UNIVERZITET U BEOGRADU
MATEMATIČKI FAKULTET

Dunja Radenković

PROCENA SEZONSKE SMRTNOSTI POPULACIJE SA PRIMENOM U ŽIVOTNOM OSIGURANJU

master rad

Beograd, 2024.

Mentor:

dr Jelena Jocković, docent
Univerzitet u Beogradu, Matematički fakultet

Članovi komisije:

prof. dr Bojana Milošević, vanredni profesor
Univerzitet u Beogradu, Matematički fakultet

dr Marija Cuparić, docent
Univerzitet u Beogradu, Matematički fakultet

Datum odbrane: ____________________

Zahvaljujem se mentorki, dr Jeleni Jocković, na smernicama prilikom odabira teme i razumevanju i pomoći prilikom izrade rada, kao i članovima komisije, prof. dr Bojani Milošević i dr Mariji Cuparić, na svim korisnim savetima i sugestijama. Veliku zahvalnost dugujem svojoj porodici na podršci i razumevanju tokom godina.

Naslov master rada: Procena sezonske smrtnosti populacije sa primenom u životnom osiguranju

Rezime: Glavni cilj ovog rada je potraga za modelima koji se bave procenom sezonske smrtnosti. Motivaciju u najvećoj meri predstavljaju nedostaci u već postojećim i poznatim modelima. Na osnovu stvarnih podataka uočava se neadekvatnost i ograničenost spomenutih modela. Ocena sezonske smrtnosti u životnom osiguranju neposredno utiče na određivanje osigurane sume koju je osiguravač u obavezi da isplati, a posledično i na premiju koju osiguranik uplaćuje. Određivanje premije je od izuzetnog značaja jer osiguravajuća kompanija ne sme odrediti iznos koji je previše visok da bi osiguranik imao motivaciju da vrši uplate. S druge strane, ne sme odrediti ni previše nizak iznos da bi bila u stanju da pokrije sve svoje stvarne troškove koji se javljaju u toku poslovanja. Na osnovu svega izrečenog jasan je značaj ove teme. Rad se sastoji iz tri celine: u prvoj celini se pravi uvod. Uvode se najbitniji pojmovi i definicije iz životnog osiguranja bez kojih ne bi bilo moguće baviti se ovom tematikom. Drugi deo se bavi predstavljanjem već postojećeg rada gde se predlaže jedno moguće rešenje. Treći deo čine pokušaji da se primenom modela poznatih iz izučavanja vremenskih serija obezbedi model koji će na pravi način prikazati sezonsku smrtnost.

Ključne reči: osiguranje, vremenske serije, smrtnost, tablice

Sadržaj

Glava 1

Uvod

Prvi vidovi osiguranja datiraju još iz Starog veka. U Hamurabijevom zakoniku su postojali članovi koji mogu da se tumače kao preteče osiguranja. Jedan od njih je predviđao oslobađanje dužnika od obaveze otplate dugova u slučaju realizacije nekog katastrofalnog događaja poput poplave, gubitka sposobnosti za rad ili smrti. Trgovci su plaćali određenu sumu poveriocu koji zauzvrat ne bi potraživao dug u slučaju da brod kojim se roba prevozi bude potopljen ili ukraden. Grci i Rimljani su formirali udruženja od velikog značaja - esnafe. Članovi esnafa su uplaćivali novac u zajedničku kasu odakle su mogli da obezbede svojim članovima i njihovim porodicama finansijsku zaštitu u slučaju npr. oštećenja imovine u požaru, pljačke ili smrti [1]. Tek u 17. veku, kada francuski matematičar Paskal postavlja ono za šta će se kasnije ispostaviti da je temelj moderne teorije verovatnoće, osiguranje počinje da se razvija po naučnom osnovu, što će imati direktan uticaj na razvoj osiguranja koje je približnije pojmu osiguranja koji danas poznajemo. Tokom godina, koncept osiguranja je evoluirao i prilagođavao se potrebama ljudi i celokupnom društveno-ekonomskom razvoju.

Ipak, motivacija i svrha osiguranja ostaje ista - zaštita osiguranika od posledica koje donosi realizacija određenog **rizika** od kog je osiguran. Rizik je neizvesnost u pogledu nekog budućeg događaja. Taj rizik mora da ispunjava određene uslove da bi mogao da bude predmet osiguranja. Pre svega, realizacija rizika mora da bude neizvesna, odnosno nijednoj strani, ni osiguraniku ni osiguravaču, ne sme biti poznato da se rizik realizovao ili da će se sa verovatnoćom 1 realizovati (ili neće realizovati) u budućnosti. Drugo, ne mogu se osigurati rizici koji dovode do gubitka kod velikog broja osiguranika. Pod tim mogu da se podrazumevaju razne pojave poput ratova, poplava, pandemija i slično. Treće, neophodno je da postoji dovoljno velika grupa

ljudi na osnovu čijeg ponašanja u prošlosti je moguće proceniti rizik na način koji će biti verodostojan. Četvrto, rizik mora da se jasno i precizno odredi. Na kraju, potrebno je da je verovatnoća da se rizik ostvari dovoljno mala da bi se osiguravaču isplatilo pružanje usluga osiguranja u tom slučaju.
Osiguranik zaključuje ugovor o osiguranju sa **osiguravačem** - pravnim licem koje ima određeni pravni status i koje ima dozvolu regulatornog tela za obavljanje poslova osiguranja. Osiguranik (ili u nekim slučajevima ugovarač osiguranja) je obavezan da plaća premije da bi mogao da ostvari pravo na isplatu **osigurane sume**. Osigurana suma je najveći iznos koji osiguravač može da plati osiguraniku. Ona ima poseban značaj u životnom osiguranju kada nije moguće odrediti odštetnu vrednost i koristi se neposredno za određivanje iznosa premije.
Osiguranje se oslanja na koncept tzv. udruživanja rizika. Radi se o formiranju zajednica gde život ili imovinu članova te zajednice ugrožava isti rizik. Na taj način vrši se raspodela rizika koji svi članovi snose u jednakoj meri. Poželjno je da zajednica bude što veća, jer će tada pojedinac da snosi manji rizik što će se ogledati i u iznosu premije [2].
Životno osiguranje je specifična vrsta osiguranja. Osim pomenute razlike u karakteru osigurane sume životno osiguranje ima izrazit štedni karakter koji je čini jednom izuzetno značajnom institucijom modernog doba. Prilikom realizacije osiguranog rizika isplaćuje se osigurana suma (jednokratno) ili renta (niz isplata u prethodno dogovorenom i jednoznačno određenom vremenskom periodu).

1.1 Osnovni elementi finansijske matematike u životnom osiguranju

Aktuarska matematika kombinuje upotrebu finansijake matematike i teorije verovatnoće. U ovom poglavlju ćemo uvesti neke od osnovnih pojmova iz finansijske matematike.

Definicija. Tok novca je niz uređenih parova (t_k, x_k), gde je $k \in \mathbf{N}$, a t_k vremenski trenuci u kojima dolazi do isplata novčanih iznosa x_k.

Definicija. Slučajni tok novca je niz uređenih parova (T_k, X_k), za $k \in \mathbf{N}$, definisanih na istom prostoru verovatnoće.

Može se primetiti da je tok novca realizacija slučajnog toka novca.

Definicija. Kamata je nadoknada za upotrebu kapitala. Ona odražava vremensku promenu vrednosti novca.

Ukoliko se sa i označi godišnja **kamatna stopa** (eng. *interest rate*), tada u zavisnosti od načina na koji se vrši kamaćenje razlikujemo slučajeve:

1. **Prosto kamaćenje**. Neka se kamaćenje vrši jednom godišnje. Ukoliko je početna uložena suma G, tada će se nakon t godina akumulirati iznos:

$$G(1 + t \cdot i).$$

2. **Složeno kamaćenje**. U slučaju da se kamaćenje vrši jednom godišnje, početna uložena suma G će nakon t godina iznositi:

$$G(1 + i)^t.$$

Sa druge strane, ukoliko se kamaćenje vrši m puta godišnje, $m \in \mathbf{N}$, tada se nakon t godina akumulira iznos:

$$G\left(1 + \frac{i}{m}\right)^{mt}.$$

Primenom binomne formule može se primetiti da je kamata dobijena putem složenog kamaćenja veća od one dobijene putem prostog:

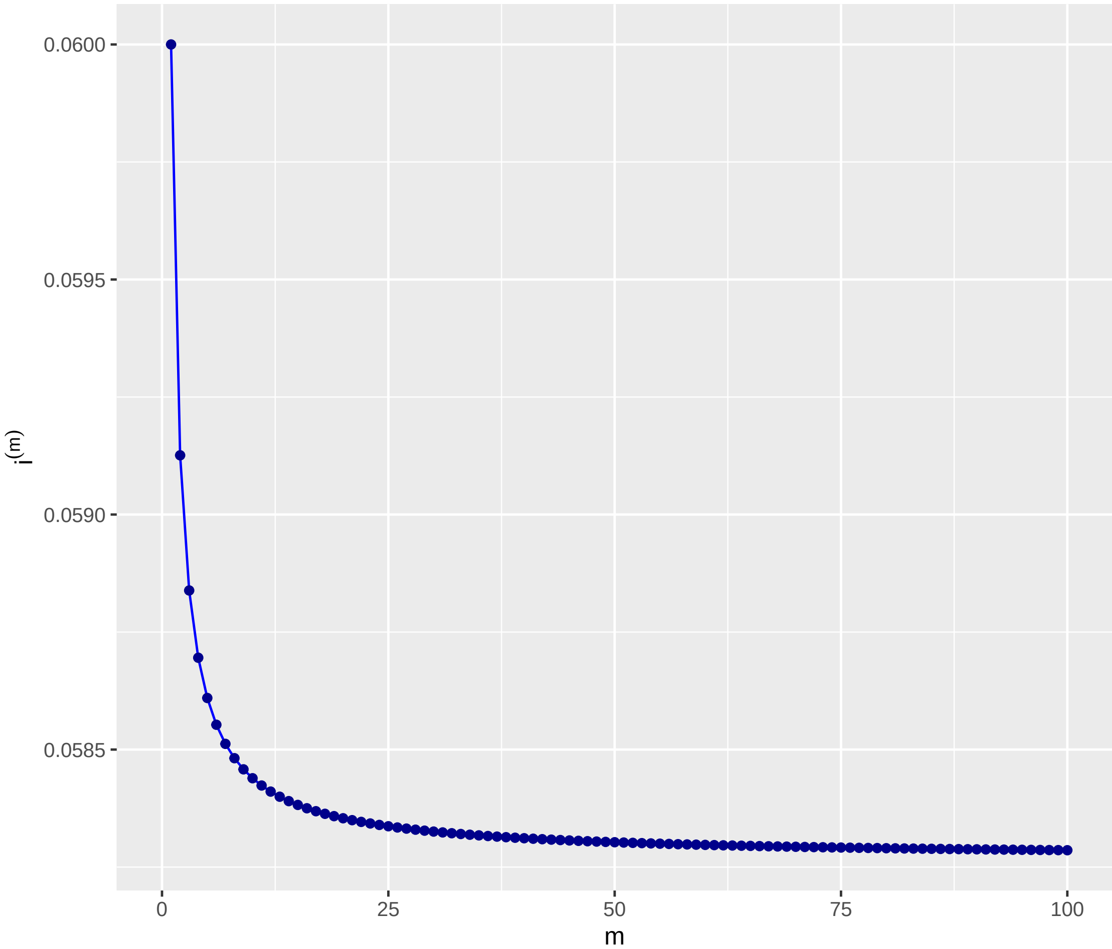
$$\begin{aligned} \left(1 + \frac{i}{m}\right)^m &= \sum_{k=0}^{m} \binom{m}{k} \left(\frac{i}{m}\right)^k 1^{m-k} \\ &= \binom{m}{0} \cdot \left(\frac{i}{m}\right)^0 + \binom{m}{1} \cdot \left(\frac{i}{m}\right)^1 + ... + \binom{m}{m} \cdot \left(\frac{i}{m}\right)^m \\ &= 1 + m \cdot \frac{i}{m} + ... + \left(\frac{i}{m}\right)^m, \end{aligned}$$

odakle je:

$$\left(1 + \frac{i}{m}\right)^m > 1 + m \cdot \frac{i}{m} = 1 + i.$$

Definicija. Efektivna kamatna stopa i je ona kamatna stopa koja pri prostom kamaćenju proizvodi isti efekat kao $i^{(m)}$ pri složenom kamaćenju, odnosno:

$$1 + i = \left(1 + \frac{i^{(m)}}{m}\right)^m. \tag{1.1}$$



Slika 1.1: Funkcija zavisnosti $i^{(m)}$ od m

Kamatna stopa $i^{(m)}$ se tada naziva **nominalna kamatna stopa**. Sada može da se izrazi vrednost efektivne i nominalne kamatne stope preko formula:

$$i = \left(1 + \frac{i^{(m)}}{m}\right)^m - 1, \tag{1.2}$$

i

$$i^{(m)} = m\left((1+i)^{\frac{1}{m}} - 1\right). \tag{1.3}$$

Na slici 1.1 se nalazi prikaz funkcionalne zavisnosti $i^{(m)}$ od m ukoliko efektivna kamatna stopa uzima, na primer, vrednost $i = 6\%$. Može se primetiti da to opadajuća funkcija, što znači da će nominalna kamata opadati sa porastom broja kamaćenja.

Prethodno razmatranje se odnosi na slučaj kada se kamaćenje vrši diskretno. U slučaju da $m \to \infty$, kaže se da se kamaćenje vrši neprekidno i tada je:

$$\delta = \lim_{m\to\infty} i^{(m)}. \tag{1.4}$$

Transformacijom izraza sa desne strane jednakosti (1.3):

$$i^{(m)} = \frac{(1+i)^{\frac{1}{m}} - (1+i)^0}{\frac{1}{m}},$$

i kombinacijom sa (1.4) dobija se:

$$\delta = \lim_{m\to\infty} \frac{(1+i)^{\frac{1}{m}} - (1+i)^0}{\frac{1}{m}} = \frac{\partial((1+i)^x)}{\partial x}|_{x=0} = \ln(1+i),$$

ili, alternativno:

$$1+i = e^{\delta},$$

što se može potvrditi i puštanjem da $\lim_{m\to\infty}$ u jednakosti (1.1).

Ukoliko se δ menja u toku vremena, tada se $\delta = \delta(t), t \geq 0$, naziva **trenutna kamatna stopa u trenutku t**.

Neka je $G(t)$ suma novca koja će se nakupiti do trenutka t. Pretpostavimo da se u početnom trenutku $t = 0$ ulaže 1 novčana jedinica, odnosno da je $G(0) = 1$. Ukoliko se posmatra interval $(t, t+dt)$, onda će akumulirani iznos u tom vremenskom periodu biti jednak $G(t)\delta(t)dt$. Tada važi:

$$dG(t) = G(t)\delta(t)dt,$$

$$\int_0^t \frac{dG(s)}{G(s)} = \int_0^t \delta(s)\ \mathrm{d}s,$$

$$\int_0^t d(\ln(G(s)) = \int_0^t \delta(s)\ \mathrm{d}s,$$

$$\ln G(t) - \ln G(0) = \int_0^t \delta(s)\ \mathrm{d}s,$$

$$G(t) = e^{\int_0^t \delta(s)\ \mathrm{d}s},$$

odnosno toliko iznosi akumulirana suma u trenutku t, $t > 0$.

Tokom vremena novac menja svoju vrednost. Ukoliko u banci sa kamatnom stopom i uložimo sumu A, ona će kroz godinu dana imati vrednost $A(1+i)$. Odatle motiv da se uvede sledeća definicija.

Definicija. Sadašnja vrednost novčanog iznosa A kroz godinu dana je

$$\frac{A}{1+i},$$

gde je i kamatna stopa.

Izraz $\frac{1}{1+i}$ se naziva diskontni faktor. Kod složenog kamaćenja m puta godišnje diskontni faktor je jednak

$$v^{(m)} = \frac{1}{1+\frac{i^{(m)}}{m}}.$$

U slučaju da se kamaćenje vrši neprekidno, sa konstantnom stopom kamaćenja, diskontni faktor je $v = e^{-\delta}$, dok je za promenljivu stopu $v = e^{-\int_0^t \delta(s)\mathrm{d}s}$.

Može se izvesti i veza između kamatne stope za prosto i složeno kamaćenje na osnovu (1.3):

$$v^{(m)} = \frac{1}{1+\frac{i^{(m)}}{m}} = \frac{1}{1+\frac{m((1+i)^{\frac{1}{m}}-1)}{m}} = \frac{1}{(1+i)^{\frac{1}{m}}} = v^{\frac{1}{m}}.$$

Ovi pojmovi su posebno bitni kod tokova novca. Osim što daju informaciju o novčanoj vrednosti, sadašnja vrednost predstavlja jedan od rešenja za problem poređenja novčanih tokova. Novčani tokovi su ekvivalentni ukoliko su im sadašnje vrednosti iste.

Računamo sadašnju i buduću vrednost novčanih tokova pri:

1. Prostom kamaćenju. Neka je $(x_0, x_1, ..., x_n)$ tok novca gde se svaki iznos uplaćuje jednom godišnje. Tada je buduća vrednost novčanih tokova pri prostom kamaćenju, a posle n godina:

   $$BV = \sum_{k=0}^{n} x_k(1+i)^{n-k},$$

   dok je sadašnja vrednost novčanih tokova pri prostom kamaćenju zapravo jednaka budućoj podeljenoj faktorom $(1+i)^n$:

   $$SV = \sum_{k=0}^{n} \frac{x_k}{(1+i)^k} = \sum_{k=0}^{n} x_k v^k.$$

2. Složenom kamaćenju. Neka je $(x_0, x_1, ..., x_n)$ tok novca takav da uplate pristižu m puta godišnje. Tada se i kamaćenje vrši m puta godišnje, a buduća vrednost toka novca pri složenom kamaćenju, a nakon n perioda odgovarajuće dužine je:
$$BV = \sum_{k=0}^{n} x_k \left(1 + \frac{i^{(m)}}{m}\right)^{n-k},$$
dok je sadašnja vrednost novčanih tokova pri složenom kamaćenju:
$$SV = \sum_{k=0}^{n} x_k {v^{(m)}}^k.$$

3. Neprekidnom kamaćenju. Ukoliko je $(x(t_0), x(t_1), ..., x(t_n))$ niz isplata u proizvoljnim trenucima $(t_0, t_1, ..., t_n)$. Tada je buduća vrednost toka novca pri neprekidnom kamaćenju:
$$BV = \sum_{k=0}^{n} x(t_k) e^{\delta(t_n - t_k)},$$
dok je vrednost u trenutku t_0 novčanih tokova pri neprekidnom kamaćenju:
$$SV = \sum_{k=0}^{n} x(t_k) e^{-\delta(t_k - t_0)}.$$

Navedeni rezultati se mogu pronaći u knjizi [4].

Perpetuiteti i anuiteti

Definicija. Perpetuitet je beskonačni niz plaćanja jednakih iznosa u jednakim vremenskim intervalima.

Osnovne vrste perpetuiteta su:

1. Jednom godišnje se uplaćuje jedna novčana jedinica. U odnosu na trenutak prve uplate, razlikujemo slučajeve:

   1.1. Prva isplata je u trenutku $t = 0$, tj. na početku prve godine. Tada je sadašnja vrednost $\ddot{a}_{\overline{\infty}|}$ niza uplata:
$$\ddot{a}_{\overline{\infty}|} = 1 + v + v^2 + v^3 + ... = \frac{1}{1 - v}.$$

1.2. Prva isplata je u trenutku $t = 1$, tj. na kraju prve godine. Tada je sadašnja vrednost $a_{\overline{\infty|}}$ niza uplata:

$$a_{\overline{\infty|}} = v + v^2 + v^3 + ... = \frac{v}{1-v}.$$

Poslednji izraz možemo zapisati i preko definicije diskontnog faktora, tj.:

$$\frac{v}{1-v} = \frac{\frac{1}{1+i}}{1-\frac{1}{1+i}} = \frac{1}{i}.$$

2. m puta godišnje se uplaćuje iznos $\frac{1}{m}$.

2.1. Prva isplata je u trenutku $t = 0$, tj. na početku prve godine. Tada je sadašnja vrednost $\ddot{a}^{(m)}_{\overline{\infty|}}$ niza uplata:

$$\ddot{a}^{(m)}_{\overline{\infty|}} = \frac{1}{m} + \frac{1}{m}v^{\frac{1}{m}} + \frac{1}{m}v^{\frac{2}{m}} + \frac{1}{m}v^{\frac{3}{m}} + ... = \frac{1}{m}\frac{1}{1-v^{\frac{1}{m}}}.$$

2.2. Prva isplata je u trenutku $t = \frac{1}{m}$, tj. na kraju prvog perioda. Tada je sadašnja vrednost $a^{(m)}_{\overline{\infty|}}$ niza uplata:

$$a^{(m)}_{\overline{\infty|}} = \frac{1}{m}v^{\frac{1}{m}} + \frac{1}{m}v^{\frac{2}{m}} + \frac{1}{m}v^{\frac{3}{m}} + ... = \frac{1}{m}v^{\frac{1}{m}}\frac{1}{1-v^{\frac{1}{m}}}.$$

Poslednji izraz možemo transformisati:

$$\frac{1}{m}v^{\frac{1}{m}}\frac{1}{1-v^{\frac{1}{m}}} = \frac{1}{m}\frac{\frac{1}{1+i}^{\frac{1}{m}}}{1-\frac{1}{1+i}^{\frac{1}{m}}} = \frac{1}{m}\frac{1}{(1+i)^{\frac{1}{m}}-1} = \frac{1}{i^{(m)}}.$$

3. Kada se vrši neprekidno kamaćenje, odnosno $m \to \infty$, sadašnja vrednost plaćanja je:

$$\bar{a}_{\overline{\infty|}} = \int_0^\infty e^{-\delta t}\, \mathrm{d}t = \frac{1}{\delta}.$$

U praksi se češće od perpetuiteta sreću anuiteti.

Definicija. Anuitet je niz plaćanja jednakih iznosa u jednakim vremenskim intervalima sa ograničenim trajanjem.

Osnovne vrste anuiteta su:

1. Jednom godišnje se uplaćuje jedna novčana jedinica, u trajanju od n godina. U odnosu na trenutak prve uplate, razlikujemo slučajeve:

1.1. Prva isplata je u trenutku $t = 0$. Sadašnja vrednost $\ddot{a}_{\overline{n}|}$ niza uplata je:

$$\ddot{a}_{\overline{n}|} = 1 + v + v^2 + v^3 + ... + v^{n-1} = \frac{1 - v^n}{1 - v}.$$

1.2. Prva isplata je u trenutku $t = 1$. Sadašnja vrednost $a_{\overline{n}|}$ niza uplata je:

$$a_{\overline{n}|} = v + v^2 + v^3 + ... + v^n = v\frac{1 - v^n}{1 - v}.$$

2. Isplaćuje se iznos $\frac{1}{m}$ m puta godišnje, u trajanju od n godina. U odnosu na trenutak prve uplate, razlikujemo slučajeve:

2.1. Prva isplata je u trenutku $t = 0$, tj. na početku prve godine. Tada je sadašnja vrednost $\ddot{a}^{(m)}_{\overline{n}|}$ niza uplata:

$$\ddot{a}^{(m)}_{\overline{n}|} = \frac{1}{m} + \frac{1}{m}v^{\frac{1}{m}} + \frac{1}{m}v^{\frac{2}{m}} + \frac{1}{m}v^{\frac{3}{m}} + ... + \frac{1}{m}v^{\frac{mn-1}{m}} = \frac{1}{m}\frac{1 - v^n}{1 - v^{\frac{1}{m}}}.$$

2.2. Prva isplata je u trenutku $t = \frac{1}{m}$, tj. na kraju prvog perioda. Tada je sadašnja vrednost $a^{(m)}_{\overline{n}|}$ niza uplata:

$$a^{(m)}_{\overline{n}|} = \frac{1}{m}v^{\frac{1}{m}} + \frac{1}{m}v^{\frac{2}{m}} + \frac{1}{m}v^{\frac{3}{m}} + ... + \frac{1}{m}v^{\frac{mn}{m}} = \frac{1}{m}v^{\frac{1}{m}}\frac{1 - v^n}{1 - v^{\frac{1}{m}}}.$$

Anuiteti se mogu zapisati u i obliku razlike dva perpetuiteta; jednog koji počinje u trenutku $t = 0$ i drugog koji počinje u trenutku $t = n$:

$$\ddot{a}_{\overline{n}|} = \ddot{a}_{\overline{\infty}|} - v^n\ddot{a}_{\overline{\infty}|}.$$

Perpetuiteti i anuiteti su značajni u životnom osiguranju jer oni predstavljaju sadašnju vrednost premija koje osiguranik uplaćuje ili rente koju isplaćuje osiguravajuća kompanija.

Glava 2

Modeli za procenu životnog veka

Procena smrtnosti zavisi od godina osiguranika i njegovog eventualnog životnog veka. Uvodi se pojam starosti osobe u trenutku sklapanja ugovora o osiguranju i nju označavamo sa x. Neka je $T(x)$ preostali životni vek osiguranika, tako da je onda $x + T(x)$ starost osobe u trenutku smrti. Tada je $T = T(x)$ apsolutno neprekidna slučajna veličina sa vrednostima u $[0, \infty)$ i funkcijom raspodele $G(t), t \geq 0$, odnosno:

$$G(t) = P\{T < t\}.$$

Dakle, $G(t)$ je verovatnoća da za svako fikirano t osoba neće preživeti narednih t godina. Funkcija gustine slučajne veličine T je $g(t) = G'(t)$. Može se definisati verovatnoća da se u malom vremenskom intervalu $(t, t+dt)$ dogodi smrt osiguranika:

$$g(t)dt = P\{t < T < t + dt\}.$$

Pored toga, naredne definicije uvode oznake koje se koriste u okviru aktuarske zajednice na međunarodnom nivou.

Definicija. Verovatnoća da će osoba starosti x preminuti u narednih t godina je:

$${}_tq_x = P\{T(x) \leq t\} = G(t),\ t \geq 0.$$

Definicija. Verovatnoća da osoba starosti x preživi narednih t godina je:

$${}_tp_x = P\{T(x) > t\} = 1 - G(t),\ t \geq 0$$

i važi da je ${}_tp_x + {}_tq_x = 1$.

Na osnovu ovih definicija, može se uvesti i oznaka za verovatnoću da osoba starosti x preživi s godina, ali ne preživi narednih $s+t$, odnosno:

$$\begin{aligned} {}_{s|t}q_x &= P\{s < T(x) < t+s\} \\ &= G(s+t) - G(s) \\ &= {}_{s+t}q_x - {}_sq_x. \end{aligned}$$

Uslovnu verovatnoću da će osoba koja je preživela $x+s$ godina preživeti i sledećih t označavamo sa:

$$ {}_tp_{x+s} = P\{T(x) > t+s | T(x) > s\} = \frac{1-G(s+t)}{1-G(s)} = \frac{{}_{t+s}p_x}{{}_sp_x}. \tag{2.1}$$

Analogno, definiše se i verovatnoća suprotnog događaja:

$$ {}_tq_{x+s} = P\{T(x) \leq t+s | T(x) > s\} = \frac{G(s+t)-G(s)}{1-G(s)} = \frac{{}_{t+s}q_x - {}_sq_x}{{}_sp_x}. \tag{2.2}$$

Odavde se može izvesti da na osnovu (2.1) važi:

$$ {}_{s+t}p_x = 1 - G(t+s) = (1-G(s))\frac{1-G(s+t)}{1-G(s)} = {}_sp_x \ {}_tp_{x+s}.$$

Kao i analogno na osnovu (2.2):

$$ {}_{s|t}q_x = G(s+t) - G(s) = (1-G(s))\frac{G(s+t)-G(s)}{1-G(s)} = {}_sp_x \ {}_tq_{x+s}. \tag{2.3}$$

Na kraju, možemo formulisati i očekivanje slučajne veličine $T(x)$:

$$\mathring{e}_x = \int_0^\infty \mathrm{tg}(t) \ \mathrm{d}t = \int_0^\infty {}_tp_x \ \mathrm{d}t.$$

Još jedan pojam koji se sreće u literaturi vezanoj za životno osiguranje je intenzitet smrtnosti (eng. *force of mortality*) kod osobe starosti $x+t$. Tačnije, posmatra se verovatnoća da osoba starost x premine u vremenskom periodu $(t, t+dt)$, pod uslovom da je već doživela t godina. Zapisujemo:

$$P\{t < T < t+dt | T > t\} = \frac{G(t+dt)-G(t)}{1-G(t)} = \frac{g(t)}{1-G(t)}dt. \tag{2.4}$$

Na osnovu tog zapisa, uvodi se definicija:

$$\mu_{x+t} = \frac{g(t)}{1 - G(t)} = -\frac{d}{dt}\ln(1 - G(t)) = -\frac{d}{dt}\ln({}_tp_x), \tag{2.5}$$

odakle sledi:

$${}_tp_x = e^{-\int_0^t \mu_{x+s}\ \mathrm{ds}}. \tag{2.6}$$

Sada možemo izvesti jednu jednostavnu aproksimaciju za verovatnoću ${}_sq_{x+t}$ u kojoj se pojavljuje intenzitet smrtnosti. Iz jednakosti (2.3) važi:

$${}_{t|s}q_x = G(s+t) - G(t) = {}_tp_x \ {}_sq_{x+t}. \tag{2.7}$$

Sa druge strane, na osnovu jednakosti (2.4) sledi i:

$$P\{t < T < t + dt\} = G(t + dt) - G(t) = {}_tp_x\mu_{x+t}dt. \tag{2.8}$$

Kada u poslednjoj jednakosti umesto dt napišemo s, dobija se:

$$G(t+s) - G(t) = {}_tp_x\mu_{x+t}s, \tag{2.9}$$

te odatle i iz (2.7) sledi:

$${}_sq_{x+t} \approx \mu_{x+t}s.$$

2.1 Preostali životni vek

Nekada je u životnom osiguranju, umesto slučajne veličine $T(x)$, logičnije posmatrati njen celobrojni deo. Zbog toga se uvodi naredna definicija:

Definicija. Celobrojni preostali životni vek (eng. *curtate future lifetime*) je slučajna veličina takva da je:

$$K = K(x) = [T(x)],$$

gde je $T(x)$ preostali životni vek osobe starosti x.

Tada je njegova funkcija raspodele:

$$P\{K = k\} = P\{k \leq T < k+1\} = {}_kp_x q_{x+k},$$

gde se koristi da je ${}_1q_x = q_x$ (analogno važi i za ${}_1p_x$).

Odavde se može izračunati i očekivanje ove slučajne veličine tj. **očekivani celobrojni preostali životni vek**:

$$e_x = \sum_{k=1}^{\infty} kP\{K = k\} = \sum_{k=1}^{\infty} k_k p_x q_{x+k},$$

što možemo zapisati i kao:

$$e_x = \sum_{k=1}^{\infty} P\{K \geq k\} = \sum_{k=1}^{\infty} {}_k p_x.$$

Ovu vrednost je često jednostavnije izračunati nego $\mathring{e}_x$ jer u većini slučajeva imamo informaciju o raspodeli slučajne veličine K, a ne T. Neka je S deo godine smrti koji je osoba preživela, odnosno:

$$T = K + S.$$

Ako se pretpostavi da su S i K nezavisne, tada je uslovna raspodela za S, pri K=k, nezavisna od k, te sledi da:

$$\begin{aligned} P\{S \leq u | K = k\} &= \frac{P\{S \leq u, K = k\}}{P\{K = k\}} && (2.10) \\ &= \frac{P\{T \leq k + u | T \geq k\} P\{T \geq k\}}{P\{K = k\}} && (2.11) \\ &= \frac{{}_u q_{x+k} \cdot {}_k p_x}{{}_k p_x \cdot q_{x+k}} && (2.12) \\ &= \frac{{}_u q_{x+k}}{q_{x+k}}, && (2.13) \end{aligned}$$

ne zavisi od k. Odnosno, definišemo $H(u)$ tako da je

$${}_u q_{x+k} = H(u) \cdot q_{x+k},$$

za $k \in \mathbf{N}_0$, $u \in [0, 1]$ i neku funkciju H.

Posebno, ukoliko pretpostavimo da je $H(u) = u$ tj. da S ima uniformnu raspodelu na intervalu [0,1], vidimo da važi aproksimacija:

$$\mathring{e}_x \approx e_x + \frac{1}{2}.$$

Pritom, na osnovu nezavisnosti K i S i pretpostavke o raspodeli S, jasno je da je disperzija slučajne veličine T tada jednaka:

$$DT = DK + DS = DK + \frac{1}{12}.$$

Pretpostavimo sada da je godina podeljena na m jednakih delova, gde je m pozitivan ceo broj. Tada je $S^{(m)}$ deo godine koji je osoba preživa u godini smrti, odnosno

$$S^{(m)} = \frac{1}{m}[m \cdot S + 1].m \in \mathbf{N}$$

Kako su po pretpostavci slučajne veličine K i S nezavisne, sledi da su i K i $S^{(m)}$ takođe nezavisne. Dodatno, ukoliko S ima apsolutno neprekidnu uniformu raspodelu na intervalu $[0,1]$, tada $S^{(m)}$ ima diskretnu uniformnu raspodelu na skupu $\{\frac{1}{m}, \frac{2}{m}, ..., \frac{m-1}{m}, 1\}$.

2.2 Tablice smrtnosti

Tablice smrtnosti su jedna od najranije razvijenih metoda za procenu smrtnosti čoveka (ili bilo koje populacije). Prve tablice smrtnosti su u drugoj polovini 17.veka razvili Edmund Halej, odnosno Džon Grant, nezavisno jedan od drugog u Poljskoj, odnosno Engleskoj. Tablice smrtnosti učestvuju u proceni smrtnosti koja je ključna za uspešno i pravilno poslovanje jedne osiguravanjuće kompanije. One nastaju na osnovu empirijskih podataka o smrtnosti populacije kao i različitih procena i ekstrapolacije i nose informaciju o raspodeli slučajne veličine K. Tablice smrtnosti su sastavljene od vrednosti ℓ_x - broja osoba koje su doživele x godina i na osnovu kojih se računaju ostale vrednosti poput q_x - verovatnoće smrtnosti u roku od jedne godine [3]. Osiguravajuće kompanije najčešće formiraju sopstvene tablice.

Tablice smrtnosti se često kreiraju za specifične grupe. Na primer, razlikovaće se tablice za ženski i muški pol, s obzirom da je poznato da se prosečan životni vek razlikuje u odnosu na pol. Isto će važiti i za zanimanja. Odnosno, verovatnoća smrtnosti za osobe iste starosti u slučaju da se jedna od njih bavi poslom koji se smatra visokorizičnim neće biti ista. Štaviše, prema našem zakonu, tokom trajanja perioda pokrića životnog osiguranja, promena zaposlenja je jedina promena koju je osiguranik u obavezi da prijavi osiguravaču [2]. Odvojene tablice smrtnosti mogu da se kreiraju i u odnosu na druge faktore kao što su rasa, socijalni status itd.

Razlikuju se dve vrste tablica smrtonsti:

1. Generacijske (kohort) tablice prate određenu grupu ljudi (kohort jedinki) koje su slične, a najčešće iste starosti. Generacijska tablica smrtnosti sadrži podatke

o broju preživelih, odnosno preminulih članova grupe od rođenja prvog člana pa do trenutka kada je i poslednji član preminuo.

2. Periodične tablice smrtnosti odražavaju smrtnost svih članova određene populacije, bez obzira na njihovu starost, u toku određenog vremenskog perioda.

Jasno je da je teže formirati kohort tablice nego periodične. Svaka od ove dve vrste tablica ima dve podvrste - detaljne i skraćene tablice smrtnosti. Detaljne tablice sadrže podatke o pojedinačnim godinama starnosti, dok skraćene čine podaci dati u okviru intervala, i to najčešće petogodišnji [5].

Sve tablice smrtonsti se formiraju na osnovu sledećih pravila:

Posmatra se ciljna grupa ljudi koji su starosti x i na kraju svake godine se beleži broj ljudi koji je preživeo tu godinu. Koriste se oznake:

- ℓ_x - broj lica koji su napunili x godina,
- $d_x = \ell_x - \ell_{x+1}$ - broj ljudi koji su napunili x godina ali su preminuli pre nego što su napunili $x + 1$. godinu.

Na osnovu ovih vrednosti se mogu izračunati verovatnoće:

- $p_x = \frac{\ell_{x+1}}{\ell_x}$,
- $q_x = \frac{d_x}{\ell_x}$,
- ${}_kq_x = \frac{\ell_x - \ell_{x+k}}{\ell_x}$,
- ${}_kp_x = \frac{\ell_{x+k}}{\ell_x}$.

Još jedan bitan pojam kod tablica smrtnosti je period odabira. Polisa životnog osiguranja se obično sklapa kod osoba koje su povoljnog zdravstenog stanja, i pritom je dodatno postupak sklapanja ugovora često praćen lekarskim pregledom. Na osnovu toga je jasno da će verovatnoća smrtnosti biti manja kod osiguranika koji su skorije sklopili ugovor, odnosno, ukoliko je sa $q_{[x]+t}$ označena verovatnoća smrti lica tokom $x + t + 1$.-ve godine, a koje je ugovor sklopilo sa x godina, važi:

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \dots$$

Međutim, može se intuitivno zaključiti da će nakon određenog broja godina ove verovatnoće biti približno iste. Taj period se naziva **period odabira**. Ukoliko ga označimo sa r, ovo se može zapisati kao:

$$q_{[x-r]+r} = q_{[x-r-1]+r+1} = q_{[x-r-2]+r+2} = ... = q_x,$$

odnosno, jednostavnije zapisano:

$$q_{[x]+t} = q_{x+t},$$

za svako $t \geq r$.

Ovako formirane tablice smrtnosti, gde su se verovatnoće smrtnosti izjednačile nakon perioda odabira, nazivamo **konačnim tablicama smrtnosti** [3].

2.3 Pretpostavke vezane za verovatnoću preživljavanja u delu godine

Kao što je već naglašeno, vrednosti iz tablica smrtnosti u potpunosti određuju raspodelu slučajne veličine K. Za račun se često koristi jednakost:

$$_kp_x = p_x p_{x+1} ... p_{x+k-1}.$$

Međutim, sve jednačine odnose se na celobrojne vrednosti. Nekad je potrebna informacija o smrtnosti samo u određenom delu godine. One se ne mogu dobiti na osnovu tablice tj. za tu svrhu potrebno je uvesti dodatne pretpostavke ukoliko nije poznata raspodela slučajne veličine T. Pretpostavke se tiču verovatoće smrtnosti $_uq_x$ i intenziteta smrtnosti μ_{x+u}, gde je $0 < u < 1$. Postoje tri pretpostavke koje se najčešće koriste:

1. $_uq_x$ je linearna funkcija po u, za $0 < u < 1$, odnosno

$$_uq_x = uq_x.$$

Na osnovu jednakosti (2.5) i poslednice pretostavke ${}_up_x = 1 - uq_x$, važi da je:

$$\begin{aligned}\mu_{x+u} &= -\frac{d}{du}\ln({}_up_x)\\ &= -\frac{1}{{}_up_x}\frac{d}{du}({}_up_x)\\ &= -\frac{1}{1-uq_x}\frac{d}{du}(1-uq_x)\\ &= -\frac{1}{1-uq_x}(-q_x)\\ &= \frac{q_x}{1-uq_x}.\end{aligned}$$

Sada je jasno da se intenzitet smrtnosti može izraziti preko vrednosti iz tablica, te je samim tim određena i raspodela slučajne veličine T. Dodatno, kada se ova pretpostavka uvrsti u (2.13), dobija se da:

$$P\{S \leq u | K = k\} = \frac{{}_uq_{x+k}}{q_{x+k}} = \frac{uq_{x+k}}{q_{x+k}} = u,$$

odakle se može zaključiti da su S i K nezavisne slučajne veličine, i da S ima uniformnu raspodelu na intervalu $[0, 1]$.

2. Intenzitet smrtnosti je konstantan, tj.

$$\mu_{x+u} = \mu_{x+\frac{1}{2}} = const., 0 < u < 1$$

Kada to uvrstimo u (2.6), odatle sledi da:

$$ {}_up_x = e^{-\int_0^u \mu_{x+s}\ \mathrm{ds}} = e^{-\int_0^u \mu_{x+\frac{1}{2}}\ \mathrm{ds}} = e^{-\mu_{x+\frac{1}{2}}\cdot u}. \tag{2.14}$$

Ukoliko se zameni vrednost $u = 1$:

$$p_x = e^{-\mu_{x+\frac{1}{2}}},$$

te važi:

$$\mu_{x+\frac{1}{2}} = -\ln p_x.$$

Kada ovu vezu uvrstimo u (2.5) jasno je da:

$${}_up_x = \left(e^{-\mu_{x+\frac{1}{2}}}\right)^u = p_x^u,$$

pa i ovu vrednost možemo dobiti iz tablica smrtnosti. Kao i za prethodnu pretpostavku, kada primenimo dobijeno na (2.13), važi:

$$P\{S \leq u|K = k\} = \frac{{}_uq_{x+k}}{q_{x+k}} = \frac{1 - {}_up_{x+k}}{1 - p_{x+k}} = \frac{1 - (p_{x+k})^u}{1 - p_{x+k}}. \tag{2.15}$$

U ovom slučaju, za razliku od prve pretpostavke, uslovna raspodela od S zavisi od k, te S i K nisu nezavisne, i S ima zasečenu eksponencijalu raspodelu. Ova pretpostavka se koristi radi pojednostavljenja računa i ukoliko postoji razlog da se veruje da je u konkretnom slučaju opravdana. Međutim, može se diskutovati i koliko je sama pretpostavka realistična, tj. da li je opravdana u opštem slučaju. Intenzitet smrtnosti će biti konstantan u toku godine ukoliko pretpostavimo da je i verovatnoća smrtnosti u toku godine konstantna. Kao što ćemo pokazati u nastavku, u praksi ovo često nije slučaj i u toku godine se pojavljuju uočljive sezonske varijacije.

3. Balducci pretpostavka. ${}_{1-u}q_{x+u}$ je linearna funkcija za $0 < u < 1$, odnosno:

$${}_{1-u}q_{x+u} = (1 - u)q_x.$$

Iz jednakosti (2.1) sledi:

$${}_up_x = \frac{p_x}{{}_{1-u}p_{x+u}} = \frac{1 - q_x}{1 - {}_{1-u}q_{x+u}} = \frac{1 - q_x}{1 - (1 - u)q_x}, \tag{2.16}$$

te odatle dobijamo raspodelu za T.

Sada može da se dobije i izraz za intenzitet smrtnosti:

$$\begin{aligned}
\mu_{x+u} &= -\frac{d}{du}\ln({}_up_x) \\
&= -\frac{d}{du}\ln(\frac{1 - q_x}{1 - (1 - u)q_x}) \\
&= -\frac{1 - (1 - u)q_x}{1 - q_x}\frac{-(1 - q_x) \cdot q_x}{(1 - (1 - u)q_x)^2} \\
&= \frac{q_x}{1 - (1 - u)q_x}.
\end{aligned}$$

Na osnovu (2.13):

$$P\{S \leq u|K = k\} = \frac{{}_u q_{x+k}}{q_{x+k}} \tag{2.17}$$

$$= \frac{1 - {}_u p_{x+k}}{q_{x+k}} \tag{2.18}$$

$$= \frac{1 - \frac{1-q_x}{1-(1-u)q_x}}{q_{x+k}} \tag{2.19}$$

$$= \frac{u}{1-(1-u)q_x}, \tag{2.20}$$

pa možemo da zaključimo da su i pod ovom pretpostavkom K i S zavisne, kao i da S nema uniformnu raspodelu.
Izuzetno, u slučaju da je verovatnoća q_{x+k} izuzetno mala, iz jednakosti (2.15) i (2.20), može da se zaključi da S ima približno uniformnu raspodelu i da su S i K (približno) nezavisne slučajne veličine.

Glava 3

Sezonske varijacije

U prethodom poglavlju je bilo reči o mogućim metodama za određivanje smrtnosti u delu godine. To su bile pretpostavke vezane za verovatnoću i intenzitet smrtnosti u toku vremenskog perioda u, $0 < u < 1$.

S obzirom da je T apsolutno neprekidna slučajna veličina, određivanjem njene raspodele se takođe određuje i smrtnost u vremenskim periodima koji su kraći od jedne godine. Intuitivno je jasno da postoji težnja da se smrtnost izrazi analitički. U poređenju sa pretpostavkama vezanim za tablice smrtnosti, ovaj princip pre svega deluje praktičnije, jer je jasno da je lakše izraziti verovatnoću preko funkcije sa nekoliko parametara nego preko tablice smrtnosti. Dodatno, dati parametri se mogu oceniti pomoću podataka o smrtnosti.
Neke od predloženih funkcija u knjizi [3] predstavljaju takozvani:

- De Muavrov zakon. Neka je w gornja granica starosti čoveka. Tada slučajna veličina T ima uniformnu raspodelu $T \sim U[0, w-x]$, gde je x pristupna starost. Funkcija gustine raspodele je:

$$g(t) = \frac{1}{w-x}. \quad 0 < t < w - x$$

  Dodatno, funkcije preživljavanja i intenziteta smrtnosti su jednake

$$s(t) = 1 - G(t) = 1 - \frac{t}{w-x},$$

$$\mu_{x+t} = \frac{g(t)}{1-G(t)} = \frac{\frac{1}{w-x}}{1 - \frac{t}{w-x}} = \frac{1}{w-x-t}. \quad 0 < t < w - x.$$

- Gompercov zakon. Funkcija intenziteta smrtnosti je:

$$\mu_{x+t} = B \cdot c^{x+t}. \quad \text{za} \quad t > 0 \quad \text{i konstante} \quad B, c > 0.$$

- Mejhemov zakon. Takođe se definiše preko funkcije intenziteta smrtnosti:

$$\mu_{x+t} = A + B \cdot c^{x+t}. \quad \text{za} \quad t > 0 \quad \text{i} \quad \text{konstante} \quad A, B, c > 0.$$

  Može se uočiti da je ovo uopštenje Gompercovog zakona, tj. da je Gompercov zakon specijalan slučaj Mejhemovog kada je $A = 0$.

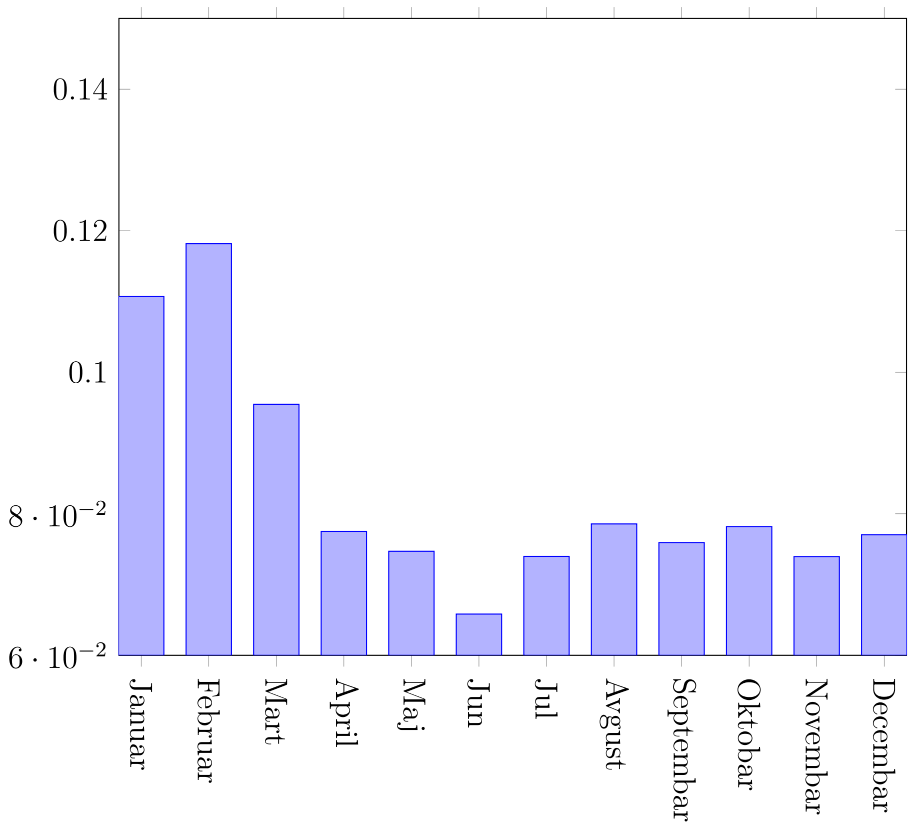
- Vejbulov zakon. Pretpostavlja se da:

$$\mu_{x+t} = k \cdot (x+t)^n. \quad \text{za} \quad t > 0 \quad \text{i} \quad \text{konstante} \quad k, n > 0.$$

Pristup koji je opisan se, međutim, ne smatra adekvatnim za primenu na realnim podacima. Jedan od razloga je što ne može na pravi način da prikaže pravilnosti u smrtnosti koje se javljaju u toku godine.

U svrhu uočavanja pomenutih pravilnosti, posmatramo podatke o smrtosti po mesecima u godini, na teritoriji Srbije, u toku 2022. godine. Podaci su preuzeti sa sajta Republikog zavoda za statistiku i mogu se videti na linku https://data.stat.gov.rs/Home/Result/18030306 [14].

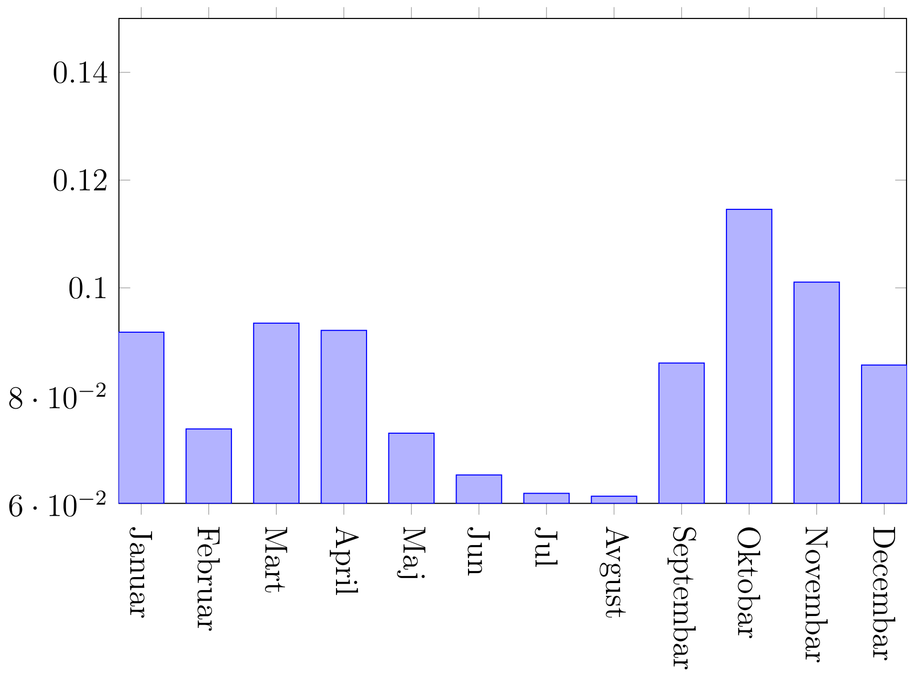
Na slici 3.1 je dat grafički prikaz ovih podataka:



Slika 3.1: Smrtnost na teritoriji Srbije 2022. godine po mesecima

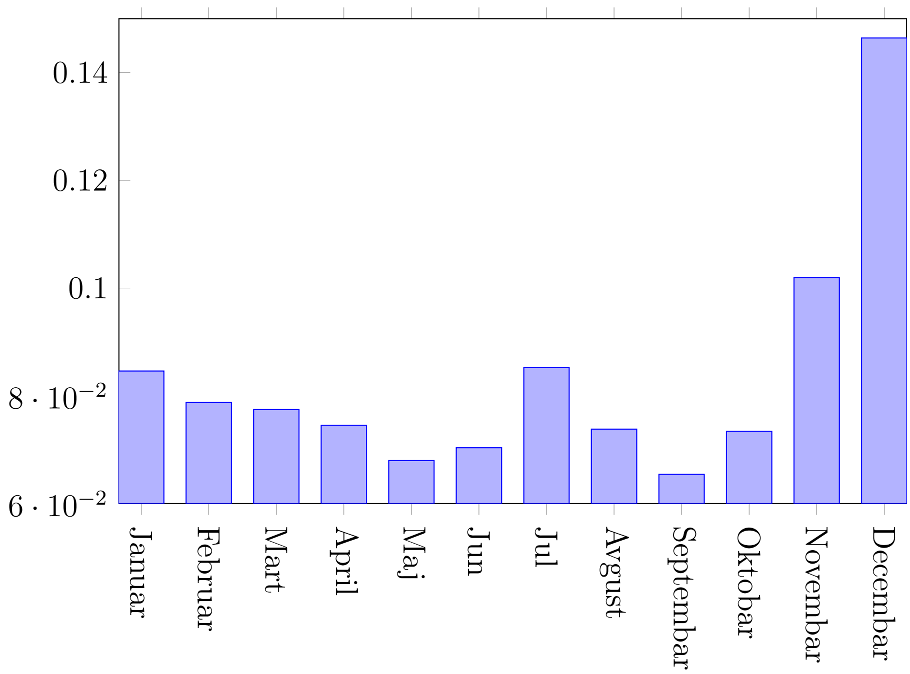
,

Na graficima 3.2, 3.3, 3.4 i 3.5 vidimo podatke iz prethodnih nekoliko godina:



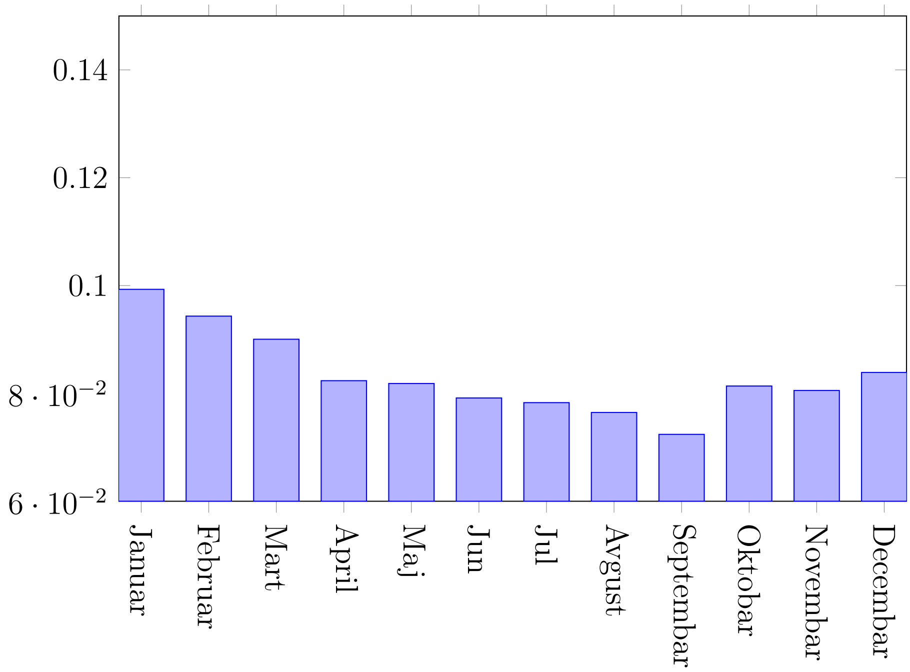
Slika 3.2: Smrtnost na teritoriji Srbije 2021. godine po mesecima

,



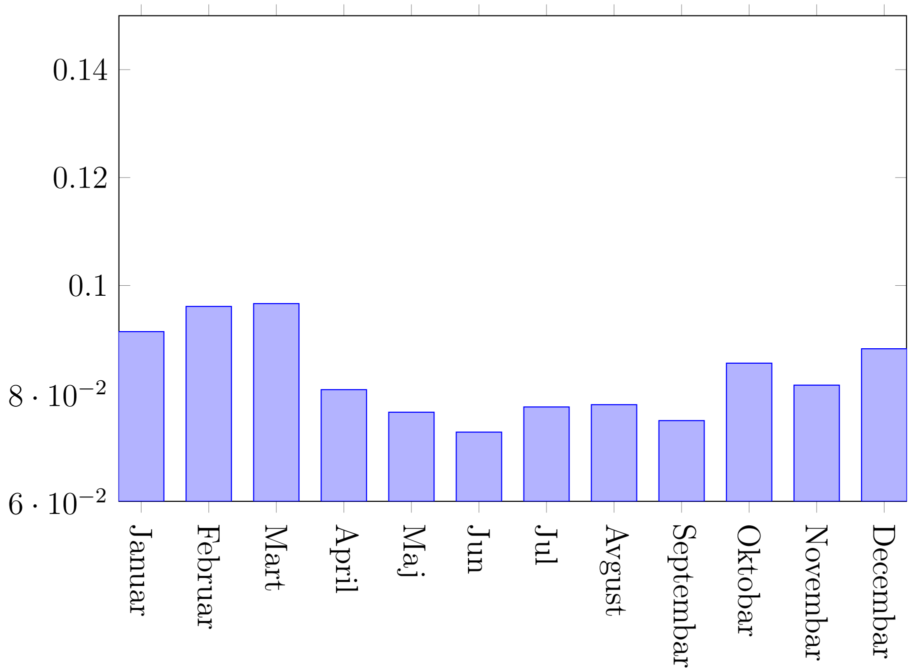
Slika 3.3: Smrtnost na teritoriji Srbije 2020. godine po mesecima

,



Slika 3.4: Smrtnost na teritoriji Srbije 2019. godine po mesecima

,



Slika 3.5: Smrtnost na teritoriji Srbije 2018. godine po mesecima

,

Jasno je da predložene funkcije neće moći na pravi način da oslikaju varijabilnost broja smrti u toku godine. Uočava se da se najveća smrtnost javlja u zimskim mesecima, i ta pravilnost je razlog zašto se traži alternativni pristup u opisivanju sezonskih varijacija.

3.1 Periodične slučajne veličine

Uvodimo pojam tzv. **pretpostavke sezonske smrtnosti** (eng. seasonal mortality assumption). Ona podrazumeva korišćenje periodične slučajne veličine za prikaz smrtnosti u toku godine [6]. Periodične (cirkularne) slučajne veličine se koriste za modeliranje pojava koje se realizuju u trenucima sa nekim periodom. Na primer, tu spadaju realizacije događaja u odnosu na dan u nedelji ili mesec u godini.

U radu [7] predstavljena je familija funkcija raspodele koju može imati periodična slučajna veličina, kao i uslovi koje je neophodno da ona ispunjava. Funkcija gustine verovatnoće $f(\theta, \mathbf{c})$ periodične slučajne veličine Θ mora da zadovoljava uslove:

1. $f(\theta, \mathbf{c}) \geq 0$,
2. $\int_0^{2\pi} f(\theta, \mathbf{c})\ \mathrm{d}\theta = 1$,
3. $f(\theta + 2k\pi; \mathbf{c}) = f(\theta; \mathbf{c})$ za celobrojnu vrednost k,

gde je $\mathbf{c}$ vektor parametara.

Ove raspodele su zasnovane na tzv. nenegativnim trigonometrijskim sumama (skraćeno NNTS) [7].

Zbog toga navodimo sledeću teoremu ([8]):

Teorema. Trigonometrijski polinom reda n:

$$T(\theta) = a_0 + \sum_{k=1}^{n}(a_k \cos k\theta + b_k \sin k\theta),$$

je za svako realno θ nenegativan ako i samo ako postoje kompleksni brojevi z_k, $k = 0, 1, 2, ...$ takvi da:

$$a_k - i \cdot b_k = 2\sum_{v=0}^{n-k} z_{v+k}\bar{z}_v,$$

kao i

$$a_0 = \sum_{k=0}^{n} |z_k|^2.$$

Dokaz ove teoreme se može pronaći u okviru rada [9].

Ukoliko se dodatno nametne uslov:

$$\sum_{k=0}^{n} |z_k|^2 = \frac{1}{2\pi},$$

tada je izrazom:

$$f(\theta, \mathbf{c}, M) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{M}(a_k \cos k\theta + b_k \sin k\theta),$$

uvedena funkcija koja je nenegativna i čiji je integral na prostoru verovatnoće jednak jedan (što omogućava poslednji uslov), odnosno, to zaista predstavlja jednu funkciju gustine.

Polinom $T(\theta)$ možemo zapisati i u obliku:

$$\begin{aligned} T(\theta) &= \left\| \sum_{k=0}^{M} c_k e^{ik\theta} \right\|^2 \\ &= \sum_{k=0}^{M} c_k e^{ik\theta} \overline{\left(\sum_{k=0}^{M} c_k e^{ik\theta} \right)} \\ &= \sum_{k=0}^{M} c_k e^{ik\theta} \sum_{k=0}^{M} \bar{c_k} e^{-ik\theta} \\ &= \sum_{k=0}^{M}\sum_{m=0}^{M} c_k \bar{c_m} e^{ik\theta} e^{-im\theta} \\ &= \sum_{k=0}^{M}\sum_{m=0}^{M} c_k \bar{c_m} e^{i(k-m)\theta}, \end{aligned}$$

koji je praktičniji za korišćenje prilikom aktuarskih obračuna.

Neka je sada N vektor koji sadrži broj smrti u svakom mesecu, i neka važi pretpostavka da su $K(x)$ i $S(x)$ međusobno nezavisne. U radu [7] je predloženo da, ukoliko je $\mathbf{c} = (c_0, ..., c_M)$ vektor M parametara, za funkciju maksimalne verodostojnosti se tada uzima $L(\mathbf{c}|M, N)$, definisana sa:

$$L(\mathbf{c}|M, N) = \prod_{r=1}^{12}(F(2\pi u_r; \mathbf{c}, M) - F(2\pi u_{r-1}; \mathbf{c}, M))^{N_r},$$

gde je $N = (N_1, ..., N_{12})$, M red trigonometrijske sume, u_r je deo godine koji odgovara periodu zaključno sa r-tim mesecom (npr. $u_1 = \frac{1}{12}$ i $u_{12} = 1$), dok je F funkcija raspodele periodične slučajne veličine.

Paket CircNNTSR u R-u sadrži razne funkcije koje su korisne za modelovanje podataka vezanih za smrtnost po mesecima u godini [10]. U ovom radu koriščena je funkcija `nntsmanifoldnewtonestimationinterval0to1` koja pomoću Njutnovog algoritma računa ocenu parametara NNTS metodom maksimalne verodostojnosti. Neophodno je da se kao argument funkciji prosledi vrednost M, tj. broj parametara čiju ocenu funkcija treba da izračuna. Na osnovu grafika na slici 3.6 može da se zaključi da slučaj $M = 4$ najviše odgovara podacima, međutim, tada rizikujemo preprilagođavanje modela. Istu funkciju možemo da primenimo i na grafik 3.7 smrtnosti po mesecima, kumulativno za period od 2015. do 2022. godine (koji je reprezentativniji), gde uočavamo jasniju pravilnost. U ovom slučaju smo koristili $M = 2$. U tabeli 3.1 su prikazani modeli sa NNTS reda 1,2,3 i 10 i njihove AIC (Akaike information criterion) i BIC (Bayesian information criterion) vrednosti. U tabeli vidimo da vrednosti AIC i BIC opadaju sa porastom broja parametara, ali da ne postoji velika razlika između manjih i većih vrednosti M. Na sličan način se ponašaju i krive koje predstavljaju ocenu funkcije gustine za različite vrednosti M, tj. ne uočavaju se značajne razlike između ocena.

Tabela 3.1: AIC i BIC vrednosti za različite modele

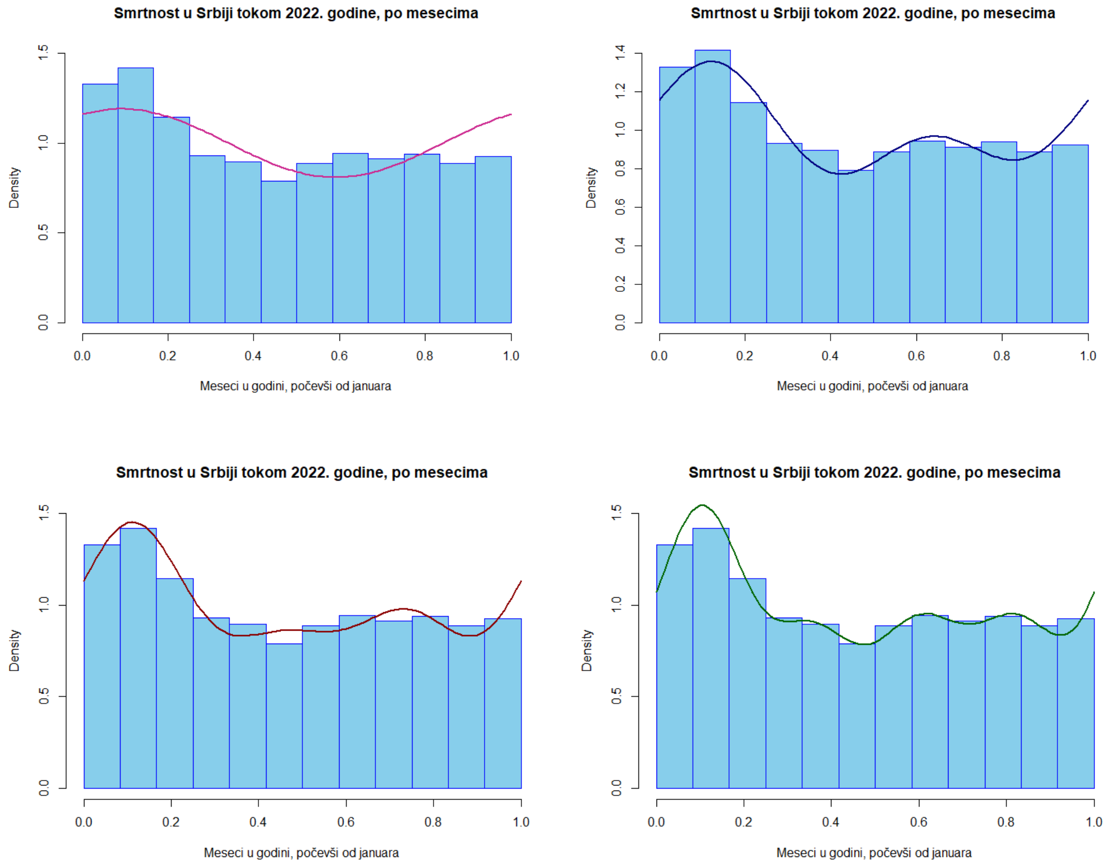
M	AIC	BIC
1	4,335,380	4,335,381
2	4,335,205	4,335,207
3	4,335,119	4,335,122
10	4,334,920	4,334,929

Jedna od poznatijih familija funkcija raspodele koje se koriste kod periodičnih slučajnih veličina je von Mises-ova familija. Ukoliko je Θ, $\Theta \in [0, 2\pi)$ slučajna veličina sa von Mises-ovom raspodelom, tada je njena funkcija gustine raspodele:

$$f(\theta, \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)},$$

gde su parametri μ - teorijska srednja vrednost ($\mu \in [0, 2\pi)$), a κ teorijska mera koncentracije u smeru μ. I_0 je Beselova funkcija prve vrste, reda nula. U slučaju da je $k = 0$, funkcija raspodele se svodi na uniformnu raspodelu na jedničnom krugu.

Sada možemo da probamo da ocenimo parametre na osnovu naših kumulativnih podataka (od 2015. godine do 2022. godine). Ugrađena funkcija u R-u `mle.vonmises()`

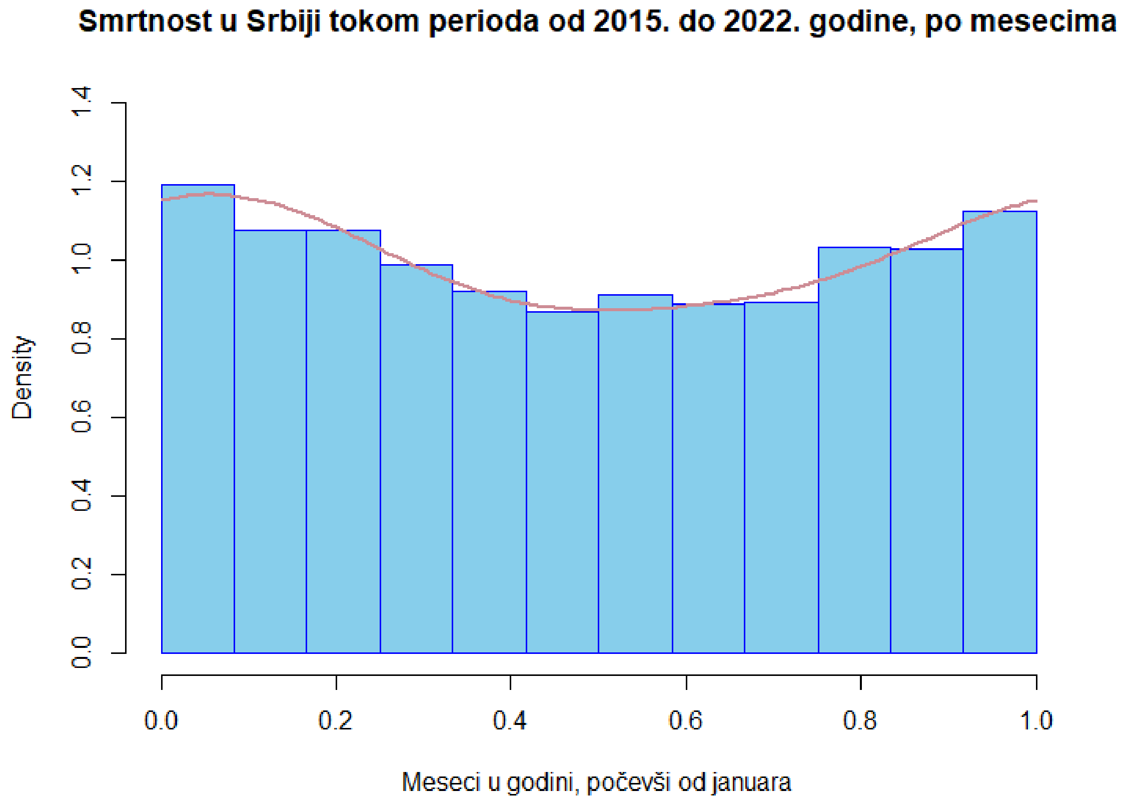


Slika 3.6: Modeli koje daje ugrađena funkcija u R-u za različite vrednosti M; M=1 (gore levo), M=2 (gore desno), M=3 (dole levo), M=4 (dole desno).

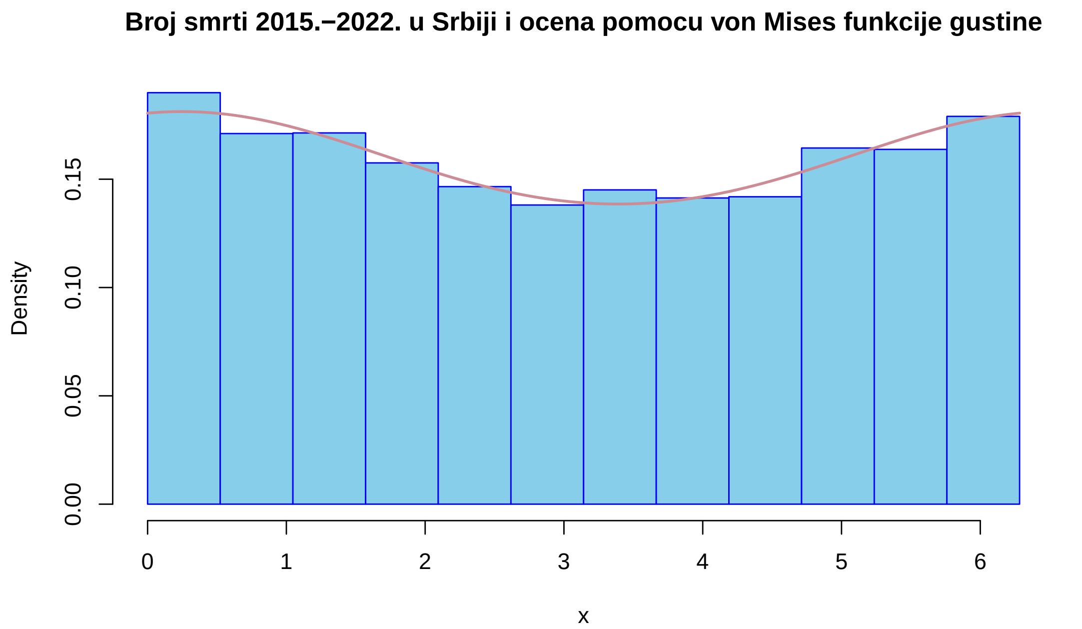
daje ocenu parametara μ i κ maksimizacijom funkcije verodostojnosti, definisanom sa:

$$L(\mu,\kappa) = \frac{1}{(2\pi I_0(\kappa))^n} e^{\kappa\cdot\sum_{k=1}^{n} cos(\theta_i-\mu)}.$$

Na osnovu dobijenih vrednosti za parametre $\mu = 0.2472$ i $\kappa = 0.1343$, možemo nacrtati funkciju gustine i uporediti sa podacima o smrtnosti, što je prikazano na grafiku 3.8:



Slika 3.7: Mortalitet u Srbiji u vremenskom periodu od 2015. do 2022. godine, od januara do decembra svake godine, sa ocenom za vrednost parametra M=2.



Slika 3.8: Ocena pomoću von Mises-ove funkcije gustine raspodele

Uopštenje ove raspodele predstavlja von Mises-Fisher-ova raspodela. Neka je $\mathbf{X}$ slučajni vektor iz ove raspodele, tako da $\mathbf{X} \in \mathbb{R}^p$ i $\|\mathbf{X}\| = 1$. Tada je funkcija gustine raspodele definisana sa:

$$f(\mathbf{x} \mid \boldsymbol{\mu}, \kappa) = C_p(\kappa) e^{\kappa \boldsymbol{\mu}^\top \mathbf{x}},$$

gde je $\boldsymbol{\mu}$ vektor teorijske srednje vrednosti, κ teorijska mera koncentracije u smeru $\boldsymbol{\mu}$, a funkcija $C_p(\kappa)$ definisana sa:

$$C_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)},$$

funkcija koja obezbeđuje da je integral preko hipersfere jednak jedan, gde figuriše i Beselova funkcije reda $\frac{p}{2} - 1$.

Glava 4

Primena modela zasnovanih na periodičnim slučajnim veličinama i analizi vremenskih serija u životnom osiguranju

Glavni razlog za potrebu modelovanja smrtnosti za period kraći od jedne godine je postojanje vrsta osiguranja koje pružaju pokriće sa trajanjem do jedne godine.

Možemo da primenimo spomenute zaključke prilikom računanja jednokratne premije kod privremenog osiguranja. Premija u životnom osiguranju se računa kao očekivana vrednost osigurane sume koja će biti isplaćena nakon realizacije osiguranog slučaja. Radi jednostavnosti se pretpostavlja da je osigurana suma jedna novčana jedinica, te se računa očekivanje diskontovane vrednosti broja 1, dok se u praksi množi sa odgovarajućim iznosom. Osigurana suma se diskontuje na osnovu trenutka koji predstavlja realizaciju slučajne veličine T. Ukoliko se pretpostavi da se kamaćenje vrši neprekidno, jednokratna premija iznosi:

$$\bar{A}^{1\;\;b}_{x:\overline{1}|} = E[e^{-\delta T(x)}] = \int_0^1 e^{-\delta t} f^b_{T(x)}(t)\ \mathrm{d}t,$$

gde je $f^b_{T(x)}$ funkcija gustine verovatnoće slučajne veličine $T(x)$, a b mesec u kom je rođena osoba. Navedena oznaka je simbol za iznos jednokratne premije kod privremenog neprekidnog osiguranja čiji je period pokrića 1 godina, a pristupna starost osiguranika x godina.

Ukoliko godinu predstavimo kao sumu 12 meseci, prethodni izraz možemo da zapišemo i u obliku:

$$\bar{A}^{1\ \ b}_{x:\overline{1}|} = \sum_{h=0}^{11} \int_{\frac{h}{12}}^{\frac{h+1}{12}} e^{-\delta t} f^b_{T(x)}(t)\ \mathrm{d}t.$$

Tada može da se uvede smena $t = \frac{h}{12} + s$, tako da je $0 < s < 1$ i $h = 0, 1, ..., 11$, odnosno:

$$\begin{aligned} \bar{A}^{1\ \ b}_{x:\overline{1}|} &= \sum_{h=0}^{11} \int_{0}^{\frac{1}{12}} e^{-\delta\left(\frac{h}{12}+s\right)} f^b_{T(x)}\left(\frac{h}{12} + s\right)\ \mathrm{d}s && (4.1) \\ &= \sum_{h=0}^{11} e^{-\delta\frac{h}{12}} P\{K(x) = 0\} \int_{0}^{\frac{1}{12}} e^{-\delta s} f^b_{S(x)|K(x)=0}\left(\frac{h}{12} + s \middle| K(x) = 0\right)\ \mathrm{d}s. && (4.2) \end{aligned}$$

Alternativno, ovu premiju možemo zapisati i u obliku:

$$\bar{A}^{1\ \ b}_{x:\overline{1}|} = \sum_{h=0}^{11} e^{-\delta\frac{h}{12}}\, {}_{\frac{h}{12}}p_x^{\ b}\, \bar{A}^{1\ \ \ \ \ b}_{x+\frac{h}{12}:\overline{\frac{1}{12}}|}. \qquad (4.3)$$

Tada, na osnovu (4.2) i (4.3) važi:

$$\bar{A}^{1\ \ \ \ \ b}_{x+\frac{h}{12}:\overline{\frac{1}{12}}|} = \frac{q_x}{{}_{\frac{h}{12}}p_x^{\ b}} \int_{0}^{\frac{1}{12}} e^{-\delta s} f^b_{S(x)|K(x)=0}\left(\frac{h}{12} + s \middle| K(x) = 0\right)\ \mathrm{d}s, \qquad (4.4)$$

odnosno, dobija se iznos jednokratne premije za osiguranje koje je uplatila osoba starosti $x + \frac{h}{12}$ i čije pokriće obuhvata narednih mesec dana.

4.1 Modelovanje sezonske smrtnosti primenom periodičnih slučajnih veličina

Sada je potrebno odrediti gustinu koja figuriše u izrazu (4.4). Neka je Θ periodična slučajna veličina koja uzima vrednosti u radijanima, iz intervala $[0, 2\pi]$. Zato se uvodi smena za slučajnu veličinu S, tako da $S = \frac{\Theta}{2\pi} \in [0, 1]$ i važi:

$$f_S(s) = f_\Theta(\theta) \left| \frac{d\Theta}{dS} \right| = f_\Theta(\theta) 2\pi = 2\pi f_\Theta(2\pi s).$$

Kako nastojimo da uključimo vrednost b u izraz za gustinu i na osnovu treće osobine funkcije gustine verovatnoće kružne slučajne veličine sledi:

$$f_S^b(s) = 2\pi f_\Theta(2\pi(s + b)).$$

Tada sledi

$$\begin{aligned} f^b_{S(x)|K(x)=0}(s|K(x) = 0) &= 2\pi f_\Theta(2\pi(s + b)) && (4.5) \\ &= 2\pi \sum_{k=0}^{M} \sum_{m=0}^{M} c_k \bar{c_m} e^{i(k-m)2\pi(s+b)} && (4.6) \\ &= 2\pi \left[\frac{1}{2\pi} + \sum_{k=1}^{M} \sum_{m=1, m\neq k}^{M} c_k \bar{c_m} e^{i(k-m)2\pi(s+b)} \right] && (4.7) \\ &= 1 + 2\pi \sum_{k=1}^{M} \sum_{m=1, m\neq k}^{M} c_k \bar{c_m} e^{i(k-m)2\pi(s+b)}. && (4.8) \end{aligned}$$

Ukoliko jednakost (4.8) uvrstimo u integral iz jednakosti (4.4):

$$\begin{aligned} &\int_0^{\frac{1}{12}} e^{-\delta s} f^b_{S(x)|K(x)=0} \left(\frac{h}{12} + s | K(x) = 0 \right) \mathrm{d}s \\ &= \int_0^{\frac{1}{12}} e^{-\delta s} \left[1 + 2\pi \sum_{k=1}^{M} \sum_{m=1, m\neq k}^{M} c_k \bar{c_m} e^{i(k-m)2\pi(\frac{h}{12}+s+b)} \right] \mathrm{d}s \\ &= \int_0^{\frac{1}{12}} e^{-\delta s} \mathrm{d}s + 2\pi \sum_{k=1}^{M} \sum_{m=1, m\neq k}^{M} c_k \bar{c_m} e^{i(k-m)2\pi(\frac{h}{12}+b)} \int_0^{\frac{1}{12}} e^{i(k-m)2\pi s - \delta s} \mathrm{d}s \\ &= \frac{1}{\delta} \left(1 - e^{-\delta \frac{1}{12}} \right) + 2\pi \sum_{k=1}^{M} \sum_{m=1, m\neq k}^{M} c_k \bar{c_m} e^{i(k-m)2\pi(\frac{h}{12}+b)} \left[\frac{1 - e^{(i(k-m)2\pi - \delta)\frac{1}{12}}}{\delta - i(k-m)2\pi} \right]. \end{aligned}$$

Takođe, iz (4.4) treba odrediti i vrednost ${}_{\frac{h}{12}}{p_x}^b$:

$$\begin{aligned}
{}_{\frac{h}{12}}{p_x}^b &= 1 - \int_0^{\frac{h}{12}} f^b_{T(x)}(t)\mathrm{d}t \\
&= 1 - q_x \int_0^{\frac{h}{12}} f_{S(x)|K(x)=0}(s+b)\mathrm{d}t \\
&= 1 - q_x \left[\frac{h}{12} + 2\pi \sum_{k=1}^{M} \sum_{m=1, m\neq k}^{M} c_k \bar{c_m} e^{i(k-m)2\pi b} \left(\frac{1 - e^{i(k-m)2\pi \frac{h}{12}}}{-i(k-m)2\pi} \right) \right].
\end{aligned}$$

Vrednosti $c_k, k = 1, 2, ..., M$ možemo dobiti iz funkcije `nntsmanifoldnewtonestimationinterval0to1` iz R-a. Kada smo računali ove podatke za mortalitet u Srbiji, ta funkcija je za $M = 2$ vratila vrednosti koje su prikazane u tabeli 4.1.

Tabela 4.1: Ocene parametara c_k

k	$\bar{c_k}$
0	0.397861673
1	0.027897577-0.008063659·i
2	0.002448305-0.003426431·i

Sada se može izračunati vrednost jednokratne premije koju uplaćuje osoba strosti $x + \frac{h}{12}$, za pokriće u trajanju od mesec dana, ukoliko je rođena u b-tom mesecu, na osnovu jednačine (4.4). Primera radi, ukoliko osoba ženskog pola starosti 54 godine i 6 meseci, rođena u martu, želi da uplati privremeno osiguranje sa osiguranom sumom od 1000 novčanih jedinica, čije pokriće obuhvata narednih mesec dana, tada je iznos mesečne premije pri neprekidnom kamaćenju sa stopom 5% jednak 0.4330. U obračunu je korišćena vrednost za verovatnoću smrtnosti iz zvaničnih tablica mortaliteta Republike Srbije.

Još jedna provera adekvatnosti modela može da bude test saglasnosti. χ^2 test saglasnosti poredi očekivanu frekvenciju sa stvarnom, i to na osnovu statistike:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i},$$

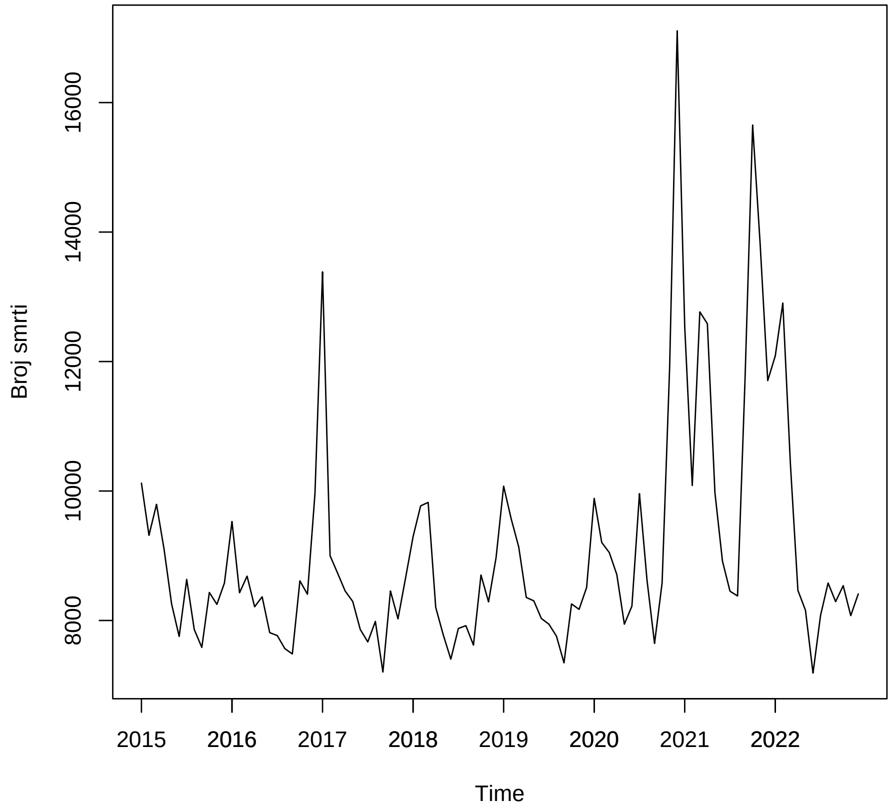
gde je O_i stvarna, a E_i očekivana frekvencija za i-tu kategoriju. Pored ovih vrednosti, računa se i broj stepeni slobode kao broj kategorija - 1. Vrednost χ^2 se koristi za izračunavanje vrednosti funkcije χ^2 raspodele koja odgovara tom kvantilu i dobijenom broju stepeni slobode. Dodatno, broj opservacija bi trebao da bude veliki jer se radi o graničnoj raspodeli. U ovom radu je korišćena ugrađena funkcija `chisq.test`. Primenom na kumulativnim podacima o smrtnosti u Srbiji dobija se p vrednost $2.2 \cdot 10^{-16}$. Međutim, za velike uzorke kao što je naš (ukupno 874.022 smrti), ovaj test može da proizvede rezultate koji nisu verodostojni, jer u tim slučajevima daje veliki značaj malim odstupanjima. Tada se može koristiti Cohen-ova ω veličina efekta koja uzima u obzir uticaj veličine uzorka na χ^2 statistiku. Formula po kojoj se računa njena vrednost prilikom poređenja teorijske raspodele i kategoričke promenljive je:

$$\omega = \sqrt{\frac{\chi^2}{n}},$$

gde je χ^2 vrednost statistike koja je izračunata u prethodnom koraku, a n veličina uzorka. Kako je vrednost statistike na našim podacima 0.0328, odnosno bliska nuli, to ukazuje na male suštinske razlike, pa može da se zaključi da podaci odgovaraju izabranoj raspodeli [11] . Ukoliko primenimo χ^2 test saglasnosti na nasumično izabran podskup početnog uzorka veličine 10.000, dobija se p vrednost 0.1542, pa zaključujemo da su podaci saglasni sa NNTS raspodelom. Ovim je u potpunosti opisan predlog modela za procenu smrtnosti i računanje premije u delu godine koji je predstavljen u radu [6].

4.2 Modelovanje sezonske smrtnosti primenom metoda analize vremenskih serija

Sada ćemo pokušati da za ocenu smrtnosti iskoristimo neke od modela karakterističnih za rad sa vremenskim serijama. Koristimo iste podatke o smrtnosti u Srbiji po mesecima kao i u prethodnom poglavlju. Od njih kreiramo vremensku seriju čiji je početak 2015. godina a kraj 2022. godina. Ta vremenska serija je prikazana na slici 4.1:

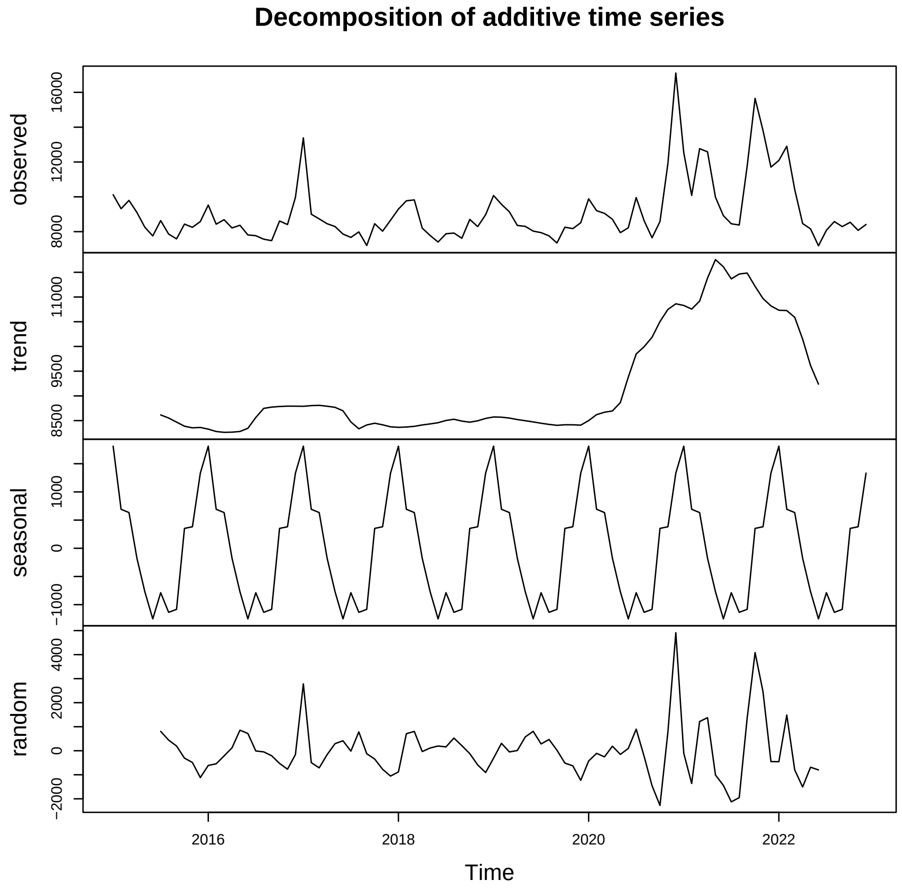


Slika 4.1: Vremenska serija koja prikazuje broj smrti u Republici Srbiji od 2015. do 2022. godine

Na grafiku je prikazano kretanje broja smrtu u svakom mesecu od početka do kraja vremenskog perioda koji posmatramo. Ideja je da na osnovu ovih podataka izvučemo zaključak o ponašanju smrtnosti tokom jedne godine koji ćemo onda moći da koristimo kao univerzalnu ocenu za bilo koju godinu.

Na grafiku se jasno uočava sezonska komponenta. Možemo da vidimo da postoji skok u januaru 2017., kao i generalno povećana smrtnost u periodu nakon početka pandemije Korona virusa. Ove promene nećemo smatrati relevantnim jer su rezultat događaja koji je redak, iznanadan, i ne može se predvideti niti modelovati.

Sa grafika je jasno da se pre radi o aditivnom, nego o multiplikativnom modelu, jer sezonska komponenta ne raste sa trendom. Na slici 4.2 posmatramo komponente ove vremenske serije:



Slika 4.2: Dekompozicija vremenske serije

Izražena je sezonska komponenta, i ne postoji uočljivi trend.
Podatke o smrtnosti koje smo koristili i prilikom pravljenja modela zasnovanog na periodičnim slučajnim veličinama, koristićemo za sve modele u nastavku.

Modeli vezani za analizu vremenskih serija

- **Regresija**

  Prilikom analize vremenske serije moguće je korisititi linearnu regresiju. Za model vremenske serije $\{x_t : t = 1, ..., n\}$ kažemo da je linearan ako vrednost koju uzima u trenutku t možemo zapisati u obliku:

$$x_t = \alpha_0 + \alpha_1 u_{1,t} + ... + \alpha_m u_{m,t} + z_t,$$

gde je $u_{i,t}$ i-ta nezavisna promenljiva vezana za trenutak t, a z_t greška. Reziduali sami čine jednu vremensku seriju $\{z_t\}$ i za njih mora da važi da:

- nisu međusobno korelisani,
- funkcija srednje vrednosti je jednaka nuli,
- imaju konstantnu disperziju, tj. homoskedastični su,
- imaju normalnu raspodelu.

Parametri α_i su zapravo oni čiju ocenu određujemo, i to metodom najmanjih kvadrata. Korišćenje regresije kod vremenskih serija se razlikuje u odnosu na klasičnu regresiju jer su reziduali često međusobno korelisani i sami čine vremensku seriju.

Ocena se formira minimiziranjem sume kvadrata reziduala:

$$\sum_{t=1}^{n} z_t^2 = \sum_{t=1}^{n} (x_t - \alpha_0 - \alpha_1 u_{1,t} - ... - \alpha_m u_{m,t})^2. \tag{4.9}$$

Konkretno, u našem slučaju, gde postoji izražena sezonska komponenta sa periodom 12, oblik modela koji odgovara podacima je:

$$x_t = \begin{cases} \alpha_1 t + \beta_1 + z_t, & t = 1, 13, \dots \\ \alpha_1 t + \beta_2 + z_t, & t = 2, 14, \dots \\ \vdots & \\ \alpha_1 t + \beta_{12} + z_t, & t = 12, 24, \dots, \end{cases}$$

gde su β_i parametri koji se ocenjuju i koji odgovaraju pojedinačnom mesecu u godini.

Dakle, kao nezavisne promenljive uzimamo mesece u godini. Funkcija `cycle` u R-u izdvaja sezonske faktore iz vremenske serije tako što za svaku vrednost vremenske serije vraća njenu poziciju u okviru sezonskog ciklusa, i na taj način omogućuje da se oni koriste kao nezavisne promenljive. Vrednosti R^2 i prilagođenog R^2 su, redom, 0.9962 i 0.9952, što znači da podaci jako dobro odgovaraju ovom modelu. Prilagođeni R^2 predstavlja modifikaciju R^2 prema formuli:

$$\bar{R^2} = 1 - \left(\frac{(1 - R^2)(n-1)}{n-k-1} \right),$$

gde je n broj opservacija a k broj prediktora, odnosno nezavisnih promenljivih. Korišćenjem prilagođenog R^2 dobija se indikator da li se u modelu koristi bespotrebno veliki broj prediktora koji ne doprinose modelu već ga samo usložnjavaju, što vidimo da ovde nije slučaj. Predviđanje budućih vrednosti na osnovu modela pomoću ugrađene funkcije `predict` je prikazano na slici 4.3, te smo dobili projekciju broja smrti u bilo kom delu godine. Formula koja je korišćena za predviđanje na osnovu linearne regresije je:

$$\hat{x_t} = \hat{\alpha_1} \cdot t + \hat{\beta_i}$$

za $i = 1, 2, ..., 12$.

Međutim, proverom grafika autokorelacione funkcije reziduala uočavaju se odstupanja od intervala poverenja, što znači da su reziduali međusobno korelisani. Značajne korelacije su prisutne i nakon primene sezonskog diferenciranja. Možemo da zaključimo da je odgovarajući model verovatno ipak nešto složeniji, pa ima smisla da istražimo i druge mogućnosti.

- **Holt-Winters model**

  Holt-Winters model je karakterističan za analizu vremenskih serija. Jednačine aditivnog Holt-Winters modela u trenutku t su oblika:

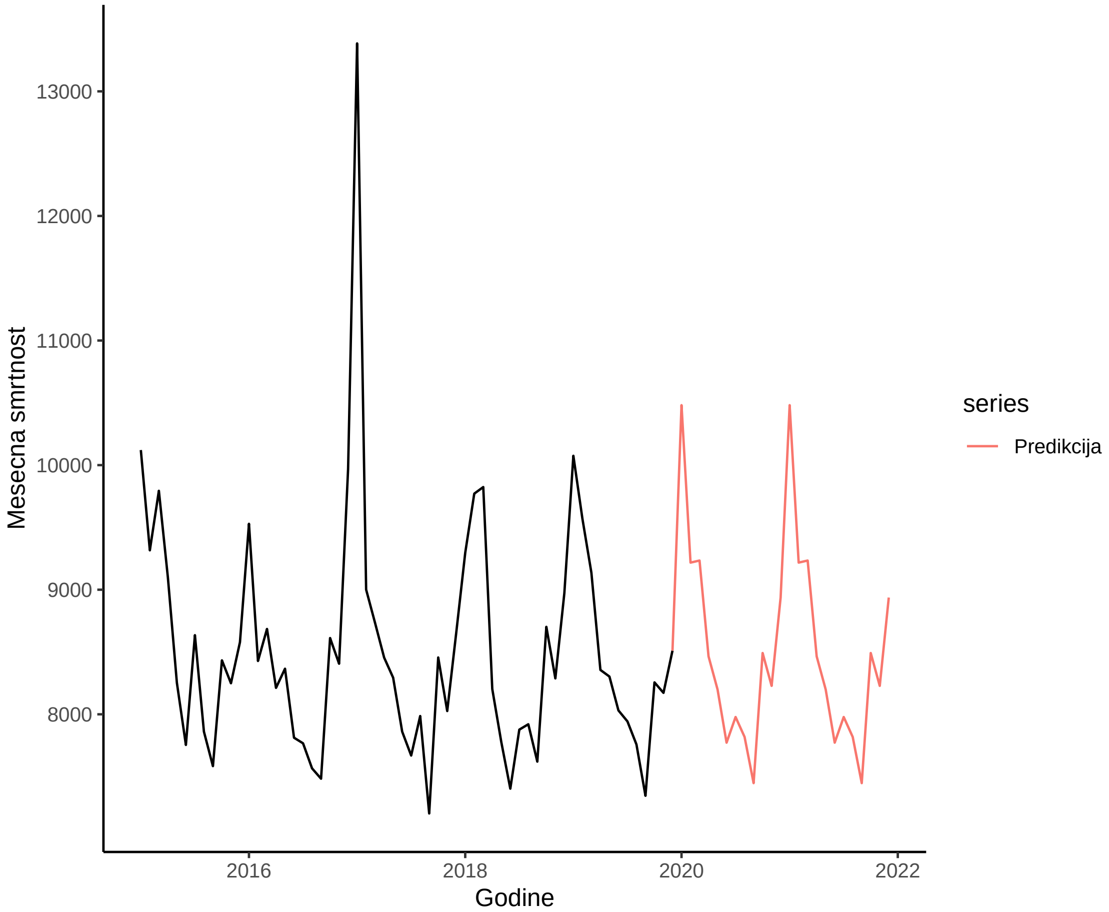
$$\begin{aligned} a_t &= \alpha(x_t - s_{t-p}) + (1-\alpha)(a_{t-1} + b_{t-1}), \\ b_t &= \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}, \\ s_t &= \gamma(x_t - a_t) + (1-\gamma)s_{t-p}, \end{aligned}$$

  gde je a_t nivo, b_t nagib, s_t sezonski efekat, a p vrednost perioda koji posmatramo. Parametri α, β i γ su parametri koje ocenjujemo i čija ocena određuje na koji način svaki od dva sabirka koji figurišu sa desne strane jednakosti ima uticaj na ocenjeni nivo, nagib i sezonski efekat.

  Konačna ocena vremenske serije u trenutku t je tada:

$$\hat{x_t} = \hat{a_{t-1}} + \hat{b_{t-1}} + \hat{s_{t-p}}, \tag{4.10}$$

  za



Slika 4.3: Prikaz vremenske serije smrtnosti u Srbiji do 2019. godine i predikcija za 2020. i 2021. pomoću regresije

$$\begin{aligned}
\hat{a}_t &= \hat{\alpha}(x_t - s_{t-p}) + (1-\hat{\alpha})(a_{t-1} + b_{t-1}),\\
\hat{b}_t &= \hat{\beta}(a_t - a_{t-1}) + (1-\hat{\beta})b_{t-1},\\
\hat{s}_t &= \hat{\gamma}(x_t - a_t) + (1-\hat{\gamma})s_{t-p}.
\end{aligned}$$

Primenom ugrađene funkcije *HoltWinters()* u R-u na vremensku seriju, dobijaju se ocenjeni parametri $\hat{\alpha}$, $\hat{\beta}$ i $\hat{\gamma}$, odnosno 0.4173584, 0, 0. Ove ocene se kao i u slučaju regresije dobijaju metodom minimizacije sume kvadrata reziduala, odnosno

$$\sum_{t=1}^{n}(x_t - \hat{x_t})^2.$$

Na osnovu toga, jednačine aditivnog Holt-Winters modela izgledaju ovako:

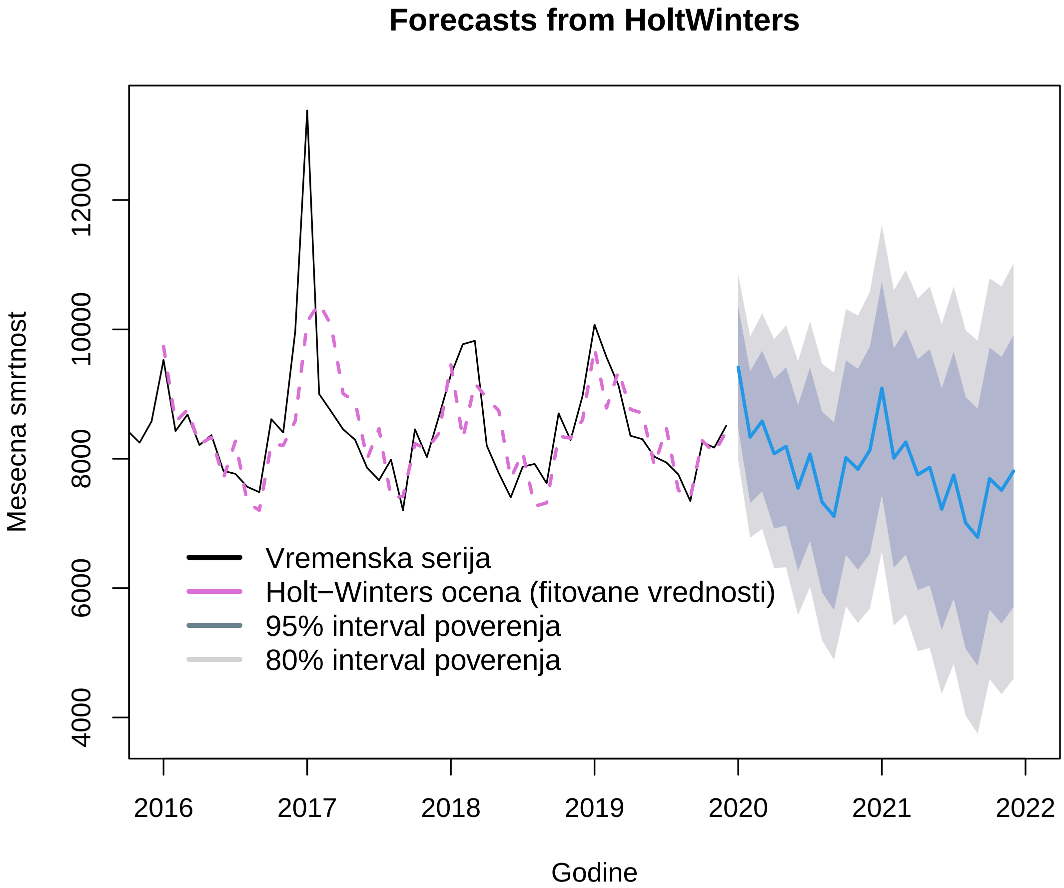
$$\hat{a_t} = 0.4173584(x_t - s_{t-p}) + 0.5826416(a_{t-1} + b_{t-1}), \tag{4.11}$$

$$\hat{b_t} = b_{t-1}, \tag{4.12}$$

$$\hat{s_t} = s_{t-p}. \tag{4.13}$$

Vrednost parametra $\hat{\beta}$ ukazuje na to da je prethodna vrednost nagiba dobra ocena za vrednost u sadašnjem trenutku, što znači da je trend linearan, jer je nagib (izvod) konstantan, odnosno trend se neće menjati nepredvidivo, dok za $\hat{\gamma}$ ukazuje da se vrednost sezonske komponente ne menja drastično iz godine u godinu.

Za predviđanje budućih vrednosti se koristi funkcija *forecast()*. Predviđanje se vrši primenom ocenjenih parametara, na osnovu jednačine 4.10. Intervali poverenja se formiraju pod uslovom da su greške međusobno nezavisne, homoskedastične i imaju normalnu raspodelu. Za proveru nezavisnosti se obično koristi autokorelaciona funkcija. Ukoliko ne postoji zavisnost između reziduala, funkcija je jednaka nuli u svim tačkama osim u nuli, gde se meri korelcija sa korakom nula, te je ta vrednost uvek jednaka 1. Grafik autokorelacione funkcije nam daje odgovor na pitanje da li postoje korelacije između reziduala. Sa grafika na slici 4.5 vidimo da ne postoje značajne korelacije. To može da se potvrdi i pomoću Ljung-Boksovog testa. Nulta hipoteza glasi da reziduali nisu međusobno korelisani, te p vrednost ovog testa koja iznosi 0.2296 potvrđuje ono što smo videli i na grafiku. Reziduali našeg modela ispunjavaju i uslov konstantne disperzije, ali ne prolaze statističke testove koji se odnose na normalnu raspodelu. Mogući razlog za to je postojanje autlajera koji su prisutni početkom 2017. godine. Uklanjanjem vrednosti iz januara 2017., reziduali prolaze Shapiro-Wilk test (p vrednost je 0.23). Na slici 4.4 predstavljena je predikcija za naredne dve godine, kao i intervali poverenja ove ocene. Interval predviđanja (npr. u nekom budućem trenutku $t + h$) će biti jednak $\hat{x_{t+h}} \pm z \cdot \hat{\sigma_h}$, gde je $\hat{x_{t+h}}$ ocena u trenutku $t+h$, z se odnosi na vrednost funkcije normalne raspodele u tački koja odgovara kvantilu određenom nivoom poverenja (za 95% nivo poverenja odgovarajući kvantil je 0.975), a $\hat{\sigma_h}$ ocenjena vrednost standardne devijacije sa korakom h.



Slika 4.4: Prikaz vremenske serije smrtnosti u Srbiji do 2019. godine i predikcija za 2020. i 2021. pomoću Holt-Winters modela

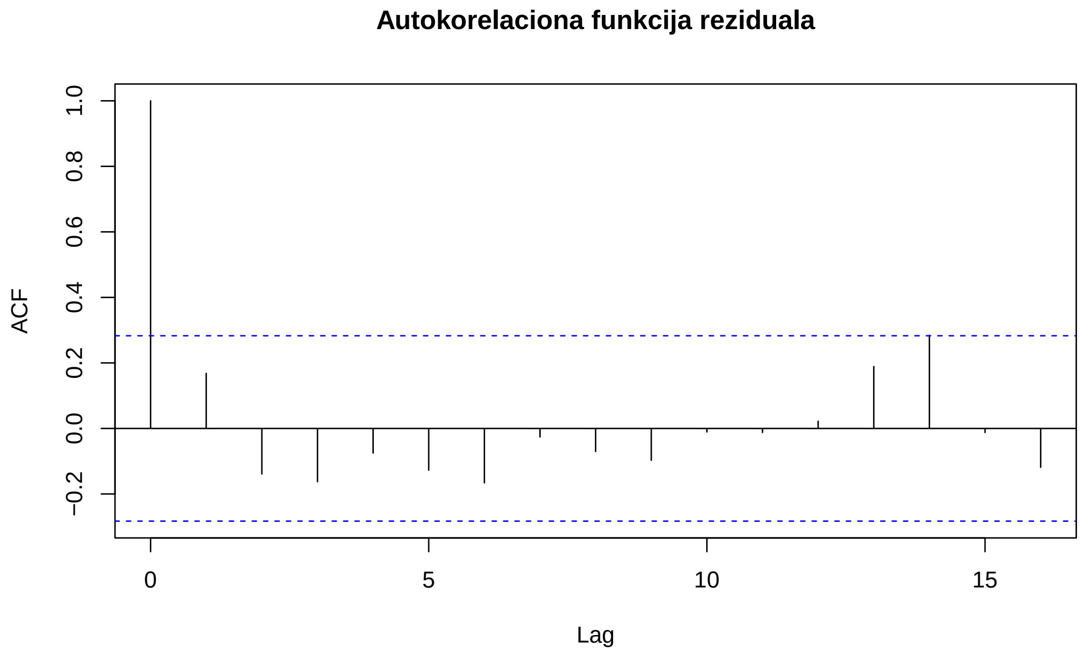
- **SARIMA modeli**

  Sada ćemo probati da modelujemo pomoću sezonskog $ARIMA$ modela. $ARMA$ model u oznaci $ARMA(p, q)$ je kombinacija autoregresionog modela reda p - $AR(p)$ i Moving Average modela reda q - $MA(q)$. Može se prikazati u polinomijalnom obliku:

$$\phi(B)x_t = \theta(B)\omega_t,$$

  gde je $\phi(B)$ autoregresivni operator koji je definisan kao

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p,$$



Slika 4.5: Autokorelaciona funkcija reziduala Holt-Winters modela

a $\theta(B)$ MA operator koji je definisan na sledeći način

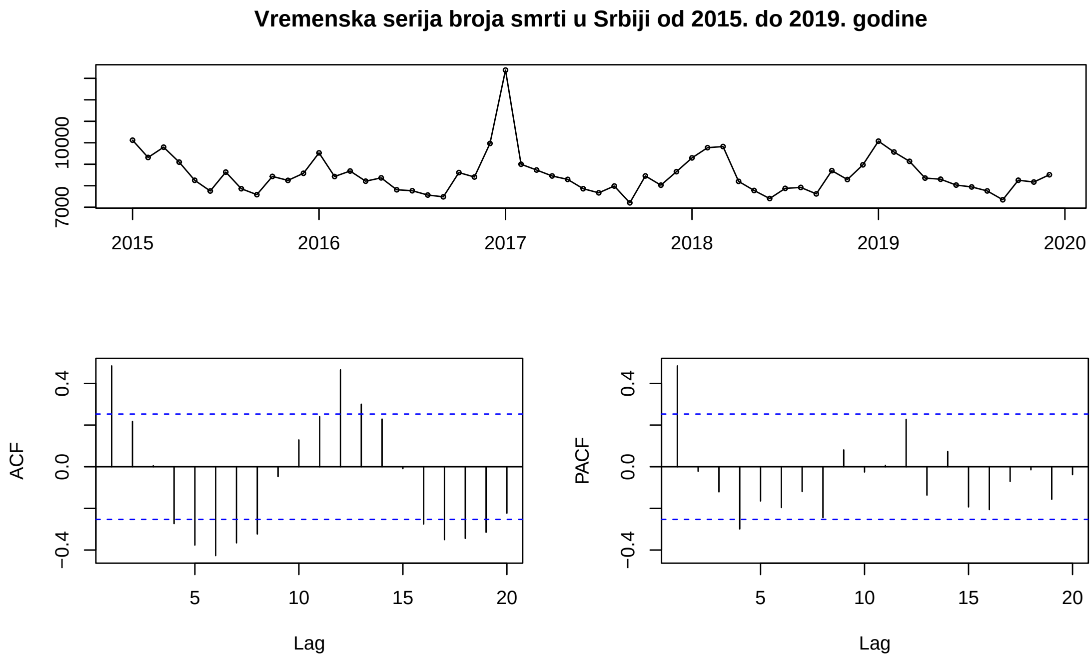
$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 - ... + \theta_q B^q,$$

ako je sa B označen tzv. *backshift* operator $Bx_t = x_{t-1}$.[12]

Kao što može da se primeti, autoregresivni model se oslanja na prethodne vrednosti u vremenskoj seriji za predikciju budućih, dok MA model koristi ω_t, za koje se pretpostavlja da imaju normalnu raspodelu, tj. $\omega_t \sim N(0, \sigma^2)$. Nesezonski ARIMA model u oznaci $ARIMA(p, d, q)$ je $ARMA$ model koji pored te dve komponente uključuje i diferenciranje reda d. On se koristi kod nestacionarnih vremenskih serija, obično ukoliko je jasno uočljiv trend.

U ovom radu je ipak od većeg intresa sezonski $ARIMA$ model. On se označava sa $ARIMA(p, d, q)(P, D, Q)_s$, gde je s - sezonski period. U ovom slučaju s će uvek biti 12, jer posmatramo podatke na mesečnom nivou u toku jedne godine. Vremenske serije sa sezonskom komponentom su nestacionarne, jer njihova funkcija srednje vrednosti nije konstantna tokom godine.

Polinomijalni zapis ovog modela je:



Slika 4.6: Prikaz vremenske serije smrtnosti i grafik ACF i PACF

$$\Phi_P(B^s)\phi_p(B)(1-B^s)^D(1-B)^d x_t = \Theta_Q(B^s)\theta_q(B)\omega_t,$$

pri čemu je D red sezonskog diferenciranja, a $\Phi_P(B^s)$ i $\Theta_Q(B^s)$ sezonski AR i MA polinomi oblika

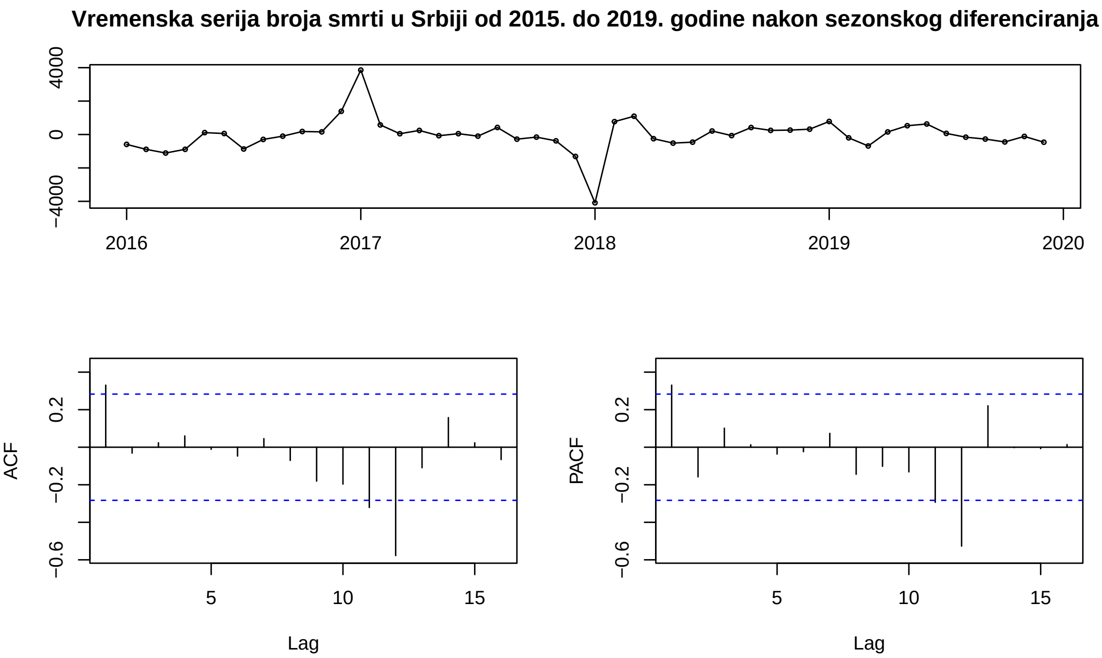
$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - ... - \Phi_P B^{Ps},$$

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + ... + \Theta_Q B^{Qs}.$$

Na slici 4.6 posmatramo izgled serije, kao i grafik autokorelacione funkcije (ACF) i parcijalne autokorelacione funkcije (PACF).

Kako je utvrđeno da postoji sezonska komponenta, pre svega se vremenska serija sezonski diferencira. To je prikazano na slici 4.7

Postojanje značajnih negativnih korelacija sa zadrškom 12 i na ACF i na PACF grafiku može ukazati na to da je sezonski MA polinom reda 1, te stavljamo da je $Q = 1$. Uočavaju se i značajne korelacije sa zadrškom 1, ali s obzirom da vrednosti nisu mnogo izvan intervala poverenja, možemo da pokušamo sa različiti varijantama. U tabeli su date AIC vrednosti za različite modele.



Slika 4.7: Prikaz vremenske serije smrtnosti i grafik ACF i PACF nakon sezonskog diferenciranja

Tabela 4.2: AIC vrednosti za različite modele

p, q	AIC
p=0, q=1	769.1463
p=1, q=1	770.214
p=1, q=0	771.1077

Najmanju AIC vrednost ima model $ARIMA(0,0,1)(0,1,1)_{12}$.

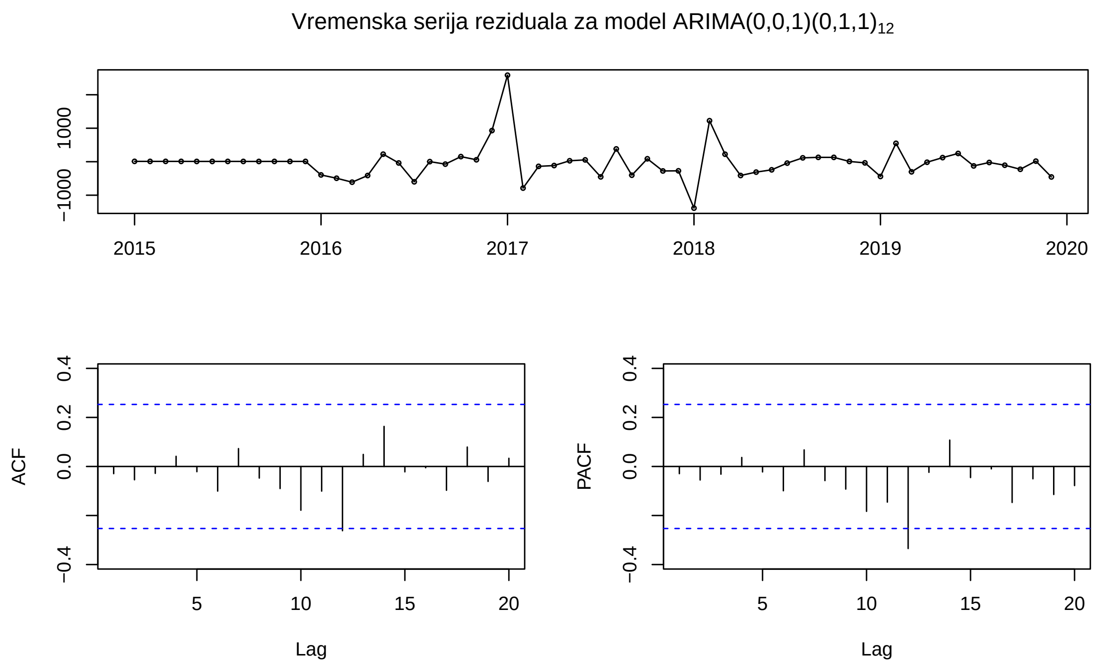
Na slici 4.8 prikazana je vremenska serija reziduala i grafika ACF i PACF. Vidimo da postoji jedna korelacija koja je značajna.

Ukoliko se doda i sezonski AR parametar, sa grafika 4.9 se vidi da više nema značajnih korelacija, pa biramo model $ARIMA(0,0,1)(1,1,1)_{12}$.

Prognoza za buduće vrednosti je prikazana na slici 4.10.

Jednačina ovog modela je:

$$(1 - \Phi_1 B^{12})(1 - B^{12})x_t = (1 + \Theta_1 B^{12})(1 + \theta_1 B)\omega_t,$$



Slika 4.8: Prikaz vremenske serije reziduala i grafik ACF i PACF modela $ARIMA(0,0,1)(0,1,1)_{12}$

odnosno:

$$x_t = x_{t-12} + \Phi_1 x_{t-12} - \Phi_1 x_{t-24} + \omega_t + \theta_1 \omega_{t-1} + \Theta_1 \omega_{t-12} + \theta_1 \Theta_1 \omega_{t-13}.$$

Kada uključimo vrednosti parametara dobijenih iz R-a, dobija se:

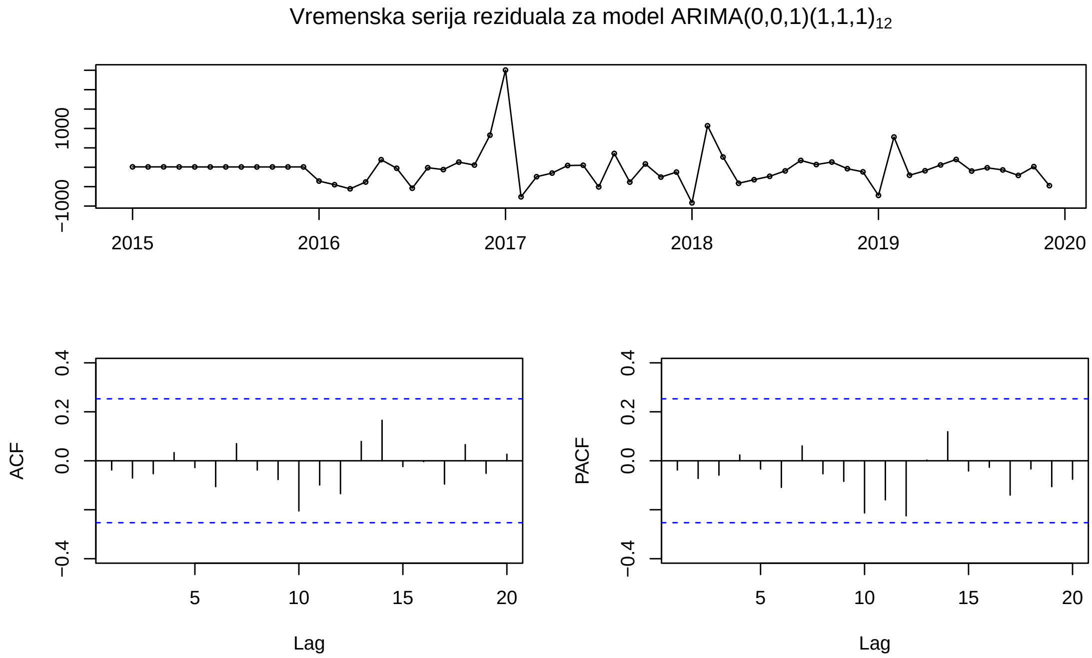
$$x_t = x_{t-12} - 0.2012 x_{t-12} + 0.2012 x_{t-24} + \omega_t + 0.3231 \omega_{t-1} - \omega_{t-12} - 0.3231 \omega_{t-13}.$$

Dakle, ako imamo tablicu smrtnosti koja sadrži podatke o broju preminulih po mesecima u toku jedne godine, tada bismo mogli da odredimo vrednosti u budućnosti za pojedinačne mesece u godini. Kako je

$${}_u q_x = \frac{\ell_x - \ell_{x+u}}{\ell_x},$$

tada možemo izračunati jednokratnu neto premiju privremenog osiguranja za osobu starosti $x + \frac{h}{12}$, gde je $h = 1, 2, ..., 12$ i pokriće u periodu dužine $\frac{1}{12}$:

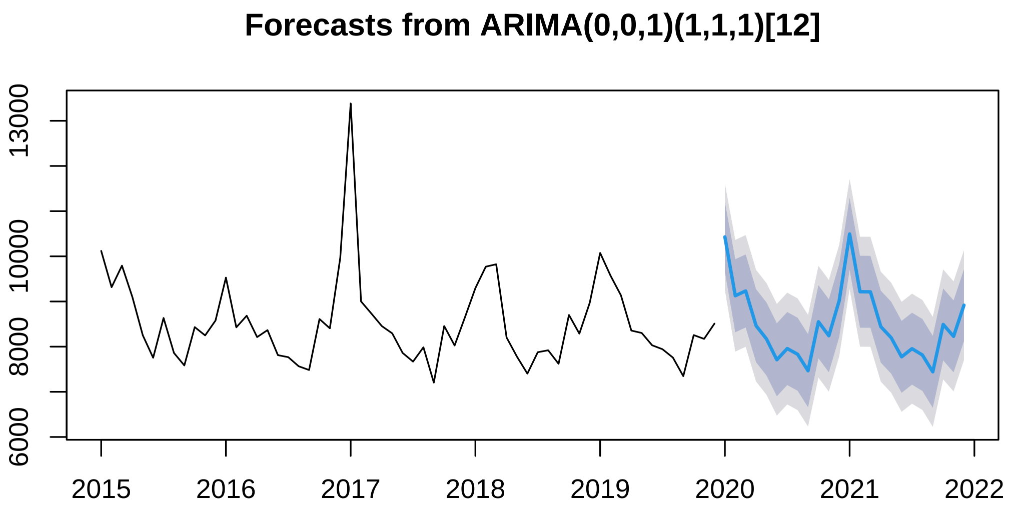
$$\bar{A}^1_{x+\frac{h}{12}:\overline{\frac{1}{12}}|} = E(Z) = E\left(I\left\{S < \frac{h+1}{12}\middle| S \geq \frac{h}{12}\right\}\right) = P\left\{S < \frac{h+1}{12}\middle| S \geq \frac{h}{12}\right\} = {}_{\frac{1}{12}}q_{x+\frac{h}{12}}.$$



Slika 4.9: Prikaz vremenske serije reziduala i grafik ACF i PACF modela $ARIMA(0,0,1)(1,1,1)_{12}$

Ukoliko bismo hteli da izračunamo premiju privremenog osiguranja za osobu starosti 54 godine i 6 meseci gde pokriće traje mesec dana, na osnovu projektovanog broja umrlih za naredni period i podataka o broju preživelih na kraju prethodnog, dobija se rezultat 1.15. I u ovom slučaju je iznos osigurane sume 1000 novčanih jedinica, te je dobijeni iznos dosta veći nego preko modela koji koristi periodične slučajne veličine.

Ukoliko bismo poredili modele prikazane u ovom radu, pre svega bi trebalo krenuti od poređenja modela zasnovanom na periodičnim slučajnim veličinama i modela zasnovanim na vremenskim serijama. Model zasnovan na periodičnim slučajnim veličinama daje ocenu same funkcije gustine, dok nam modeli zasnovani na analizi vremenskih serija daju ocenu broja smrti u toku godine. Iako je prva ocena analitička i pogodnija za dalji rad, modeli vremenskih serija su jednostavniji a rezultati interpretabilniji. Dodatnu komplikaciju kod periodičnih slučajnih veličina predstavlja parametar M koji je potrebno oceniti i proslediti funkciji koja računa odgovarajuće koeficijente. Ukoliko bismo uporedili AIC vrednosti NNTS i $ARIMA$ modela, videli bismo da su dosta manje vrednosti prisutne kod $ARIMA$ modela.



Slika 4.10: Prikaz vremenske serije i predviđanje budućih vrednosti pomoću $ARIMA(0,0,1)(1,1,1)_{12}$ modela

Što se tiče modela koji se najčešće koriste prilikom rada sa vremenskim serijama, u ovom poglavlju smo detaljno prikazali više mogućnosti. Na osnovu predstavljenih analiza kako rezultata, tako i grafika autokorelacione funkcije, kao najbolji model $ARIMA$ model bismo izabrali $ARIMA(0,0,1)(1,1,1)_{12}$ model. Otklonili smo sve značajne korelacije u rezidualima i dobili model koji najviše odgovara našim podacima. S obzirom da reziduali linearne regresije nisu ispunjavali porebne uslove, kao najboljeg kandidata bismo tada izabrali ili $ARIMA(0,0,1)(1,1,1)_{12}$ model ili Holt-Winters model.

Zaključak

Motivacija ovog rada je bila da se pronađe odgovarajući model koji će na jasan i egzaktan način da predstavi pravilnosti koje se javljaju u smrtnosti populacije u toku godine. To bi za cilj imalo unapređenje samog obračuna premije koji bi tada obezbedio preciznije rezultate nego u opštem slučaju. S druge strane, s obzirom na karakter polisa životnog osiguranja, i činjenicu da su to uglavnom ugovori sklopljeni na duži vremenski period, ovo bi svoju primenu našlo eventualno u slučajevima gde postoji kraći period pokrića.

U prvoj glavi rada dat je teorijski uvod sa elementima finansijske mtematike koji su potrebni u svrhu bavljenja ovom tematikom. U drugoj glavi se nalazi pregled pojmova koji se sreću prilikom rada u oblasti osiguranja i aktuarstva. Ove dve glave čine osnovu za ostatak rada. U trećoj glavi je prikazano jedno rešenje početnog problema - posmatranje periodične slučajne veličine. Data je motivacija za njeno korišćenje i opisani su uslovi koje mora da ispunjava njena funkcija gustine raspodele. Četvrto poglavlje je podeljeno na dve vrste modelovanja - zasnovano na periodičnoj slučajnoj veličini i zasnovano na analizi vremenskih serija. Kao primer smo posmatrali podatke o broju umrlih u Republici Srbiji od 2015. godine. Razmatrano je nekoliko modela koji se koriste u svrhu proučavanja vremenskih serija, provereno da li ispunjavaju sve neophodne uslove i upoređen je njihov kvalitet. Na kraju je izabran model $ARIMA(0,0,1)(1,1,1)_{12}$ koji je najviše pogodovao podacima.

Ideja za dalje istraživanje bi mogla da se početno bazira na radu [13]. Smatram da bi za nastavak izučavanja ove teme intresantno bilo da se pronađe novi pristup, kao što je opisano u tom radu. Prustup se bazira na tome da se više ne posmatra ukupna smrtnost na nivou populacije - već individualno, koristeći funkciju preživljavanja.

Bibliografija

[1] James R. Fegan. A short history of mortality, 2000.

[2] Nataša Petrović Tomić. *Osnovi prava osiguranja*. Pravni fakultet Univerziteta u Beogradu, 2021.

[3] Hans U. Gerber. *Life Insurance Mathematics*. Springer Berlin, Heidelberg, 1997.

[4] Slobodanka Janković and Bojana Milošević. *Elementi finansijske matematike*. Univerzitet u Beogradu, Matematički fakultet, 2017.

[5] Steve Selvin. *Life tables*. Practical Guides to Biostatistics and Epidemiology. Cambridge University Press, 2008.

[6] J.J. Fernández-Durán and M.M. Gregorio-Domínguez. Seasonal mortality for fractional ages in short term life insurance. *Scandinavian Actuarial Journal*, 2015(3):266–277, 2015.

[7] J. J. Fernández-Durán. Circular distributions based on nonnegative trigonometric sums. *Biometrics*, 60(2):499–503, 2004.

[8] Borislav D. Bojanov. *Approximation Theory: A Volume Dedicated to Blagovest Sendov*. 2002.

[9] Leopold Fejér. Über trigonometrische polynome. *Journal für die reine und angewandte Mathematik*, 146:53–82, 1916.

[10] J.J. Fernández-Durán and M.M. Gregorio-Domínguez. Circnntsr: An r package for the statistical analysis of circular, multivariate circular, and spherical data using nonnegative trigonometric sums. *Journal of Statistical Software*, 70(6):1–19, 2016.

[11] M.A. Gallo, B.E. Wheeler, and I.M. Silver. *Fundamentals of Statistics for Aviation Research*. Aviation Fundamentals. Taylor & Francis, 2023.

[12] R.H. Shumway and D.S. Stoffer. *Time Series Analysis and Its Applications: With R Examples*. Springer Texts in Statistics. Springer International Publishing, 2017.

[13] Stephen J. Richards, Stefan J. Ramonat, Gregory T. Vesper, and Torsten Kleinow. Modelling seasonal mortality with individual data. *Scandinavian Actuarial Journal*, 2020(10):864–878, November 2020.

[14] Republički zavod za statistiku. Podaci o broju umrlih prema polu i mesecu smrti. `https://data.stat.gov.rs/Home/Result/18030306`. Accessed: 2023-07-30.

Biografija autora

Dunja Radenković je rođena u Beogradu 29. jula 1999. godine, gde je i završila osnovno i srednjoškolsko obrazovanje. Matematički fakultet je upisala 2018. godine, na smeru Statistika, finansijska i aktuarska matematika. Osnovne studije je završila 2022. godine, kada je upisala i master studije na istom smeru.
U februaru 2023. godine se zaposlila kao deo aktuarskog tima u kompaniji PwC, gde usavršava svoje znanje u toj oblasti radeći na brojim međunarodnim projektima. Od aprila 2024. godine je pridruženi član Udruženja aktuara Srbije.