemo izostaviti taj indeks, pa  $p_x$  predstavlja vjerojatnost da  $(x)$  doživi barem dob  $x + 1$ . Slično,  $q_x$  je vjerojatnost da  $(x)$  umre prije dobi  $x + 1$ . U aktuarskoj se terminologiji  $q_x$  zove **stopa smrtnosti** u dobi  $x$ .

Sljedeće jednakosti slijede izravno iz gornjih definicija i prethodnih rezultata u ovom poglavlju:

$${}_tp_x + {}_tq_x = 1,$$

$${}_u|_tq_x = {}_up_x - {}_{u+t}p_x,$$

$${}_{t+u}p_x = {}_tp_x {}_up_{x+t}, \text{ iz formule (1.5),} \quad (1.15)$$

$$\mu_x = -\frac{1}{{}_xp_0} \frac{d}{dx} {}_xp_0, \text{ iz formule (1.9).} \quad (1.16)$$

Slično,

$$\mu_{x+t} = -\frac{1}{{}_tp_x} \frac{d}{dt} {}_tp_x \implies \frac{d}{dt} {}_tp_x = -{}_tp_x \mu_{x+t}, \quad (1.17)$$

$$\mu_{x+t} = \frac{f_x(t)}{S_x(t)} \implies f_x(t) = {}_tp_x \mu_{x+t}, \text{ iz formule (1.10)} \quad (1.18)$$

$${}_tp_x = \exp \left\{ -\int_0^t \mu_{x+s} ds \right\}, \text{ iz formule (1.11).} \quad (1.19)$$

Kako je  $F_x$  funkcija distribucije i  $f_x$  pripadna funkcija gustoće, slijedi

$$F_x(t) = \int_0^t f_x(s) ds,$$

što u aktuarskoj notaciji možemo zapisati kao

$${}_tq_x = \int_0^t {}_sp_x \mu_{x+s} ds. \quad (1.20)$$

Ovo je vrlo važna formula i može se interpretirati na sljedeći način: uzmimo neko vrijeme  $s$ , gdje je  $0 \leq s < t$ . Vjerojatnost da je  $(x)$  živ u trenutku  $s$  je  ${}_s p_x$ , a vjerojatnost da  $(x)$  umre u starosti između  $x + s$  i  $x + s + ds$  (pri čemu je doživio dob  $x + s$ , a  $ds$  je vrlo malen) grubo govoreći iznosi  $\mu_{x+s} ds$ . Dakle,  ${}_s p_x \mu_{x+s} ds$  može se interpretirati kao vjerojatnost da  $(x)$  umre između dobi  $x + s$  i  $x + s + ds$ . Kako bismo dobili vjerojatnost smrti prije dobi  $x + t$ , to možemo sumirati na svim mogućim intervalima od  $s$  do  $s + ds$  u kojima može nastupiti smrt, što zahtijeva integriranje budući da se radi o infitezimalnim intervalima. Ova vrsta interpretacije je važna u tome što može biti primijenjena i u kompliciranijim situacijama.

U posebnom slučaju kad je  $t = 1$ , formula (1.20) postaje

$$q_x = \int_0^1 {}_s p_x \mu_{x+s} ds.$$

Kada je  $q_x$  malen, slijedi da je  $p_x$  blizu 1, pa je i  ${}_s p_x$  blizu 1 za  $0 \leq s < 1$ , pa imamo

$$q_x \approx \int_0^1 \mu_{x+s} ds \approx \mu_{x+1/2} = \mu_{x+1/2} \cdot ((x+1) - x),$$

gdje druga aproksimacija slijedi iz pravila pravokutnika za numeričku integraciju (gdje se vrijednost funkcije  $\mu$  na segmentu aproksimira s vrijednošću funkcije u polovištu tog segmenta).

**Primjer 1.4.** *Kao u primjeru 1.1, neka je  $F_0(x) = 1 - (1 - \frac{x}{120})^{\frac{1}{6}}$  za  $0 \leq x \leq 120$ . Izračunajmo  $q_x$  i  $\mu_{x+1/2}$  za  $x = 20$  i  $x = 110$  te komentirajmo rezultate.*

*Rješenje.* Imamo

$$p_x = \frac{S_0(x+1)}{S_0(x)} = \left(1 - \frac{1}{120-x}\right)^{\frac{1}{6}},$$

što daje  $q_{20} = 0.00167$  i  $q_{110} = 0.01741$ , a iz rješenja primjera 1.2 imamo  $\mu_{20\frac{1}{2}} = 0.00168$ , odnosno  $\mu_{110\frac{1}{2}} = 0.01754$ . Vidimo da je  $\mu_{x+1/2}$  dobra aproksimacija za  $q_x$  kad je stopa smrtnosti mala, ali i ne tako dobra aproksimacija, barem u terminima apsolutne greške, kad stopa smrtnosti nije blizu 0.  $\square$

## 1.4 Očekivanje i standardna devijacija od $T_x$

U ovom potpoglavlju promatramo očekivani budući životni vijek od  $(x)$ :  $\mathbb{E}[T_x]$ , što se u aktuarskoj notaciji označava s  $\overset{\circ}{e}_x$ , a naziva još i **potpuno očekivanje života**. Kako bismo evaluirali  $\overset{\circ}{e}_x$ , čitamo iz formula (1.17) i (1.18)

$$f_x(t) = {}_t p_x \mu_{x+t} = -\frac{d}{dt} {}_t p_x. \quad (1.21)$$

Iz definicije očekivanja imamo

$$\dot{e}_x = \int_0^\infty t f_x(t) dt = \int_0^\infty t {}_t p_x \mu_{x+t} dt.$$

Sada možemo iskoristiti (1.21) kako bismo evaluirali ovaj integral (koristimo parcijalnu integraciju uz  $u = t$  i  $dv = \left(\frac{d}{dt} {}_t p_x\right) dt$ ):

$$\dot{e}_x = -\int_0^\infty t \left(\frac{d}{dt} {}_t p_x\right) dt = -\left(t {}_t p_x \Big|_0^\infty - \int_0^\infty {}_t p_x dt\right).$$

U prvom potpoglavlju pretpostavili smo  $\lim_{t \rightarrow \infty} t {}_t p_x = 0$ , pa gornja jednakost postaje

$$\dot{e}_x = \int_0^\infty {}_t p_x dt. \quad (1.22)$$

Slično, za  $\mathbb{E}[T_x^2]$  imamo (koristimo parcijalnu integraciju uz  $u = t^2$  i  $dv = \left(\frac{d}{dt} {}_t p_x\right) dt$ )

$$\begin{aligned} \mathbb{E}[T_x^2] &= \int_0^\infty t^2 {}_t p_x \mu_{x+t} dt \\ &= -\int_0^\infty t^2 \left(\frac{d}{dt} {}_t p_x\right) dt \\ &= -\left(t^2 {}_t p_x \Big|_0^\infty - \int_0^\infty {}_t p_x 2t dt\right) \\ &= 2 \int_0^\infty t {}_t p_x dt. \end{aligned} \quad (1.23)$$

Dakle, imamo integralne izraze za  $\mathbb{E}[T_x]$  i  $\mathbb{E}[T_x^2]$ . Za neke distribucije varijable  $T_x$ , možemo integrirati direktno, dok u ostalim slučajevima trebamo koristiti tehnike numeričke integracije kako bismo evaluirali integrale u formulama (1.22) i (1.23). Varijanca od  $T_x$  tada se može izračunati kao

$$\text{Var}(T_x) = \mathbb{E}[T_x^2] - (\dot{e}_x)^2.$$

**Primjer 1.5.** Neka je kao u primjeru 1.1  $F_0(x) = 1 - \left(1 - \frac{x}{120}\right)^{\frac{1}{6}}$  za  $0 \leq x \leq 120$ . Izračunajmo  $\dot{e}_x$  i  $\text{Var}(T_x)$  za

(a)  $x = 30$



(b)  $x = 80$ .

Rješenje. Imamo

$${}_t p_x = \frac{S_0(x+t)}{S_0(x)} = \left(1 - \frac{t}{120-x}\right)^{\frac{1}{6}}.$$

Sjetimo se da ta formula vrijedi za  $0 \leq t \leq 120 - x$ , budući da u ovom modelu doživljenje dobi veće od 120 nije moguće. Tehnički, imamo

$${}_t p_x = \begin{cases} \left(1 - \frac{t}{120-x}\right)^{\frac{1}{6}} & \text{za } x+t \leq 120, \\ 0 & \text{za } x+t > 120. \end{cases}$$

Stoga je gornja granica integracije u jednadžbi (1.22) jednaka  $120 - x$ , i

$$\dot{e}_x = \int_0^{120-x} \left(1 - \frac{t}{120-x}\right)^{\frac{1}{6}} dt.$$

Iz supstitucije  $y = 1 - \frac{t}{120-x}$  slijedi  $t = (120-x)(1-y)$ , što daje

$$\dot{e}_x = (120-x) \int_0^1 y^{\frac{1}{6}} dy = \frac{6}{7}(120-x).$$

Odatle je  $\dot{e}_{30} = 77.143$  i  $\dot{e}_{80} = 34.286$ . Uočimo da je u ovom modelu očekivanje budućeg životnog vijeka uvijek jednako  $\frac{6}{7}$  godina potrebnih do doživljenja dobi 120. Kako bismo izračunali varijancu, potreban nam je  $\mathbb{E}[T_x^2]$ , pa iz (1.23) imamo

$$\mathbb{E}[T_x^2] = 2 \int_0^{120-x} {}_t p_x dt = 2 \int_0^{120-x} t \left(1 - \frac{t}{120-x}\right)^{\frac{1}{6}} dt.$$

Uz istu supstituciju kao gore imamo

$$\mathbb{E}[T_x^2] = 2(120-x)^2 \int_0^1 \left(y^{\frac{1}{6}} - y^{\frac{7}{6}}\right) dy = 2(120-x)^2 \left(\frac{6}{7} - \frac{6}{13}\right).$$

Konačno,

$$\begin{aligned} \text{Var}(T_x) &= \mathbb{E}[T_x^2] - (\dot{e}_x)^2 \\ &= (120-x)^2 \left(2 \left(\frac{6}{7} - \frac{6}{13}\right) - \left(\frac{6}{7}\right)^2\right) \\ &= (120-x)^2 (0.056515) \\ &= ((120-x)(0.23773))^2 \end{aligned}$$

pa je  $Var(T_{30}) = 21.396^2$  i  $Var(T_{80}) = 9.509^2$ .

Budući da znamo da u ovom modelu svi životi završavaju do dobi 120, ima smisla da je nesigurnost oko budućeg životnog vijeka veća za mlađe dobi nego starije.  $\square$

Svojstvo modela u prethodnom primjeru je da možemo doći do egzaktnih formula za veličine od interesa, kao što je  $\dot{e}_x$ , ali za mnogo modela to nije moguće. Na primjer, kada modeliramo smrtnost koristeći Gompertzov zakon, ne postoji eksplicitna formula za  $\dot{e}_x$  te moramo koristiti numeričku integraciju kako bismo izračunali momente od  $T_x$ .

## 1.5 Cjelobrojno buduće trajanje života

$K_x$  i  $e_x$

U brojnim primjenama u osiguranju ne zanima nas samo buduće trajanje života pojedinca, nego i podatak koji se zove cjelobrojno buduće trajanje života. To je slučajna varijabla koja se definira kao najveće cijelo preostalog budućeg trajanja života i označava s  $K_x$  za osobu staru  $x$  godina. Dakle, imamo

$$K_x = \lfloor T_x \rfloor.$$

Možemo razmišljati o cjelobrojnem budućem trajanju života kao o cijelom broju godina koje će pojedinac još doživjeti u budućnosti. Za ilustraciju, razmotrimo situaciju gdje osoba starosti  $x$  ima pravo na isplate 1 u trenucima  $1, 2, 3, \dots$ , uz uvjet da je u tim trenucima i dalje živa. Tada je broj izvršenih isplata jednak broju cijelih godina koje je ( $x$ ) doživio nakon trenutka 0, i upravo to je cjelobrojno buduće trajanje života.

Do vjerojatnosne distribucije od  $K_x$  dolazimo na način da uzimamo da je za  $k = 0, 1, 2, \dots$   $K_x = k$  ako i samo ako ( $x$ ) umre između navršениh dobi  $x + k$  i  $x + k + 1$ . Stoga za  $k = 0, 1, 2, \dots$

$$\begin{aligned} \mathbb{P}(K_x = k) &= \mathbb{P}(k \leq T_x < k + 1) \\ &= {}_k|q_x \\ &= {}_k p_x - {}_{k+1} p_x \\ &= {}_k p_x - {}_k p_x \cdot {}_k p_{x+k} \\ &= {}_k p_x \cdot q_{x+k}. \end{aligned}$$

Očekivana vrijednost od  $K_x$  označava se s  $e_x$ , pa imamo

$$\begin{aligned}
\mathbb{E}[K_x] &= e_x \\
&= \sum_{k=0}^{\infty} k\mathbb{P}(K_x = k) \\
&= \sum_{k=0}^{\infty} k(kp_x - {}_{k+1}p_x) \\
&= (1p_x - 2p_x) + 2(2p_x - 3p_x) + 3(3p_x - 4p_x) + \dots \\
&= \sum_{k=1}^{\infty} kp_x.
\end{aligned} \tag{1.24}$$

Primijetimo da je donja granica sumacije  $k = 1$ . Slično imamo

$$\begin{aligned}
\mathbb{E}[K_x^2] &= \sum_{k=0}^{\infty} k^2\mathbb{P}(K_x = k) \\
&= (1p_x - 2p_x) + 4(2p_x - 3p_x) + 9(3p_x - 4p_x) + 16(4p_x - 5p_x) + \dots \\
&= 1p_x + 3_2p_x + 5_3p_x + 7_4p_x + \dots \\
&= 2 \sum_{k=1}^{\infty} k_k p_x - \sum_{k=1}^{\infty} k p_x \\
&= 2 \sum_{k=1}^{\infty} k_k p_x - e_x.
\end{aligned}$$

Kao što je i s potpunim očekivanjem trajanja života, postoji vrlo malo distribucija od  $K_x$  čiji se  $\mathbb{E}[K_x]$  i  $\mathbb{E}[K_x^2]$  mogu analitički izračunati. Za realističnije modele kao što je Gompertzov, te vrijednosti možemo izračunati pomoću primjerenih softvera, kao što je Excel. Iako u principu moramo evaluirati beskonačnu sumu, u nekoj dobi vjerojatnost doživljenja postaje dovoljno mala da bi se mogla zanemariti, pa tu dob možemo smatrati efektivnom ograničavajućom dobi modela.

## Potpuno i cjelobrojno očekivano buduće trajanje života, $\overset{\circ}{e}_x$ i

$e_x$

Kako je cjelobrojno buduće trajanje života cijeli dio budućeg životnog vijeka, prirodno je zapitati se postoji li jednostavna veza između  $\overset{\circ}{e}_x$  i  $e_x$ . Vezu možemo aproksimativno zapisati kao

$$\dot{e}_x = \int_0^\infty {}_t p_x dt = \sum_{j=0}^{\infty} \int_j^{j+1} {}_t p_x dt.$$

Ako aproksimiramo svaki od ovih integrala koristeći pravilo trapeza za numeričku integraciju, imamo

$$\int_j^{j+1} {}_t p_x dt \approx \frac{1}{2}({}_j p_x + {}_{j+1} p_x),$$

dakle

$$\dot{e}_x \approx \sum_{j=0}^{\infty} \frac{1}{2}({}_j p_x + {}_{j+1} p_x) = \frac{1}{2} + \sum_{j=1}^{\infty} {}_j p_x.$$

Dobili smo aproksimaciju koja se često koristi u praksi:

$$\dot{e}_x \approx e_x + \frac{1}{2}. \quad (1.25)$$

## 1.6 Završne napomene

Unatoč tomu što su zakoni smrtnosti kao što je Gompertzov model privlačni zahvaljujući svojoj jednostavnosti, oni rijetko mogu valjano predstavljati smrtnost u cijelom rasponu životnih dobi. Jednostavno proširenje Gompertzovog zakona je Makehamov zakon (Makeham, 1860.), koji modelira intenzitet smrtnosti kao

$$\mu_x = A + Bc^x. \quad (1.26)$$

Ovo je vrlo slično Gompertzovom zakonu, ali dodan je fiksni pribrojnik koji ne ovisi o dobi, koji omogućuje bolju prilagodbu za slučajeve smrti uslijed nezgode. Dodatni pribrojnik poboljšava prilagodbu početnog modela podacima o smrtnosti za mlađe dobi.

U novije vrijeme, Gompertz-Makehamov pristup je dodatno poopćen u  $GM(r, s)$  (Gompertz-Makeham) formulu,

$$\mu_x = h_r^1(x) + \exp(h_s^2(x)),$$

gdje su  $h_r^1$  i  $h_s^2$  polinomi u  $x$ , stupnjeva  $r$  i  $s$ , respektivno. I Gompertzov i Makehamov zakon su posebni slučajevi  $GM$  formule.

U odjeljku 1.2. istaknuli smo važnost intenziteta smrtnosti. Ovdje treba istaknuti i kako je prilikom analize podataka o smrtnosti intenzitet smrtnosti prirodna veličina za procjenu, dok distribucija životnog vijeka nije. Za generalnije distribucije, veličina  $\frac{f_0(x)}{S_0(x)}$  koju aktuari zovu intenzitet smrtnosti u dobi  $x$ , poznata je i pod imenom **stopa hazarda** u analizi doživljenja i **stopa neuspjeha** u teoriji pouzdanosti.

## Poglavlje 2

# Tablice smrtnosti

### 2.1 Osnovno o tablicama smrtnosti

Uz dani model doživljenja i vjerojatnosti doživljenja  ${}_t p_x$ , možemo konstruirati **tablicu smrtnosti** za taj model počevši od neke inicijalne dobi  $x_0$  sve do maksimalne dobi  $\omega$ . Definiramo funkciju  $\{l_x\}$  za  $x_0 \leq x \leq \omega$  na sljedeći način: neka je  $x_0$  proizvoljan pozitivan broj (koji nazivamo **bazom** tablice) i za  $0 \leq t \leq \omega - x_0$  definiramo

$$l_{x_0+t} = l_{x_0} {}_t p_{x_0}.$$

Iz te definicije za  $x_0 \leq x \leq x+t \leq \omega$ , imamo

$$\begin{aligned} l_{x+t} &= l_{x_0} {}_{x+t-x_0} p_{x_0} \\ &= l_{x_0} {}_{x-x_0} p_{x_0} {}_t p_x \\ &= l_x {}_t p_x, \end{aligned}$$

odakle slijedi

$${}_t p_x = l_{x+t}/l_x. \tag{2.1}$$

Za bilo koji  $x \geq x_0$ , možemo interpretirati  $l_{x+t}$  kao očekivani broj preživjelih u dobi  $x+t$  od ukupno  $l_x$  nezavisnih pojedinaca dobi  $x$ . Ova interpretacija prirodnija je ako je  $l_x$  cijeli broj, i temelji se na činjenici da je broj preživjelih do dobi  $x+t$  slučajna varijabla binomne distribucije s parametrima  $l_x$  i  ${}_t p_x$ . Naime, pretpostavimo da imamo  $l_x$  nezavisnih života starosti  $x$ , pri čemu svaki život ima vjerojatnost doživljenja dobi  $x+t$  jednaku  ${}_t p_x$ . Tada je broj preživjelih do dobi  $x+t$  binomna slučajna varijabla  $\mathcal{L}_t$  s parametrima  $l_x$  i  ${}_t p_x$ . Očekivana vrijednost broja preživjelih tada je

$$\mathbb{E}[\mathcal{L}_t] = l_x {}_t p_x = l_{x+t}.$$

Uvijek koristimo tablicu u formi  $l_y/l_x$ , zbog čega je i baza tablice proizvoljna – ne bi bilo razlike za model doživljenja kad bismo, primjerice, sve vrijednosti  $l_x$  pomnožili sa 100.

Prema (2.1) možemo koristiti funkciju  $l_x$  kako bismo izračunali vjerojatnosti doživljenja, isto kao i vjerojatnosti smrtnosti. Na primjer,

$$q_{30} = 1 - \frac{l_{31}}{l_{30}} = \frac{l_{30} - l_{31}}{l_{30}}, \quad (2.2)$$

te

$${}_{15}|{}_{30}q_{40} = {}_{15}p_{40} {}_{30}q_{55} = \frac{l_{55}}{l_{40}} \left(1 - \frac{l_{85}}{l_{55}}\right) = \frac{l_{55} - l_{85}}{l_{40}}. \quad (2.3)$$

U principu, tablica smrtnosti definirana je za sve  $x$  od inicijalne dobi  $x_0$  do granične dobi  $\omega$ . U praksi je vrlo često da je tablica smrtnosti predstavljena, a u nekim slučajevima i definirana samo za cjelobrojne dobi. U toj je formi tablica smrtnosti koristan način sažimanja distribucije životnog vijeka, budući da svojim jednim stupcem omogućava računanje vjerojatnosti doživljenja ili smrti kroz cjelobrojni raspon godina, počevši od cjelobrojne dobi. Uobičajeno je u tablici smrtnosti tabeliranoj po cjelobrojnim dobima uz  $l_x$  prikazati i vrijednosti  $d_x$ , gdje je

$$d_x = l_x - l_{x+1}. \quad (2.4)$$

Te vrijednosti omogućuju nam računanje  $q_x$  – iz (2.4) imamo

$$d_x = l_x \left(1 - \frac{l_{x+1}}{l_x}\right) = l_x(1 - p_x) = l_x q_x.$$

Do ovog izraza također možemo doći ako interpretiramo  $d_x$  kao očekivani broj smrti između dobi  $x$  i  $x + 1$  od  $l_x$  života starosti točno  $x$ , pa koristeći ponovno binomnu distribuciju, imamo

$$d_x = l_x q_x. \quad (2.5)$$

**Primjer 2.1.** *Tablica 2.1 predstavlja isječak tablice smrtnosti. Izračunajmo*

(a)  $l_{40}$ ,

(b)  ${}_{10}p_{30}$ ,

(c)  $q_{35}$ ,

$x$	$l_x$	$d_x$
30	10000.00	34.78
31	9965.22	38.10
32	9927.12	41.76
33	9885.35	45.81
34	9839.55	50.26
35	9789.29	55.17
36	9734.12	60.56
37	9673.56	66.49
38	9607.07	72.99
39	9534.08	80.11

Tablica 2.1: Isječak tablice smrtnosti

(d)  ${}_5q_{30}$ ,

(e) vjerojatnost da osoba starosti 30 umre između navršenih 35 i 36 godina.

Rješenje. (a) Iz jednakosti (2.4),

$$l_{40} = l_{39} - d_{39} = 9453.97.$$

(b) Iz formule (2.1) imamo

$${}_{10}p_{30} = \frac{l_{40}}{l_{30}} = \frac{9453.97}{10000} = 0.94540.$$

(c) Iz formule (2.5) imamo

$$q_{35} = \frac{d_{35}}{l_{35}} = \frac{55.17}{9789.29} = 0.00564.$$

(d) Prema jednakosti (2.2) slijedi

$${}_5q_{30} = \frac{l_{30} - l_{35}}{l_{30}} = 0.02107.$$

(e) Ta vjerojatnost jednaka je  ${}_5|q_{30}$ , pa prema (2.3) slijedi

$${}_5q_{30} = \frac{l_{35} - l_{36}}{l_{30}} = \frac{d_{35}}{l_{30}} = 0.00552.$$

□

## 2.2 Pretpostavke o razlomljenoj dobi

Tablica smrtnosti  $\{l_x\}_{x \geq x_0}$  pruža točno iste informacije kao odgovarajuća distribucija doživljenja,  $S_{x_0}$ . Ipak, tablica smrtnosti tabelirana po cjelobrojnim dobima ne sadrži sve informacije iz pripadnog modela doživljenja, budući da vrijednosti  $l_x$  za cjelobrojne dobi  $x$  nisu dovoljne za izračunavanje vjerojatnosti koje sadrže necjelobrojne dobi, poput  ${}_{0.7}p_{30.5}$ . Uz dane vrijednosti  $l_x$  samo u cjelobrojnim dobima, potrebna nam je dodatna pretpostavka ili neke dodatne informacije kako bismo izračunali vjerojatnosti za necjelobrojne dobi. Specifičnije, potrebna su nam dodatne pretpostavke o vjerojatnosnoj distribuciji slučajne varijable budućeg životnog vijeka na intervalima između cjelobrojnih dobi. Koristimo termin **pretpostavka o razlomljenoj dobi** kako bismo opisali takve pretpostavke. U ovom potpoglavlju pretpostavljamo da je tablica smrtnosti dana samo za cjelobrojne dobi i opisujemo dvije najkorisnije pretpostavke o razlomljenoj dobi.

### Uniformna distribucija smrti

Pretpostavka o uniformnoj distribuciji smrti (eng. uniform distribution of deaths, odnosno UDD) najčešći je oblik pretpostavke o razlomljenoj dobi. Ona može biti formulirana na dva različita, ali ekvivalentna načina.

#### UDD1

Za cijeli broj  $x$  i za  $0 \leq s \leq 1$ , pretpostavimo

$${}_s q_x = s q_x. \quad (2.6)$$

#### UDD2

Podsjetimo se da je  $K_x$  najveće cijelo od  $T_x$  i definirajmo slučajnu varijablu kao razlomljeni dio od  $T_x$ , odnosno

$$T_x = K_x + R_x.$$

Pretpostavka UDD2 tada je da za cijeli broj  $x$  vrijedi  $R_x \sim U(0,1)$  te je  $R_x$  nezavisna od  $K_x$ .

Pokažimo ekvivalenciju tih dviju pretpostavki. Pretpostavimo najprije da vrijedi UDD1. Tada za cijeli  $x$  i  $0 \leq s < 1$  imamo



$$\begin{aligned}
\mathbb{P}(R_x \leq s) &= \sum_{k=0}^{\infty} \mathbb{P}(R_x \leq s, K_x = k) \\
&= \sum_{k=0}^{\infty} \mathbb{P}(k \leq T_x \leq k + s) \\
&= \sum_{k=0}^{\infty} {}_k p_x {}_s q_{x+k} \\
&\stackrel{\text{UDD1}}{=} \sum_{k=0}^{\infty} {}_k p_x s (q_{x+k}) \\
&= s \sum_{k=0}^{\infty} {}_k p_x q_{x+k} \\
&= s \sum_{k=0}^{\infty} \mathbb{P}(K_x = k) \\
&= s.
\end{aligned}$$

Ovime smo dokazali da vrijedi  $R_x \sim U(0,1)$ . Kako bismo dokazali nezavisnost  $R_x$  i  $K_x$ , primijetimo

$$\begin{aligned}
\mathbb{P}(R_x \leq s, K_x = k) &= \mathbb{P}(k \leq T_x \leq k + s) \\
&= {}_k p_x s (q_{x+k}) \\
&= s {}_k p_x q_{x+k} \\
&= \mathbb{P}(R_x \leq s) \mathbb{P}(K_x = k),
\end{aligned}$$

pri čemu zadnja jednakost slijedi iz  $R_x \sim U(0,1)$ . Ovime smo dokazali da UDD1 povlači UDD2.

Kako bismo dokazali obratnu implikaciju, pretpostavimo da vrijedi UDD2. Tada za cijeli  $x$  i  $0 \leq s < 1$  imamo

$$\begin{aligned}
{}_s q_x &= \mathbb{P}(T_x \leq s) \\
&= \mathbb{P}(K_x = 0, R_x \leq s) \\
&= \mathbb{P}(R_x \leq s) \mathbb{P}(K_x = 0),
\end{aligned}$$

gdje zadnja jednakost slijedi iz pretpostavke o nezavisnosti  $K_x$  i  $R_x$ . Dakle,

$${}_s q_x = s q_x. \quad (2.7)$$

Formulacija UDD2 objašnjava zašto se ova pretpostavka zove uniformna distribucija smrti, ali u praktičnim primjenama korisnija je formulacija UDD1.

Izravna posljedica je da za  $0 \leq s < 1$  vrijedi

$$l_{x+s} = l_x - s d_x. \quad (2.8)$$

To slijedi budući da je

$${}_s q_x = 1 - \frac{l_{x+s}}{l_x},$$

a uvrštavanje  $s q_x$  umjesto  ${}_s q_x$  daje

$$s \frac{d_x}{l_x} = \frac{l_x - l_{x+s}}{l_x},$$

odakle množenjem s  $l_x$  i izražavanjem  $l_{x+s}$  dobivamo upravo jednakost (2.8).

Dakle, pretpostavljamo da je  $l_{x+s}$  linearno opadajuća funkcija od  $s$ . Diferenciranjem jednakosti (2.6) u odnosu na  $s$ , dobivamo

$$\frac{d}{ds} {}_s q_x = q_x, \quad 0 \leq s \leq 1,$$

pri čemu znamo da lijeva strana predstavlja funkciju gustoće od  $T_x$  u točki  $s$ , s obzirom da diferenciramo funkciju distribucije. Iz (1.20) uostalom znamo da je funkcija gustoće za  $T_x$  u  $s$  jednaka  ${}_s p_x \mu_{x+s}$ , pa pod UDD za  $0 \leq s < 1$  vrijedi

$$q_x = {}_s p_x \mu_{x+s}. \quad (2.9)$$

Lijeva strana ne ovisi o  $s$ , što znači da je funkcija gustoće konstanta za  $0 \leq s < 1$ , što također slijedi iz pretpostavke o uniformnoj distribuciji od  $R_x$ . Budući da je  $q_x$  konstanta u odnosu na  $s$ , a  ${}_s p_x$  opadajuća funkcija od  $s$ , vidimo da je  $\mu_{x+s}$  rastuća funkcija od  $s$ , što je primjereno za sve dobi koje su od interesa osiguravateljima. Ipak, ako primijenimo aproksimaciju na uzastopne dobi, dobivamo funkciju intenziteta smrtnosti koja nije neprekidna, s prekidima u cjelobrojnim dobima, kao što je ilustrirano u primjeru 2.4. Iako je to nepoželjno, ne smatra se ozbiljnim nedostatkom.

**Primjer 2.2.** *Uz dano  $p_{40} = 0.999473$ , izračunajmo  ${}_{0.4}q_{40.2}$  pod pretpostavkom uniformne distribucije smrti.*

*Rješenje.* Primijetimo da fundamentalni rezultat u (2.7) zahtijeva da  $x$  bude cijeli broj. Zato ćemo vjerojatnost  ${}_{0.4}q_{40.2}$  zapisati na način da sadrži samo vjerojatnosti u cjelobrojnim dobima:

$$\begin{aligned}
 {}_{0.4}q_{40.2} &= 1 - {}_{0.4}p_{40.2} = 1 - \frac{l_{40.6}}{l_{40.2}} \\
 &= 1 - \frac{l_{40} - 0.6p_{40}}{l_{40} - 0.2p_{40}} = 1 - \frac{0.6p_{40}}{0.2p_{40}} \\
 &= 1 - \frac{1 - 0.6q_{40}}{1 - 0.2q_{40}} = 2.108 \times 10^{-4}.
 \end{aligned}$$

□

**Primjer 2.3.** *Iskoristimo istu tablicu smrtnosti kao u primjeru 2.1 i pretpostavku UDD kako bismo izračunali*

(a)  ${}_{1.7}q_{33}$ ,

(b)  ${}_{1.7}q_{33.5}$ .

*Rješenje.* (a) Primijetimo da je

$${}_{1.7}q_{33} = 1 - {}_{1.7}p_{33} = 1 - (p_{33})({}_{0.7}p_{34}).$$

$p_{33}$  možemo direktno izračunati iz tablice smrtnosti koristeći  $l_{34}/l_{33} = 0.995367$  i  ${}_{0.7}p_{34} = 1 - 0.7q_{34} = 0.996424$ , pa je  ${}_{1.7}q_{33} = 0.008192$ .

(b) Kako bismo izračunali  ${}_{1.7}q_{33.5}$  koristeći UDD, raspisujemo:

$$\begin{aligned}
 {}_{1.7}q_{33.5} &= 1 - {}_{1.7}p_{33.5} \\
 &= 1 - \frac{l_{35.2}}{l_{33.5}} \\
 &= 1 - \frac{l_{35} - 0.2d_{35}}{l_{33} - 0.5d_{33}} \\
 &= 0.008537.
 \end{aligned}$$

□

**Primjer 2.4.** *Pod pretpostavkom o uniformnoj distribuciji smrti, izračunajmo  $\lim_{t \rightarrow 1^-} \mu_{40+t}$  koristeći  $p_{40} = 0.999473$  te  $\lim_{t \rightarrow 0^+} \mu_{41+t}$  koristeći  $p_{41} = 0.999429$ .*

*Rješenje.* Iz formule (2.9) imamo  $\mu_{x+t} = q_x / {}_t p_x$ . Uvrštavanjem  $x = 40$  dobivamo

$$\lim_{t \rightarrow 1^-} \mu_{40+t} = q_{40} / p_{40} = 5.273 \times 10^{-4},$$

dok uvrštavanje  $x = 41$  daje

$$\lim_{t \rightarrow 0^+} \mu_{41+t} = q_{41} = 5.71 \times 10^{-4}.$$

□

**Primjer 2.5.** Uz dato  $q_{70} = 0.010413$ ,  $q_{71} = 0.011670$  i pod pretpostavkom UDD izračunajmo  ${}_{0.7}q_{70.6}$ .

*Rješenje.* Budući da su smrti uniformno distribuirane između dobi 70 i 71 te između dobi 71 i 72, imamo

$${}_{0.7}q_{70.6} = {}_{0.4}q_{70.6} + (1 - {}_{0.4}q_{70.6}){}_{0.3}q_{71}.$$

Koristeći isti argument kao u rješenju primjera 2.3, odnosno  $p_{70} = {}_{0.6}p_{70} {}_{0.4}p_{70.6}$ , imamo

$${}_{0.4}q_{70.6} = 1 - {}_{0.4}p_{70.6} = 1 - \frac{p_{70}}{{}_{0.6}p_{70}} = 1 - \frac{1 - q_{70}}{1 - 0.6q_{70}} = 4.191 \times 10^{-3},$$

a budući da je  ${}_{0.3}q_{71} = 0.3q_{71} = 3.501 \times 10^{-3}$ , dobivamo  ${}_{0.7}q_{70.6} = 7.678 \times 10^{-3}$ .  $\square$

## Konstantni intenzitet smrtnosti

Druga pretpostavka o razlomljenoj dobi je da je intenzitet smrtnosti konstantan između cjelobrojnih dobi. Dakle, za cjelobrojni  $x$  i  $0 \leq s < 1$ , pretpostavljamo da je  $\mu_{x+s}$  neovisno o  $s$  i obilježavamo ga s  $\mu_x^*$ . Vrijednost  $\mu_x^*$  možemo pronaći koristeći

$$p_x = \exp \left\{ - \int_0^1 \mu_{x+s} ds \right\}.$$

Dakle, pretpostavka da je za  $0 \leq s < 1$   $\mu_{x+s} = \mu_x^*$  daje  $p_x = e^{-\mu_x^*}$ , tj.  $\mu_x^* = -\log p_x$ . Nadalje, pod pretpostavkom konstantnog intenziteta smrtnosti za  $0 \leq s < 1$  imamo

$${}_s p_x = \exp \left\{ - \int_0^s \mu_x^* du \right\} = e^{-\mu_x^* s} = (p_x)^s.$$

Slično, za  $t, s > 0$  i  $t + s < 1$ ,

$${}_s p_{x+t} = \exp \left\{ - \int_0^s \mu_x^* du \right\} = (p_x)^s.$$

Dakle, pod pretpostavkom o konstantnom intenzitetu smrtnosti, vjerojatnost doživljenja perioda od  $s < 1$  godina u dobi  $x + t$  nezavisna je od  $t$ , pod uvjetom da je  $s + t < 1$ .

Pretpostavka o konstantnom intenzitetu smrtnosti vodi do step funkcije intenziteta smrtnosti na rasponu uzastopnih dobi života. Po svojoj prirodi ova pretpostavka daje konstantan intenzitet smrtnosti na intervalu od navršene dobi  $x$  do navršene dobi  $x + 1$ , premda bismo očekivali da intenzitet smrtnosti neprestano raste. Ipak, ako pravi intenzitet smrtnosti polako raste tijekom godine (između navršavanja uzastopnih dobi), pretpostavka o konstantnom intenzitetu smrtnosti je razumna.

**Primjer 2.6.** Uz dano  $p_{40} = 0.999473$ , izračunajmo  ${}_{0.4}q_{40.2}$  pod pretpostavkom konstantnog intenziteta smrtnosti.

*Rješenje.*  ${}_{0.4}q_{40.2} = 1 - {}_{0.4}p_{40.2} = 1 - (p_{40})^{0.4} = 2.108 \times 10^{-4}$  □

**Primjer 2.7.** Uz dano  $q_{70} = 0.010413$  i  $q_{71} = 0.011670$ , izračunajmo  ${}_{0.7}q_{70.6}$  pod pretpostavkom konstantnog intenziteta smrtnosti.

*Rješenje.* Kao u rješenju primjera 2.5, pišemo

$${}_{0.7}q_{70.6} = {}_{0.4}q_{70.6} + (1 - {}_{0.4}q_{70.6}){}_{0.3}q_{71},$$

gdje je  ${}_{0.4}q_{70.6} = 1 - (p_{70})^{0.4} = 4.178 \times 10^{-3}$ , a  ${}_{0.3}q_{71} = 1 - (p_{71})^{0.3} = 3.515 \times 10^{-3}$ , što daje  ${}_{0.7}q_{70.6} = 7.679 \times 10^{-3}$ . □

Primijetimo da smo u primjerima 2.2 i 2.6, odnosno primjerima 2.5 i 2.7 koristili po dvije različite metode za rješavanje istog problema, i rješenja se podudaraju do na osmu, odnosno do na petu decimalu, respektivno. Općenito vrijedi da pretpostavke o uniformnoj distribuciji smrti i konstantnom intenzitetu smrtnosti daju vrlo slične rezultate. Razlog tomu možemo vidjeti iz sljedećih aproksimacija. Pod pretpostavkom konstantnog intenziteta smrtnosti,

$$q_x = 1 - e^{-\mu^*} \approx \mu^*$$

pod uvjetom da je  $\mu^*$  malen i za  $0 < t < 1$ ,

$${}_tq_x = 1 - e^{-\mu^*t} \approx \mu^*t.$$

Drugim riječima, aproksimacija od  ${}_tq_x$  jednaka je aproksimaciji od  $q_x$  pomnoženoj s  $t$ , što je upravo ono što daje i pretpostavka o uniformnoj distribuciji smrti.

## 2.3 Nacionalne tablice smrtnosti

Tablice smrtnosti koje se temelje na iskustvu mortaliteta na cijeloj populaciji jedne države redovno se objavljuju u mnogim državama svijeta. Obično se izrađuju odvojene tablice za muškarce i žene, a ponekad i za neke druge grupe pojedinaca, npr. u ovisnosti o pušačkim navikama.

Tablica 2.2 prikazuje vrijednosti  $q_x \times 10^5$ , gdje je  $q_x$  vjerojatnost smrti unutar jedne godine za odabrane navršene dobi  $x$ , odvojeno za muškarce i žene, za populacije Australije, Engleske i Walesa te SAD-a. Ove tablice konstruirane su koristeći evidenciju smrti u pojedinoj dobi, ili mali broj uzastopnih godina i procjenu populacije u sredini tog perioda. Relevantno razdoblje istaknuto je u naslovu svakog od tri

x	Australaska tablica smrtnosti 2000.-2002.		Engleska tablica smrtnosti 1990.-1992.		Američka tablica smrtnosti 2002.	
	Muškarci	Žene	Muškarci	Žene	Muškarci	Žene
0	567	466	814	632	764	627
1	44	43	62	55	53	42
2	31	19	38	30	37	28
10	13	8	18	13	18	13
20	96	36	84	31	139	45
30	119	45	91	43	141	63
40	159	88	172	107	266	149
50	315	202	464	294	570	319
60	848	510	1392	830	1210	758
70	2337	1308	3930	2190	2922	1899
80	6399	4036	9616	5961	7028	4930
90	15934	12579	20465	15550	16805	13328
100	24479	23863	38705	32489	-	-

Tablica 2.2: Vrijednosti  $q_x \times 10^5$  iz nekih nacionalnih tablica smrtnosti

stupca. Podaci za najstarije dobi iznimno su nepouzdana. Iz tog razloga, tablica za SAD ne prikazuje vrijednosti  $q_x$  za dob 100 i starije.

Za sve tri nacionalne tablice te neovisno o spolu, vrijednosti  $q_x$  prate isti uzorak kao funkcija dobi,  $x$ . Iz tablice 2.2 opažamo sljedeće činjenice:

- Vrijednost  $q_0$  relativno je visoka. Stopa smrtnosti po rođenju, tzv. perinatalna smrtnost, visoka je zbog komplikacija koje mogu proizaći iz kasnijih stupnjeva trudnoće i samog procesa poroda. Vrijednost  $q_x$  ne postiže ponovno tako visoku razinu sve do dobi od otprilike 55 godina (vidi i [1, str. 51., slika 3.1])
- Stopa smrtnosti mnogo je niža nakon navršene prve godine – manja je od 10% svog iznosa u prvoj godini, te pada do dobi oko 10 godina.
- Na spomenutoj slici [1, str. 51., slika 3.1] vidimo da muška i ženska smrtnost značajno divergiraju u kasnim tinejdžerskim godinama, pri čemu muška smrtnost strmije raste. Ovakva pojava nije prisutna samo među mladim odraslim muškarcima iz različitih populacija diljem svijeta, nego je i značajka povijesnih populacija u državama poput Ujedinjenog Kraljevstva, gdje su se podaci o smrtnosti prikupljali kroz poprilično mnogo vremena. Ovaj fenomen ponekad

se naziva "izbočina nesreća" (eng. "accident hump"), budući da mnogo smrti u toj dobi kod muškaraca nastupi uslijed nesreće.

- Stopa smrtnosti raste od dobi 10, pri čemu je zbog spomenute "izbočine nesreća" rast relativno velik između dobi 10 i 20 za muškarce, nakon čega slijedi skromniji rast između dobi 20 i 40, te postojani rast nakon dobi 40.
- Za svaku dob, svih šest vrijednosti  $q_x$  usporedive su, pri čemu je za svaku državu stopa smrtnosti za žene u pojedinoj dobi uvijek malo manja nego za muškarce.
- Gompertzov model predstavljen u prvom poglavlju relativno je jednostavan, s obzirom da zahtijeva samo dva parametra te je u sklopu njega intenzitet smrtnosti jednostavnog funkcijskog oblika,  $\mu_x = Bc^x$ . U prvom smo poglavlju ustanovili da ovaj model ne pristaje dobro u svim dobima. Iz slike [1, str. 51., slika 3.1] jasno je da ovaj model ne može dobro opisati perinatalnu smrtnost, niti "izbočinu nesreća". Ipak, stope smrtnosti u kasnijim dobima bolje se ponašaju te se Gompertzov model često pokazuje korisnim u starijim dobima. Primjerice, na slici [1, str. 52., slika 3.2] možemo vidjeti krivulju stope smrtnosti muškaraca starije dobi u SAD-u 2002. godine zajedno s Gompertzovom krivuljom prilagođenom tablici smrtnosti za isto razdoblje. Možemo uočiti kako Gompertzova krivulja poprilično dobro prijanja stvarnim podacima, što je posebno impresivno uzmemo li u obzir da je Gompertz ovaj model predložio 1825.

Zadnja opaska vezana uz tablicu 2.2 je da uočimo da smo uspoređivali tri nacionalne tablice smrtnosti koristeći vjerojatnosti umiranja unutar jedne godine,  $q_x$ , a ne intenzitet smrtnosti,  $\mu_x$ . To je iz razloga što se vrijednosti  $\mu_x$  ne objavljuju ni za koju dob u američkim tablicama smrtnosti. Također, vrijednosti  $\mu_x$  ne objavljuju se za dob 0 u drugim dvjema nacionalnim tablicama – postoje tehnički problemi u procjeni  $\mu_x$  unutar godine u kojoj se intenzitet smrtnosti rapidno mijenja, kao što je to između dobi 0 i 1.

## 2.4 Dodatne opaske vezane uz tablice smrtnosti

U ovom poglavlju najprije ćemo sažeto ponoviti neke poznate rezultate. Pretpostavimo da poznamo vjerojatnosti doživljenja  ${}_t p_x$  (a priori ili u danom modelu koji je opisan s  $F_x(t), S_x(t), \mu_x$ ). Neka je za  $x_0 \leq x \leq \omega$   $\{l_x\}$  tablica smrtnosti. Kao što iz zadanih veličina  $l_x$  možemo naći vjerojatnost doživljenja ((2.1)), tako iz  $l_x$  možemo naći i vjerojatnost smrti. Ovdje dajemo poopćenja jednadžbi (2.2) i (2.3):

$$q_x = 1 - p_x = 1 - \frac{l_{x+1}}{l_x} = \frac{l_x - l_{x+1}}{l_x}, \quad {}_tq_x = 1 - {}_tp_x = \frac{l_x - l_{x+t}}{l_x}, \quad (2.10)$$

$${}_u p_x - {}_{u+t} p_x = \frac{l_{x+u}}{l_x} - \frac{l_{x+u+t}}{l_x} = \frac{l_{x+u} - l_{x+u+t}}{l_x}. \quad (2.11)$$

Najčešće je početna dob  $x_0 = 0$ , a tablice smrtnosti sadrže samo  $l_x$  za cjelobrojne  $x$ , kao što smo ranije spomenuli. U Hrvatskoj se uzima  $l_0 = 100000$ ; u LAT A1967-70 (u daljnjem LAT) je  $l_0 = 34489$ . (Tablica LAT dolazi kao prilog knjizi [2].) Međutim, to je nevažno jer ionako su u svim računima relevantni odgovarajući kvocijenti. Uočimo da je u LAT  $l_{109} = 0.12332121$  i podrazumijeva se da je  $l_{110} = 0$ ; dakle  $\omega = 110$ .

Što se tiče veličine  $d_x$ , uz formule (2.4) i (2.5) uočimo da je  $l_\omega = 0$  pa je  $d_{\omega-1} = l_{\omega-1}$ . Također,

$$l_x = \sum_{k=x}^{\omega-1} d_k \quad (2.12)$$

Naime,  $\sum_{k=x}^{\omega-1} d_k = \sum_{k=x}^{\omega-1} (l_k - l_{k+1}) = l_x - l_\omega = l_x$ . Još je očitiji zdravorazumski dokaz: imamo njih  $l_x$  u dobi  $x$ . Svi moraju umrijeti: netko u  $[x, x + 1 >$ , netko u  $[x + 1, x + 2 > \dots$ . Na ovaj način, polazeći od  $T_x, F_x(t), S_x(t), f_x(t), \mu_x, {}_t p_x, {}_t q_x$  (a u stvari  ${}_t p_0$ , tj.  $S_0(t)$  je dovoljno), izgradimo tablice koje zorno opisuju smrtnost naše populacije, odnosno modela. Problem je međutim u tome što mi u praksi nemamo ništa od toga – barem ne a priori. U praksi imamo empirijske podatke i iz njih izvedemo veličine  $p_x$  i  $q_x$ , konstruiramo tablice, a nakon toga sve ostalo – u obrnutom poretku. U stvari, tražimo teorijski model. Prve tablice smrtnosti konstruirao je Sir Edmund Halley 1693. Kod nas se one objavljuju kao službeni akt u Narodnim novinama (vidi [4]). Mogu i ne moraju biti odvojene po spolu. Mi ćemo u našim primjerima koristiti LAT.

Dakle neki popis stanovništva ili drugi empirijski izvor pomažu nam da utvrdimo, odnosno procijenimo veličine  $q_x$ . Sad  $l_0$  utvrdimo proizvoljno, a onda je  $d_0 = l_0 q_0$ ,  $l_1 = l_0 - d_0$  i induktivno dalje:  $d_x = l_x q_x$ ,  $l_{x+1} = l_x - d_x, \forall x = 1, 2, \dots, \omega - 1$ . Nakon toga stavi se  ${}_t q_x = \frac{l_x - l_{x+t}}{l_x}$ ,  ${}_t p_x = \frac{l_{x+t}}{l_x}$ . Time smo rekonstruirali, odnosno reetablirali formule (2.1)-(2.5) te (2.12). Sad dalje modeliramo na sljedeći način:

## Glatka interpolacija – Karup-King

Najprije interpoliramo: od čvorova  $l_0, l_1, l_2, \dots$  proizvedemo neprekidnu, čak i derivabilnu funkciju  $l(x) = l_x, x \in [0, \infty]$ . To se može, ali ipak ne uz Lagrangeov interpolacijski polinom budući da on neće dati monotonu funkciju, što  $l_x$  po prirodi stvari mora biti.



Sada stavimo

$$\mu_x = -\frac{1}{l_x} \frac{d}{dx} l_x = -(\ln(l_x))', x > 0 \quad (2.13)$$

Sad je bitno provjeriti da smo konzistentni. Naime, ako smo veličine  $l_x$  izračunali iz unaprijed danog modela u kojem smo imali  $F_x(t)$ ,  $S_x(t)$  i  $\mu_x$ , onda moramo vidjeti da je formula (2.13) suglasna s ranijom definicijom intenziteta smrtnosti  $\mu_x$  iz jednakosti (1.9):  $\mu_x = \frac{-1}{S_0(x)} \frac{d}{dx} S_0(x)$ . Vidimo da (2.1) povlači  $l_x = l_0 p_0$ , a kako je iz (1.12)  ${}_t p_0 = S_0(t)$ , slijedi  $-\frac{1}{l_x} \frac{d}{dx} l_x = -\frac{1}{l_0 S_0(x)} \frac{d}{dx} (l_0 S_0(x)) = -\frac{1}{S_0(x)} \frac{d}{dx} S_0(x)$ . Dalje je lagano. Sad kad imamo  $\mu_x$ , onda najprije imamo  $\int_x^{\omega(\infty)} l_t \mu_t dt = \int_x^{\omega(\infty)} -l'_t dt = -l_\omega + l_x = l_x$ ; dakle

$$l_x = \int_x^{\infty} l_t \mu_t dt, \quad (2.14)$$

a odavde

$$d_x = l_x - l_{x+1} = \int_x^{x+1} l_t \mu_t dt. \quad (2.15)$$

Dalje,

$$\begin{aligned} \int_0^x \mu_t dt &= -\int_0^x \frac{d}{dt} (\ln(l(t))) dt = -(\ln(l(x)) - \ln(l_0)) = -\ln \frac{l_x}{l_0} \\ \implies l_x &= l_0 \exp \left( -\int_0^x \mu_t dt \right). \end{aligned} \quad (2.16)$$

Dalje, (2.16) za  $x+n$  umjesto  $x$  daje  $l_{x+n} = l_0 \exp \left( -\int_0^{x+n} \mu_t dt \right)$ . Dijeljenjem s (2.16) dobivamo

$$\begin{aligned} \frac{l_{x+n}}{l_x} &= \exp \left( -\int_0^{x+n} \mu_t dt + \int_0^x \mu_t dt \right) \\ &\stackrel{(2.1)}{\implies} {}_n p_x = \exp \left( -\int_x^{x+n} \mu_t dt \right). \end{aligned} \quad (2.17)$$

Primijetimo da i (2.16) i (2.17) vrijede i za necjelobrojne  $x$  i  $n$ . Dalje,

$$\begin{aligned} \int_x^{x+n} l_t \mu_t dt &= -\int_x^{x+n} l'_t dt = l_x - l_{x+n} \\ \implies {}_n q_x &= \frac{l_x - l_{x+n}}{l_x} = \frac{1}{l_x} \int_x^{x+n} l_t \mu_t dt \\ &= \int_x^{x+n} \frac{l_t}{l_x} \mu_t dt = \int_x^{x+n} {}_{t-x} p_x \mu_t dt. \end{aligned}$$

Nakon supstitucije  $y = t - x$  dobivamo

$${}_nq_x = \int_0^n {}_y p_x \mu_{x+y} dy. \quad (2.18)$$

Posebno, za  $n = 1$  dobiva se

$$q_x = \int_0^1 {}_y p_x \mu_{x+y} dy. \quad (2.19)$$

S obzirom da sve ovisi (i može biti izraženo pomoću)  $\mu_x$ , bilo bi važno tu funkciju zaista i znati. Međutim, ono čime raspolažemo samo su vrijednosti  $l_x$  za cjelobrojne  $x$ , te  $p_x, q_x, {}_t p_x, {}_t q_x$  za  $x, t \in \mathbb{N}$ . Interpolaciju za  $l_x$  ne znamo eksplicitno provesti, a pogotovo se ne možemo nadati nekom analitičkom zapisu funkcije  $\mu_x$ . Zato pribjegavamo različitim interpolacijama, odnosno procjenama.

(A) Iz (2.17) znamo da je  ${}_n p_x = \exp\left(-\int_x^{x+n} \mu_t dt\right)$ . Posebno, za  $n = 1$  imamo  $\ln p_x = -\int_x^{x+1} \mu_t dt \approx -\mu_{x+\frac{1}{2}}$  ako smo uzeli  $\mu_t \approx \mu_{x+\frac{1}{2}}, \forall t \in [x, x+1]$ . Slično,  $-\frac{1}{2}(\ln p_x + \ln p_{x-1}) = \frac{1}{2}\left(\int_x^{x+1} \mu_t dt + \int_{x-1}^x \mu_t dt\right) = \frac{1}{2}\left(\int_{x-1}^{x+1} \mu_t dt\right) \approx \frac{1}{2} \cdot 2 \cdot \mu_x = \mu_x$  ako smo uzeli  $\mu_t \approx \mu_x, \forall t \in [x-1, x+1]$ . Dakle,

$$\mu_x \approx -\frac{1}{2}(\ln p_x + \ln p_{x-1}). \quad (2.20)$$

(B) Ovdje pretpostavimo da je  $l_x$  kvadratna funkcija na  $[x-1, x+1]$ . Tada je

$$\mu_x \approx \frac{l_{x-1} - l_{x+1}}{2l_x}. \quad (2.21)$$

Naime,  $\mu_x = -\frac{1}{l_x}(l_x)'$ . A sad se lako vidi: ako je  $f(x) = \alpha x^2 + \beta x + \gamma$  na segmentu  $x \in [x_0 - h, x_0 + h]$ , onda je  $f'(x_0) = \frac{f(x_0+h) - f(x_0-h)}{2h}$ . Sad se to primijeni na  $f(x) = l_x, x \in [x_0 - 1, x_0 + 1]$ .

(C) Uzmimo da je  $l_x$  dovoljno puta derivabilna. Taylorov razvoj oko  $x$  daje (pišemo znak jednakosti umjesto aproksimacije zanemarujući repove)  $l_{x+h} = l_x + hl'_x + \frac{h^2}{2!}l''_x + \frac{h^3}{3!}l'''_x + \frac{h^4}{4!}l^{(iv)}_x$ . Odavde je

$$\begin{aligned} l_{x-2} &= l_x - 2l'_x + 2l''_x - \frac{8}{6}l'''_x + \frac{16}{24}l^{(iv)}_x, \\ l_{x+2} &= l_x + 2l'_x + 2l''_x + \frac{8}{6}l'''_x + \frac{16}{24}l^{(iv)}_x, \\ l_{x-1} &= l_x - l'_x + \frac{1}{2}l''_x - \frac{1}{6}l'''_x + \frac{1}{24}l^{(iv)}_x, \\ l_{x+1} &= l_x + l'_x + \frac{1}{2}l''_x + \frac{1}{6}l'''_x + \frac{1}{24}l^{(iv)}_x. \end{aligned}$$

Od prve jednakosti oduzmemo drugu, od treće četvrtu pa dobivamo  $l_{x-2} - l_{x+2} = -(4l'_x + \frac{8}{3}l'''_x)$ ,  $l_{x-1} - l_{x+1} = -(2l'_x + \frac{1}{3}l'''_x)$ . Odavde je  $l_{x-2} - l_{x+2} - 8(l_{x-1} - l_{x+1}) = 12l'_x$ . Sad  $\mu_x = -\frac{l'_x}{l_x}$  daje

$$\mu_x \approx \frac{1}{12l_x}(8(l_{x-1} - l_{x+1}) - (l_{x-2} - l_{x+2})). \quad (2.22)$$

## 2.5 Karupov interpolacijski polinom

Neka je dan konačan niz točaka  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), n \in \mathbb{N}$ , na grafu nepoznate funkcije  $f$ ; dakle vrijedi  $f(x_i) = y_i, i = 0, 1, \dots, n$ . Općenito, problem interpolacije je problem nalaženja funkcije  $g$  definirane na segmentu  $[x_0, x_n]$ , ili šire, takve da vrijedi  $g(x_i) = y_i, i = 0, 1, \dots, n$ . (Ovdje i u daljnjem pretpostavljamo da je  $x_0 < x_1 < \dots < x_n$ .) Posebno, želimo da funkcija  $g$  ima čim više dobrih svojstava. Uzmemo li, npr. funkciju  $g$  kao Lagrangeov interpolacijski polinom,  $g$  će biti ne samo neprekidna, nego i beskonačno puta derivabilna. Problem je, međutim, u tome što vrijednosti  $g'(x_i)$  nisu ni u kakvoj vezi s  $f'(x_i)$ . Ako pak želimo da se u točkama  $x_i$  podudaraju i derivacije funkcija  $f$  i  $g$  (možda ne samo prve derivacije, nego eventualno i neke derivacije višeg reda), možemo se poslužiti Hermiteovim interpolacijskim polinomom. No, tu se javljaju dva problema. Prvo, Hermiteov polinom je polinom visokog stupnja. Drugo, i važnije, mi možda i želimo neku kontrolu nad derivacijama funkcije  $g$ , ali najčešće ne poznajemo derivacije funkcije  $f$  pa Hermiteova metoda nije upotrebljiva.

Karupova metoda kao rezultat daje **derivabilnu** funkciju  $g$  koja je po dijelovima (na svakom pojedinom segmentu  $[x_i, x_{i+1}]$ ) jednaka restrikciji nekog polinoma 3. stupnja. U čemu je prednost u odnosu na Lagrangeov interpolacijski polinom? (Sjetimo se da je on beskonačno puta derivabilan.) Prije svega u kontroli stupnja; Lagrangeov polinom bit će stupnja  $n - 1$ , dok Karupova metoda operira s polinomima stupnja 3, ma kako  $n$  bio velik. Drugo, ukoliko imamo beskonačan niz točaka  $(x_0, y_0), (x_1, y_1), (x_2, y_2), \dots$  Lagrangeova metoda više nije primjenjiva. Doduše, možemo konstruirati Lagrangeove interpolacijske polinome za svakih  $n$  točaka i pomičući "prozor" (tj. najprije za  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ , potom  $(x_{n+1}, y_{n+1}), (x_{n+2}, y_{n+2}), \dots, (x_{2n}, y_{2n})$  itd.) i sve to "zalijepiti", ali sada gubimo derivabilnost u točkama lijepljenja. S druge strane, Karupova metoda primjenjuje se induktivno, bez razlike, tj. neovisno o tome je li niz čvorova konačan ili beskonačan.

U daljnjem ćemo pretpostaviti da je dan niz  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), n \in \mathbb{N}$ , pri čemu je  $x_{i+1} - x_i = 1, \forall i = 0, 1, \dots, n - 1$ . (Slučajevi kad čvorovi nisu ekvidistantni ili kad je  $x_{i+1} - x_i = \text{const.} \neq 1$  rješavaju se odgovarajućom zamjenom varijabli.)

Ideja: Pogledamo 4 uzastopne točke:  $(x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2})$ , a interpoliramo samo na segmentu  $[x_i, x_{i+1}]$ . Lagrangeovom metodom interpolirali bismo polinomom 1. stupnja, a Karupovom metodom polinomom 3. stupnja:  $y(h) = a + bh + ch^2 + dh^3$ ,  $h = x - x_i \in [0, 1]$ , pri čemu koeficijente  $a, b, c, d$  tek trebamo odrediti. Uočimo da za  $h = 0$  i  $h = 1$  dobivamo  $x = x_i$  i  $x = x_{i+1}$ , respektivno. Sad uvjeti  $y(0) = y_i, y(1) = y_{i+1}$  daju jednakosti

$$(1) \quad a = y_i.$$

$$(2) \quad a + b + c + d = y_{i+1}.$$

Nadalje, zahtijevat ćemo da derivacije u rubnim točkama, dakle  $y'(0)$  i  $y'(1)$  poprimaju unaprijed zadane vrijednosti koje ćemo odrediti na sljedeći način.

Pogledajmo prvo lijevi kraj, tj. točku  $x_i$ . Pogledajmo Lagrangeov interpolacijski polinom  $L_i^{(3)}$  koji prolazi točkama  $(x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1})$ . Uočimo da je naš lijevi kraj  $x_i$  ovdje središnja točka. Sad je plan da postignemo

$$(3) \quad y'(0) = (L_i^{(3)})'(x_i).$$

Dakle, želimo da derivacija funkcije  $y$  u lijevom kraju segmenta s kojim radimo (to je  $[x_i, x_{i+1}]$ ) bude jednaka derivaciji Lagrangeovog polinoma u prozoru duljine 2 kojemu je  $x_i$  središnja točka. Potpuno analogno postupimo u desnom kraju  $x_{i+1}$ . Pogledamo točke  $(x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2})$ , pogledamo pripadajući Lagrangeov polinom  $L_{i+1}^{(3)}$  i zahtijevamo

$$(4) \quad y'(1) = (L_{i+1}^{(3)})'(x_{i+1}).$$

### Lema 2.1.

$$L_i^{(3)}(x) = \frac{1}{2}(x - x_i)(x - x_{i+1})y_{i-1} - (x - x_{i-1})(x - x_{i+1})y_i + \frac{1}{2}(x - x_{i-1})(x - x_i)y_{i+1}$$

*Dokaz.* Očito vrijedi  $L_i^{(3)}(x_{i-1}) = y_{i-1}, L_i^{(3)}(x_i) = y_i, L_i^{(3)}(x_{i+1}) = y_{i+1}$  (imajmo na umu da je  $x_{i+1} - x_i = x_i - x_{i-1} = 1$ ).  $\square$

### Lema 2.2.

$$(L_i^{(3)})'(x_i) = \frac{1}{2}(y_{i+1} - y_{i-1})$$

*Dokaz.*

$$\begin{aligned} (L_i^{(3)})'(x) &= \frac{1}{2}(x - x_{i+1})y_{i-1} + \frac{1}{2}(x - x_i)y_{i-1} \\ &\quad - (x - x_{i+1})y_i - (x - x_{i-1})y_i \\ &\quad + \frac{1}{2}(x - x_i)y_{i+1} + \frac{1}{2}(x - x_{i-1})y_{i+1}, \end{aligned}$$

a sad je uvrštavanjem  $x_i$  tvrdnja očita. □

**Korolar 2.1.**

$$(L_{i+1}^{(3)})'(x_{i+1}) = (y_{i+2} - y_i)$$

Sad uvjet (3) i Lema 2.2 daju

$$(5) \quad b = \frac{1}{2}(y_{i+1} - y_{i-1}),$$

a iz uvjeta (4) i Korolara 2.1 dobivamo

$$(6) \quad b + 2c + 3d = \frac{1}{2}(y_{i+2} - y_i)$$

Sustav jednadžbi (1),(2),(5),(6) lako je riješiti. Dobiva se

$$\begin{aligned} a &= y_i, \\ b &= \frac{1}{2}(y_{i+1} - y_{i-1}), \\ c &= \frac{1}{2}(-y_{i+2} + 4y_{i+1} - 5y_i + 2y_{i-1}), \\ d &= \frac{1}{2}(y_{i+2} - 3y_{i+1} + 3y_i - y_{i-1}) \end{aligned}$$

Time smo dobili funkciju  $y(h)$ . Lako se vidi da  $y(h)$  možemo pisati u obliku (ako sredimo izlučujući zadane veličine  $y_{i-1}, y_i, y_{i+1}, y_{i+2}$ )

$$y(h) = -\frac{1}{2}h(h-1)^2y_{i-1} + \frac{1}{2}(h-1)(3h^2-2h-2)y_i - \frac{1}{2}h(3h^2-4h-1)y_{i+1} + \frac{1}{2}h^2(h-1)y_{i+2}$$

Na ovaj način završili smo interpolaciju na segmentu  $[x_i, x_{i+1}]$ . Dobivenu funkciju na tom segmentu (to je gornji  $y(h)$ ) označimo sada s  $g_i$ .

Sad se premjestimo jedan korak udesno. Interpoliramo na segmentu  $[x_{i+1}, x_{i+2}]$ , a u konstrukciji koristimo točke  $(x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2}), (x_{i+3}, y_{i+3})$ . Uočimo da su ovdje relevantni Lagrangeovi polinomi  $L_{i+1}^{(3)}$  te  $L_{i+2}^{(3)}$ . Dobivenu funkciju na tom segmentu označimo s  $g_{i+1}$ . Uočimo da je po konstrukciji  $g'_{i+1}(x_{i+1}) = (L_{i+1}^{(3)})'(x_{i+1})$ ; dakle, dobili smo  $g'_{i+1}(x_{i+1}) = g'_i(x_{i+1})$ .

Zaključak: ako zalijepimo  $g_i$  s  $g_{i+1}$ , dobit ćemo funkciju na segmentu  $[x_i, x_{i+2}]$  koja je derivabilna i u spojnoj točki  $x_{i+1}$  (jasno, derivabilnost u ostalim točkama uopće nije sporna).

U primjeni krećemo od segmenta  $[x_0, x_1]$  na kojem brinemo samo o derivaciji na desnom kraju i onda se korak po korak pomičemo udesno, kako je prethodno opisano.

## 2.6 Spragueov interpolacijski polinom

Spragueova metoda interpolacije je poopćenje Karupove metode i koristi se također zbog svojih prednosti u odnosu na Lagrangeovu metodu (jednake derivacije u točkama interpolacije, tj. neprekidnost derivacije) i Hermiteovu metodu (ne moraju eksplicitno biti poznate derivacije u danim točkama). Obje metode oslanjaju se na aproksimaciju derivacije u danim točkama na temelju Lagrangeovih interpolacijskih polinoma. Između ostalog, ove metode koriste se u aktuarskoj matematici.

Kao i kod Karupove metode, pretpostavit ćemo da su točke interpolacije ekvidistantne s razmakom 1 te koristimo iste oznake. Ovdje za interpolaciju na segmentu  $[x_i, x_{i+1}]$  umjesto polinoma trećeg, konstruiramo polinom petog stupnja, čije su prve i druge derivacije u rubnim točkama jednake odgovarajućim derivacijama Lagrangeovih polinoma  $L_i^{(5)}$ , odnosno  $L_{i+1}^{(5)}$ . Da bismo osigurali neprekidnost derivacija prvog i drugog reda aproksimacije na segmentu  $[x_i, x_{i+1}]$ , gledamo prethodne dvije i naredne dvije točke, tj. promatramo 6 uzastopnih točaka:  $(x_{i-2}, y_{i-2}), (x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2}), (x_{i+3}, y_{i+3})$ . Stavimo  $y(h) = a + bh + ch^2 + dh^3 + eh^4 + fh^5$ , gdje je  $h = x - x_i \in [0, 1]$ , kao kod Karupa. Također, za  $h = 0$  i  $h = 1$  dobivamo  $x = x_i$ , tj.  $x = x_{i+1}$ , odnosno interpolacijski polinom mora prolaziti kroz točke  $(x_i, y_i)$  i  $(x_{i+1}, y_{i+1})$ . Zapišimo derivacije aproksimacije:

$$\begin{aligned} y'(h) &= b + 2ch + 3dh^2 + 4eh^3 + 5fh^4, \\ y''(h) &= 2c + 6dh + 12eh^2 + 20fh^3. \end{aligned}$$

Analogno kao u Lemi 2.2 i Korolaru 2.1 može se pokazati

$$\begin{aligned} (L_i^{(5)})'(x_i) &= \frac{1}{12}(y_{i-2} - 8y_{i-1} + 8y_{i+1} - y_{i+2}) \\ (L_{i+1}^{(5)})'(x_{i+1}) &= \frac{1}{12}(y_{i-1} - 8y_i + 8y_{i+2} - y_{i+3}), \end{aligned}$$

odnosno za drugu derivaciju

$$\begin{aligned} (L_i^{(5)})''(x_i) &= \frac{1}{12}(-y_{i-2} + 16y_{i-1} - 30y_i + 16y_{i+1} - y_{i+2}) \\ (L_{i+1}^{(5)})''(x_{i+1}) &= \frac{1}{12}(-y_{i-1} + 16y_i - 30y_{i+1} + 16y_{i+2} - y_{i+3}). \end{aligned}$$

Dakle, imamo sustav od 6 jednadžbi:

$$\begin{aligned}
 y(0) &= a = y_i \\
 y(1) &= a + b + c + d + e + f = y_{i+1} \\
 y'(0) &= b = \frac{1}{12}(y_{i-2} - 8y_{i-1} + 8y_{i+1} - y_{i+2}) \\
 y'(1) &= b + 2c + 3d + 4e + 5f = \frac{1}{12}(y_{i-1} - 8y_i + 8y_{i+2} - y_{i+3}) \\
 y''(0) &= 2c = \frac{1}{12}(-y_{i-2} + 16y_{i-1} - 30y_i + 16y_{i+1} - y_{i+2}) \\
 y''(1) &= 2c + 6d + 12e + 20f = \frac{1}{12}(-y_{i-1} + 16y_i - 30y_{i+1} + 16y_{i+2} - y_{i+3})
 \end{aligned}$$

Rješenje tog sustava je:

$$\begin{aligned}
 a &= y_i \\
 b &= \frac{1}{12}(y_{i-2} - 8y_{i-1} + 8y_{i+1} - y_{i+2}) \\
 c &= \frac{1}{24}(-y_{i-2} + 16y_{i-1} - 30y_i + 16y_{i+1} - y_{i+2}) \\
 d &= \frac{1}{24}(-9y_{i-2} + 39y_{i-1} - 70y_i + 66y_{i+1} - 33y_{i+2} + 7y_{i+3}) \\
 e &= \frac{1}{24}(13y_{i-2} - 64y_{i-1} + 126y_i - 124y_{i+1} + 61y_{i+2} - 12y_{i+3}) \\
 f &= \frac{1}{24}(-5y_{i-2} + 25y_{i-1} - 50y_i + 50y_{i+1} - 25y_{i+2} + 5y_{i+3})
 \end{aligned}$$

Sređivanjem dobivamo izraz za  $y(h)$  preko 6 zadanih veličina  $y_{i-2}, \dots, y_{i+3}$ :

$$\begin{aligned}
 y(h) &= -\frac{1}{24}h(h-1)^3(5h+2)y_{i-2} + \frac{1}{24}h(h-1)(25h^3 - 39h^2 + 16)y_{i-1} \\
 &\quad - \frac{1}{12}(h-1)(25h^4 - 38h^3 - 3h^2 + 12h + 12)y_i + \frac{1}{12}h(25h^4 - 62h^3 + 33h^2 + 8h + 8)y_{i+1} \\
 &\quad - \frac{1}{24}h(h-1)(25h^3 - 36h^2 - 3h - 2)y_{i+2} + \frac{1}{24}h^3(h-1)(5h-7)y_{i+3}
 \end{aligned}$$

Ovime smo demonstrirali postupak nalaženja interpolacije na segmentu  $[x_i, x_{i+1}]$ . Ostatak procesa analogan je kao kod Karupove interpolacije.

## Poglavlje 3

# Primjena rezultata uz Karupovu interpolaciju

U ovome poglavlju, koristeći podatke za muškarce iz [4] izvest ćemo konkretnu formulu za  $l_x$  u svakoj dobi, pri čemu ćemo provesti Karupovu interpolaciju.

### 3.1 Računanje koeficijenata

Neka su dane vrijednosti  $l_n$  za  $n = 0, 1, 2, \dots, 100, 101$ , pri čemu je  $l_0 = 100000$ ,  $l_{101} = 0$ . Tražimo funkciju  $k$  na segmentu  $[0, 101]$  tako da  $k$  bude derivabilna na  $(0, 101)$  te da vrijedi  $k(n) = l_n, \forall n$  i  $k'(t) \leq 0, \forall t \in (0, 101)$ . Ovdje ćemo radi potpunosti na jednom mjestu ponoviti neke rezultate iz prethodnog poglavlja, no s drukčijim oznakama. Stavimo  $k_i = k|_{[i, i+1]}, i = 0, 1, 2, \dots, 100$ . Svaki  $k_i$  tražimo u obliku  $k_i(i+h) = f_i(h), f_i(h) = a_i + b_i h + c_i h^2 + d_i h^3$  ( $i = 1, 2, \dots, 99$ ) uz uvjete

$$f_i(0) = l_i, f_i(1) = l_{i+1}, f_i'(0) = \frac{1}{2}(l_{i+1} - l_{i-1}), f_i'(1) = \frac{1}{2}(l_{i+2} - l_i) \quad (3.1)$$

Uvjeti (3.1) daju jedinstveno određene  $a_i, b_i, c_i, d_i$  jednadžbama

$$\begin{aligned} a_i &= l_i \\ b_i &= \frac{1}{2}(l_{i+1} - l_{i-1}) \\ a_i + b_i + c_i + d_i &= l_{i+1} \\ b_i + 2c_i + 3d_i &= \frac{1}{2}(l_{i+2} - l_i) \end{aligned} \quad (3.2)$$



Jedinstveno rješenje sustava (3.2) je

$$\begin{aligned} a_i &= l_i, \\ b_i &= \frac{1}{2}(l_{i+1} - l_{i-1}), \\ c_i &= \frac{1}{2}(-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1}), \\ d_i &= \frac{1}{2}(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1}), \end{aligned} \quad (3.3)$$

tako da dobivamo

$$\begin{aligned} f_i(h) &= l_i + \frac{1}{2}(l_{i+1} - l_{i-1})h + \frac{1}{2}(-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1})h^2 \\ &\quad + \frac{1}{2}(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1})h^3, \end{aligned} \quad (3.4)$$

$$\begin{aligned} f'_i(h) &= \frac{1}{2}(l_{i+1} - l_{i-1}) + (-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1})h \\ &\quad + \frac{3}{2}(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1})h^2. \end{aligned} \quad (3.5)$$

Dodatno moramo konstruirati funkcije  $f_0(h)$  i  $f_{100}(h)$ . Uvjeti na funkciju  $f_0$  su  $f_0(0) = l_0$ ,  $f_0(1) = l_1$ ,  $f'_0(1) = f'_1(0) = \frac{1}{2}(l_2 - l_0)$ , odakle slijedi rezultat

$$f_0(h) = l_0 + \frac{1}{2}(-l_2 + 4l_1 - 3l_0)h + \frac{1}{2}(l_2 - 2l_1 + l_0)h^2, \quad (3.6)$$

$$f'_0(h) = \frac{1}{2}(-l_2 + 4l_1 - 3l_0) + (l_2 - 2l_1 + l_0)h. \quad (3.7)$$

Sad još uočimo da vrijedi

$$f'_0(h) < 0 \iff (l_2 - 2l_1 + l_0)h < -\frac{1}{2}(-l_2 + 4l_1 - 3l_0) \quad (3.8)$$

Kad se uzme u obzir da je  $l_0 = 100000$ ,  $l_1 = 99488$ ,  $l_2 = 99426$ , nejednadžba (3.8) postaje  $450h < 737$ , tj.  $h < 1,638$ , a to pokazuje da je  $f'_0(h) < 0, \forall h \in (0, 1)$ , dakle,  $f_0$  opada na cijelom segmentu  $[0, 1]$ , kako smo i željeli.

Analogno, moramo definirati  $f_{100}(h)$  za  $h \in [0, 1]$  tako da zadovoljava uvjete  $f_{100}(0) = l_{100}$ ,  $f_{100}(1) = 0$ ,  $f'_{100}(0) = f'_{99}(1) = \frac{1}{2}(l_{101} - l_{99}) = \frac{1}{2}l_{99}$ . Tražimo  $f_{100}$  u

obliku  $f_{100}(h) = a_{100} + b_{100}h + c_{100}h^2$  pa imamo uvjete  $\begin{cases} a_{100} = l_{100} \\ a_{100} + b_{100} + c_{100} = 0 \\ b_{100} = -\frac{1}{2}l_{99} \end{cases}$  iz

kojih slijedi  $c_{100} = -a_{100} - b_{100} = -l_{100} + \frac{1}{2}l_{99}$  pa dobivamo

$$f_{100}(h) = l_{100} - \frac{1}{2}l_{99}h + (-l_{100} + \frac{1}{2}l_{99})h^2, \quad (3.9)$$

$$f'_{100}(h) = -\frac{1}{2}l_{99} + (-2l_{100} + l_{99})h. \quad (3.10)$$

Još želimo da  $f_{100}$  opada na cijelom intervalu  $[0, 1]$ , dakle želimo da vrijedi  $f'_{100}(h) < 0, \forall h \in (0, 1)$ . Uočimo da je  $f'_{100}(h) < 0 \iff (-2l_{100} + l_{99})h < \frac{1}{2}l_{99}$ . Kad se uzme u obzir da je  $l_{99} = 309$  i  $l_{100} = 188$ , dobivamo nejednadžbu  $-67h < \frac{1}{2}309$ , odnosno  $h > -\frac{309}{134}$ , što znači da  $f$  opada za sve  $h > -\frac{309}{134}$ , pa posebno i na segmentu  $[0, 1]$ .

Ovime smo konstruirali funkciju  $k$  na cijelom segmentu  $[0, 101]$ . Osigurali smo da vrijedi  $k(n) = l_n, \forall n = 0, 1, 2, \dots, 101$  i postigli da je  $k$  derivabilna na cijelom intervalu  $(0, 101)$ . Naime, derivabilnost na svakom intervalu  $(i, i+1)$  neupitna je, a u čvornim točkama  $i = 1, 2, \dots, 100$  imamo lijevu i desnu derivaciju i one su jednake jer vrijedi  $f'_i(1) = f'_{i+1}(0), \forall i = 0, 1, \dots, 100$ . Sad je još preostalo osigurati da  $k$  bude padajuća funkcija na cijelom segmentu  $[0, 101]$ . To znači da sve funkcije  $f_i$  budu padajuće na segmentu  $[0, 1]$ . Već smo postigli da su funkcije  $f_0$  i  $f_{100}$  (dvije rubne, jedine dvije koje su kvadratne) padajuće. Ostaje vidjeti što je s funkcijama  $f_i$  za  $i = 1, 2, \dots, 99$ .

Uzmimo proizvoljan  $i, 1 \leq i \leq 99$ . Znamo da je  $f_i(0) = l_i > l_{i+1} = f_i(1)$  i vrijedi  $f_i(0), f_i(1) > 0$ . Funkcija  $f_i$  je derivabilna na cijelom  $\mathbb{R}$  i osim toga vrijedi  $f'_i(0) = \frac{1}{2}(l_{i+1} - l_{i-1}) < 0, f'_i(1) = \frac{1}{2}(l_{i+2} - l_i) < 0$ . Kako je  $f'_i$  neprekidna funkcija, znamo da je  $f'_i(h) < 0$  i u nekoj okolini oko  $h = 0$ , a također i u nekoj okolini oko točke  $h = 1$ . Zanima nas ponašanje funkcije  $f'_i$  na  $[0, 1]$ . Jedini način da  $f'_i$  poprimi pozitivnu vrijednost negdje u intervalu  $(0, 1)$ , s obzirom da je  $f'_i$  kvadratna parabola, je da joj je tjeme u nekoj točki  $h_0 \in (0, 1)$ , da ta točka bude maksimum, te da vrijedi  $f''_i(h_0) > 0$ . Nađimo to tjeme; treba riješiti jednadžbu  $f''_i(h_0) = 0$ . Iz (3.5) slijedi da je

$$(-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1}) + 3(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1})h_0 = 0. \quad (3.11)$$

Još uočimo da je relevantan predznak od  $f'''_i(h_0)$ . Iz (3.11) dobivamo

$$f'''_i(h) = 3(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1}). \quad (3.12)$$

Razmotrimo prvo mogućnost da je (to se zaista može dogoditi u praksi)

$$l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1} = 0. \quad (3.13)$$

Ako vrijedi (3.13), onda zapravo imamo

$$f_i(h) = l_i + \frac{1}{2}(l_{i+1} - l_{i-1})h + \frac{1}{2}(-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1})h^2 \quad (3.14)$$

i

$$f'_i(h) = \frac{1}{2}(l_{i+1} - l_{i-1}) + (-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1})h. \quad (3.15)$$

Ako dakle vrijedi (3.13), onda je  $f'_i$  linearna funkcija pa nema problema – s obzirom da su vrijednosti  $f'_i(0)$  i  $f'_i(1)$  negativne, jasno je da će tada  $f'_i$  biti negativna i za sve  $h$  između 0 i 1.

Dakle, jedini način da se problem dogodi je da istodobno budu ispunjeni sljedeći uvjeti:

$$l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1} \neq 0 \quad (3.16)$$

$$h_0 = \frac{l_{i+2} - 4l_{i+1} + 5l_i - 2l_{i-1}}{3(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1})} \in (0, 1) \quad (\text{ne smeta ako je } h_0 = 0 \text{ ili } h_0 = 1) \quad (3.17)$$

$$l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1} < 0 \quad (\text{ovo znači da } f'_i \text{ u točki } h_0 \text{ ima maksimum}) \quad (3.18)$$

$$\begin{aligned} & \frac{1}{2}(l_{i+1} - l_{i-1}) + (-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1})h_0 \\ & + \frac{3}{2}(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1})h_0^2 > 0 \end{aligned} \quad (3.19)$$

(ovaj uvjet znači da je  $f'_i(h_0) > 0$ )

Ukoliko su svi uvjeti (3.16)-(3.19) ispunjeni, graf funkcije  $f'_i$  na segmentu  $[0, 1]$  dvaput siječe os apscisa i imamo problem – funkcija  $f'_i$  nije monotona, tj. nerastuća.

Provjerom na stvarnim podacima vrijede li sva ta 4 uvjeta za pojedine  $i, i = 1, 2, \dots, 99$ , nalazimo da je jedini problem u funkciji  $f_1$ . Nađemo da je

$$\begin{aligned} f_1(h) &= 99488 - 287h + 435.5h^2 - 210.5h^3, \\ f'_1(h) &= -287 + 871h - 631.5h^2, \\ f''_1(h) &= 871 - 1263.5h, \end{aligned}$$

pa izlazi  $h_0 = 0.68963$ : to je točka maksimuma funkcije  $f'_1$  i  $f'_1(h_0) = 13.333$ . Sad uočimo sljedeće: svaku od funkcija  $f_i, i = 1, 2, \dots, 99$  možemo zamijeniti funkcijama  $g_i$ , također zadanim na intervalu  $[0, 1]$  sa

$$g_i(h) = a_i + b_i h + (c_i + \gamma)h^2 + (d_i - 2\gamma)h^3 + \gamma h^4 \quad (3.20)$$

Uočimo da je

$$g'_i(h) = b_i + 2(c_i + \gamma)h + 3(d_i - 2\gamma)h^2 + 4\gamma h^3 \quad (3.21)$$

Sad primjećujemo da je

$$\begin{aligned} g_i(0) &= a_i = f_i(0) \\ g'_i(0) &= b_i = f'_i(0) \\ g_i(1) &= a_i + b_i + c_i + \gamma + d_i - 2\gamma + \gamma = a_i + b_i + c_i + d_i = f_i(1) \\ g'_i(1) &= b_i + 2c_i + 2\gamma + 3d_i - 6\gamma + 4\gamma = b_i + 2c_i + 3d_i = f'_i(1). \end{aligned}$$

Dakle, i funkcija  $g_i$  zadovoljava tražene rubne uvjete i sad možemo pokušati naći  $\gamma$  tako da  $g_i$  bude padajuća na čitavom intervalu  $(0, 1)$ . To znači da tražimo  $\gamma$  takav da bude  $g'_i(h) < 0, \forall h \in (0, 1)$ .

Znamo da je

$$\begin{aligned} g'_i(0) &= b_i = \frac{1}{2}(l_{i+1} - l_{i-1}) \\ g'_i(1) &= b_i + 2c_i + 3d_i = \frac{1}{2}(l_{i+2} - l_i) \end{aligned}$$

U konkretnom slučaju imamo  $i = 1, b_1 = -287, c_1 = 435.5, d_1 = -210.5$ , pa je

$$\begin{aligned} g'_1(h) &= -287 + (871 + 2\gamma)h + (-631.5 - 6\gamma)h^2 + 4\gamma h^3 \\ g''_1(h) &= 871 + 2\gamma + (-1263 - 12\gamma)h + 12\gamma h^2. \end{aligned}$$

Kritične točke za funkciju  $g'_1(h)$  rješenje su jednadžbe  $g''_1(h) = 0$ ; dakle

$$h_{1,2} = \frac{1}{24\gamma}(1263 + 12\gamma \pm ((1263 + 12\gamma)^2 - 48\gamma(871 + 2\gamma))^{\frac{1}{2}}) \quad (3.22)$$

Funkcija pod korijenom neka se označi sa  $s(\gamma)$ . Pokazuje se da  $s(\gamma)$  ima minimum u točki  $\gamma_0 = 119.75$  te da je  $s(\gamma) > 0$ . Dakle, pod korijenom u (3.22) uvijek imamo pozitivnu vrijednost i najmanju vrijednost dobivamo za  $\gamma = 119.75$ . Uzmimo

$$\gamma = 120 \quad (3.23)$$

Sada imamo

$$\begin{aligned} g'_1(h) &= -287 + 1111h - 1351.5h^2 + 480h^3 \\ g''_1(h) &= 1111 - 2703h + 1440h^2 \end{aligned}$$

Kritične točke funkcije  $g'_1$  su

$$h_{1,2} = -\frac{1}{2880}(2703 \pm \sqrt{906849}) = \frac{1}{2880}(2703 \pm 952, 286),$$

odnosno  $h_1 \approx 0.608$  te  $h_2 \approx 1.269$ . Vidimo da je samo  $h_1$  kandidat za ekstrem funkcije  $g'_1$  na intervalu  $(0, 1)$ , stoga direktnim uvrštavanjem provjeravamo i dobijemo  $g'_1(h_1) \approx -3.230$ . Budući da u točkama 0 i 1 funkcija  $g'_1$  također ima negativne vrijednosti, to znači da  $g'_1$  ne mijenja predznak na  $(0, 1)$ , odnosno da je funkcija  $g_1$  padajuća na  $(0, 1)$ , što smo i željeli osigurati. Ovime smo izveli konkretnu interpolaciju i opravdali sve potrebne detalje.

Rekapitulirajmo: interpoliranu funkciju  $l$  označimo s  $k$  koja je složena od  $k_i$ .

1.

$$k_0(h) = l_0 + \frac{1}{2}(-l_2 + 4l_1 - 3l_0)h + \frac{1}{2}(l_2 - 2l_1 + l_0)h^2,$$

$$k'_0(0) = \frac{1}{2}(-l_2 + 4l_1 - 3l_0), \quad k'_0(1) = \frac{1}{2}(l_2 - l_0) = -287$$

2.

$$k_1(h) = l_1 + \frac{1}{2}(l_2 - l_0)h + \frac{1}{2}(-l_3 + 4l_2 - 5l_1 + 2l_0 + 240)h^2$$

$$+ \frac{1}{2}(l_3 - 3l_2 + 3l_1 - l_0 - 480)h^3 + 120h^4$$

$$k'_1(0) = \frac{1}{2}(l_2 - l_0) = -287, \quad k'_1(1) = \frac{1}{2}(l_3 - l_1)$$

3.

$$k_i(h) = l_i + \frac{1}{2}(l_{i+1} - l_{i-1})h + \frac{1}{2}(-l_{i+2} + 4l_{i+1} - 5l_i + 2l_{i-1})h^2$$

$$+ \frac{1}{2}(l_{i+2} - 3l_{i+1} + 3l_i - l_{i-1})h^3$$

$$k'_i(0) = \frac{1}{2}(l_{i+1} - l_i), \quad k'_i(1) = \frac{1}{2}(l_{i+2} - l_i)$$

4.

$$k_{100}(h) = l_{100} - \frac{1}{2}l_{99}h + (-l_{100} + \frac{1}{2}l_{99})h^2$$

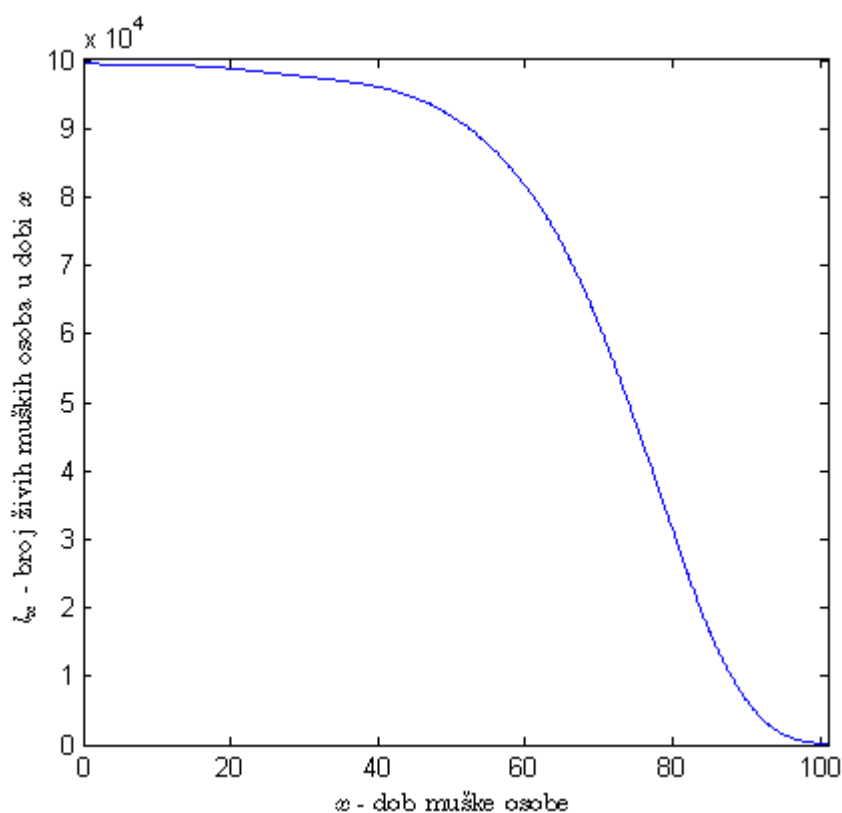
$$k'_{100}(0) = -\frac{1}{2}l_{99}, \quad k'_{100}(1) = -2l_{100} + \frac{1}{2}l_{99}$$

## 3.2 Grafički prikaz

Interpoliranu funkciju  $l$  označenu s  $k$  i sastavljenu od  $k_i$  na gore prikazan način definirali smo u programu jezika Matlab i potom grafički prikazali. Rezultat možemo vidjeti na slici 3.1. Vidimo da je funkcija  $l_x$  uistinu i naočigled izgladnena.

Kao što je već ranije rečeno, teško je doći do analitičkog zapisa funkcije  $\mu_x$  i stoga ju aproksimiramo pomoću vrijednosti  $l_x$ . Točnije, ovdje smo koristeći vrijednosti  $l_x$  za sve  $x$  između 0 i 1 s korakom 0.05 te aproksimacijske formule (2.21) i (2.22) konstruirali pripadne grafove za  $\mu_x$ . Slika 3.2 dobivena je uz aproksimaciju  $\mu_x$  iz formule (2.21), tj.

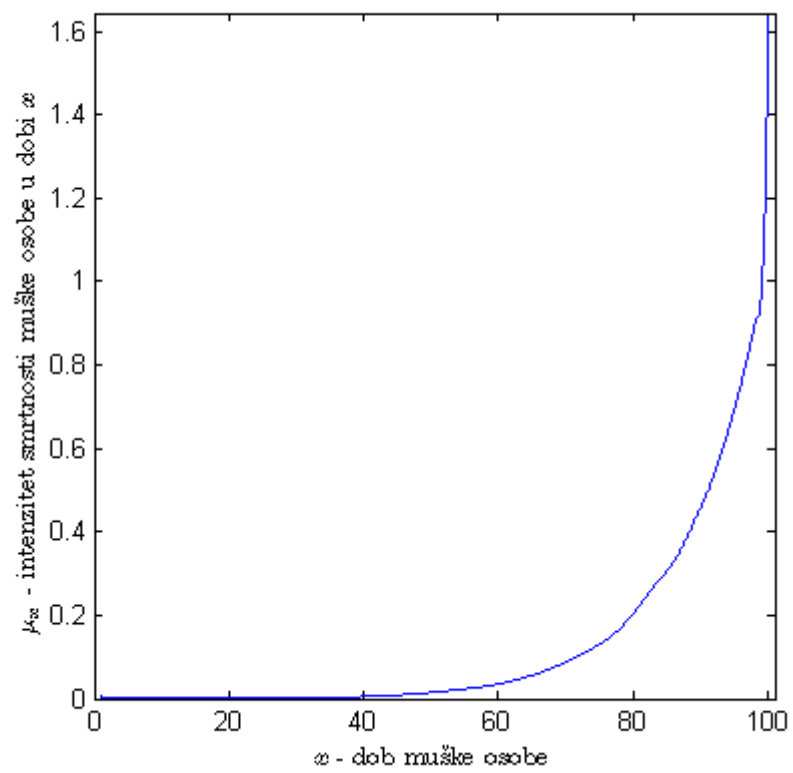
$$\mu_x \approx \frac{l_{x-1} - l_{x+1}}{2l_x},$$

Slika 3.1: Izgladena funkcija  $l_x$  dobivena Karupovom interpolacijom

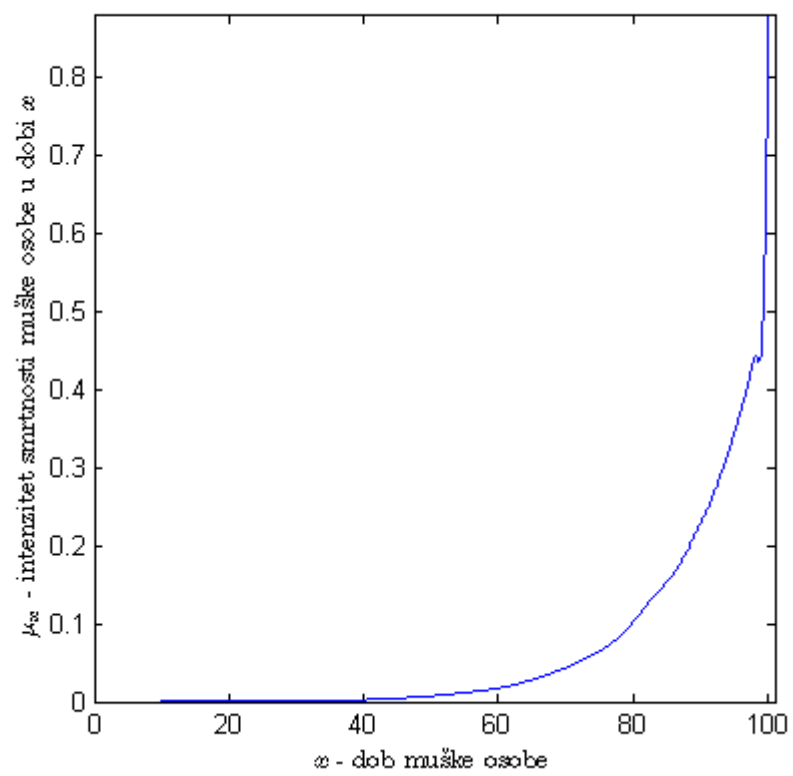
a slika 3.3 uz aproksimaciju iz (2.22), tj.

$$\mu_x \approx \frac{1}{12l_x}(8(l_{x-1} - l_{x+1}) - (l_{x-2} - l_{x+2})).$$

Kod obje aproksimacije u slučaju pojavljivanja nule u nazivniku cijeli smo rezultat izjednačavali s nulom. Vidimo da se nijedna od ove dvije aproksimacije ne ponaša tako dobro kao aproksimacija  $l_x$ , što je razumljivo s obzirom da je sam proces dobivanja aproksimacije  $l_x$  puno sofisticiraniji te je ovdje u oba slučaja  $\mu_x$  dobivena kao aproksimacija na temelju aproksimacije. Ipak, s obzirom na nedostupnost egzaktnih formula za  $\mu_x$ , i ovakve aproksimacije u praksi mogu dobro poslužiti.



Slika 3.2: Aproksimacija  $\mu_x \approx \frac{l_{x-1} - l_{x+1}}{2l_x}$



Slika 3.3: Aproksimacija  $\mu_x \approx \frac{1}{12l_x}(8(l_{x-1} - l_{x+1}) - (l_{x-2} - l_{x+2}))$



# Bibliografija

- [1] D.C.M. Dickson, M. Hardy, M.R. Hardy i H.R. Waters, *Actuarial Mathematics for Life Contingent Risks*, International Series on Actuarial Science, Cambridge University Press, 2013, ISBN 9781107044074.
- [2] A. Neill, *Life Contingencies*, Butterworth-Heinemann, 1977, ISBN 9780750609173.
- [3] N. Sarapa, *Teorija vjerojatnosti*, Udžbenici Sveučilišta u Zagrebu, Školska knjiga, 1992, ISBN 9788603007411.
- [4] Hrvatski zavod za mirovinsko osiguranje, *Tablice aktuarske matematike*, [http://www.mirovinsko.hr/UserDocsImages/Upravnovijece/sjednice/6\\_31012013/tablice\\_aktuarske\\_matematike\\_potpisano.pdf](http://www.mirovinsko.hr/UserDocsImages/Upravnovijece/sjednice/6_31012013/tablice_aktuarske_matematike_potpisano.pdf).

# Sažetak

U ovome radu najprije predstavljamo modele doživljenja, pri čemu definiramo osnovne pojmove aktuarske matematike i uvodimo pripadnu notaciju. Poseban naglasak stavljamo na tablice smrtnosti i ono što iz njih možemo izvesti, kao što su primjerice vjerojatnosti smrti i doživljenja te aproksimacija intenziteta smrtnosti. Središnje mjesto rada zauzima upravo interpolacija empirijskih podataka o broju živih osoba pojedine cjelobrojne dobi pomoću Karupove metode, što na kraju rada i grafički prikazujemo.

# Summary

In this thesis we firstly introduce survival models, wherein we define fundamental notions of actuarial mathematics and introduce actuarial notation. Special attention is given to the life tables and all that can be derived from them, such as death and survival probabilities or approximation of the force of mortality. The centrepiece of the thesis is the very interpolation of empirical data about the number of persons alive at particular integer age using Karup's method, which is graphically displayed at the end of the thesis.

# Životopis

Matea Galović rođena je 24. lipnja 1992. u Zagrebu. Nakon završetka Osnovne škole Jure Kaštelana, upisuje V. gimnaziju u Zagrebu, gdje uz postizanje odličnog uspjeha sudjeluje i na brojnim natjecanjima u znanju, kako iz matematike, tako i ostalih predmeta. Godine 2011. upisuje preddiplomski studij Matematika na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu koji završava 2014. s odličnim uspjehom te iste godine na tom fakultetu upisuje diplomski studij Financijska i poslovna matematika. Od 2013. mentor je udruge Mladi nadareni matematičari "Marin Getaldić" koja se bavi promocijom i unapređenjem matematike. Godine 2014. nagrađena je Rektorovom nagradom Sveučilišta u Zagrebu za rad u suradnji s kolegom Matkom Ljuljem, a godine 2016. pohvalnicom Fakultetskog vijeća Prirodoslovno-matematičkog fakulteta za izuzetan uspjeh u studiju.