

UNIVERZITET U BEOGRADU

Matematički fakultet



MASTER RAD

MULTIPLIKATIVNI MODELI VREMENSKIH SERIJA

Mentor:

dr Jelena Jocković

Autor:

Sonja Jovanović 1057/2020

Beograd, septembar 2021.

SADRŽAJ

1. UVOD.....	4
2. VREMENSKE SERIJE.....	5
2.1. UVODNA ANALIZA VREMENSKIH SERIJA	5
2.2. PODELA VREMENSKIH SERIJA.....	5
2.3. KOMPONENTE VREMENSKIH SERIJA	6
2.4. CILJEVI ANALIZE VREMENSKIH SERIJA	7
2.5. OSNOVNE OZNAKE I POJMOVI U ANALIZI VREMENSKIH SERIJA	7
2.6. AUTOKOVARIJACIONA I AUTOKORELACIONA FUNKCIJA	10
2.7. TESTOVI SLUČAJNOSTI.....	12
3. METOD DEKOMPOZICIJE VREMENSKIH SERIJA	13
3.1. OSNOVNI KORACI DEKOMPOZICIJE.....	14
4. MODELI VREMENSKIH SERIJA.....	16
4.1. OPIS MODELA	16
4.1.1. ADITIVNI MODEL	16
4.1.2. MULTIPLIKATIVNI MODEL	16
4.1.3. MEŠOVITI (KOMBINOVANI) MODEL.....	17
4.1.4. KAKO ODABRATI ODGOVARAJUĆI MODEL.....	18
4.2. MODELI AR, MA, ARMA, ARIMA.....	18
4.2.1. AUTOREGRESIONI MODELI (AR).....	19
4.2.2. MODELI POKRETNIH PROSEKA (MA)	20
4.2.3. AUTOREGRESIONI MODELI POKRETNIH PROSEKA (ARMA)	20
4.2.4. AUTOREGRESIONI MODELI POKRETNIH PROSEKA ZA INTEGRISANE VREMENSKE SERIJE (ARIMA)	20
4.3. SEZONSKE VREMENSKE SERIJE	21
4.3.1. SEZONSKI ARIMA MODELI	22
4.3.2. PRIMERI SEZONSKIH ARIMA MODELA.....	23
4.3.2.1. Sezonski autoregresioni model prvog reda.....	23
4.3.2.2. Sezonski integrisani model prvog reda.....	23
4.3.2.3. Sezonski model pokretnih proseka prvog reda.....	23
4.3.2.4. „Vazduhoplovni model“	24
5. PRIMENA METODA DEKOMPOZICIJE I IZBOR MODELA.....	25
5.1. PRIMER 1	26
5.2. PRIMER 2	39
5.3. PRIMER 3	46
5.4. PRIMER 4	52
6. SIMULACIJA MODELA	57
6.1. SIMULACIJA ADITIVNIH MODELA	57

6.2. SIMULACIJA MULTIPLIKATIVNIH MODELA	66
6.3. SIMULACIJA MEŠOVITIH MODELA.....	74
7. ZAKLJUČAK.....	84
8. LITERATURA.....	86

1. UVOD

Analiza vremenskih serija predstavlja jednu od statističkih disciplina koja beleži najdinamičniji razvoj poslednjih decenija. Ona je predmet istraživanja u različitim naučnim oblastima, a pogotovo u ekonomiji. Njeni metodi omogućavaju donošenje relevantnih ekonomskih zaključaka.

Jedan od ciljeva analize vremenskih serija je izbor modela vremenske serije koji opisuje njeno kretanje na zadovoljavajući način. Model se zatim koristi da se na osnovu prošlih opservacija prognoziraju buduće. Prognoziranje budućeg kretanja vremenske serije jeste jedan od osnovnih zadataka analize vremenskih serija, stoga je jako bitno odabrati odgovarajući model vremenske serije.

Može se pretpostaviti da je vremenska serija sačinjena od četiri komponente: trenda, ciklične komponente, sezonske komponente i slučajne komponente.

Prema aditivnom modelu, varijacije pojave predstavljaju zbir komponenti, dok prema multiplikativnom modelu one predstavljaju njihov proizvod. Mešoviti model kombinuje sabiranje i množenje komponenti.

U ovom radu će biti prikazan i metod dekompozicije vremenskih serija, koji sa jedne strane predstavlja razlaganje vremenske serije na njene komponente, dok sa druge strane predstavlja mogućnost da se pojedine komponente po potrebi eliminišu iz vremenske serije. Jedan od glavnih ciljeva dekompozicije je ocena sezonskih efekata koji se mogu koristiti za kreiranje i prikazivanje sezonski prilagođenih vrednosti.

Mnoge poslovne i ekonomske vremenske serije održavaju i objavljuju vladine agencije koje koriste dekompoziciju vremenskih serija da bi kreirali desezonirane vremenske serije. Za metod dekompozicije je takođe jako bitno odabrati odgovarajući model.

Prilikom analize ekonomskih pojava najčešće se koristi multiplikativni model pa je zato on veoma važan za analizu.

Multiplikativni modeli su značajni i u raznim drugim oblastima, te se na mnoge načine upotrebljavaju u praksi. Neki od primera naučnih radova u kojima se koristi multiplikativni model navedeni su u literaturi ([21], [22], [23]). U njima se može videti primena multiplikativnih modela na predviđanje dnevnog toka novca, relativne stope preživljavanja, pa čak i u biologiji, u uzgoju biljaka. Stoga je ovo značajno polje praktičnog istraživanja.

Međutim, primeri multiplikativnih modela su mnogo manje zastupljeni u literaturi od aditivnih, pa ih je i teže prepoznati u praksi.

Cilj ovog rada je da se na raznovrsnim primerima realizovanih vremenskih serija prikaže kako odlučiti da li je u pitanju aditivni, multiplikativni ili neki od mešoviti modela. Time će i multiplikativni modeli biti lakše prepoznati.

U radu će biti prikazani primeri modela nad raznim skupovima podataka, ali takođe i simulacije modela nad proizvoljnim podacima. Za praktični deo rada biće korišćen programski jezik R. Usput, biće detaljnije objašnjene i razne mogućnosti ovog programskog paketa.

U ovom poglavlju upotrebljena je i literatura [2], [3], [4].

2. VREMENSKE SERIJE

2.1. UVODNA ANALIZA VREMENSKIH SERIJA

Vremenska serija predstavlja uređeni niz podataka, pri čemu se uređenje ostvaruje u odnosu na vreme i najčešće u jednakim vremenskim intervalima.

Izbor jedinice vremena se po pravilu vrši prema samoj pojavi i svrsi ispitivanja, pa tako to može biti jedna sekunda, sat, a nekada čak i cela decenija.

Sa primerima vremenskih serija se susrećemo u svakodnevnom životu: praćenje dnevne temperature, fluktuacije deviznog kursa, godišnje kretanje prinosa pojedinih poljoprivrednih kultura i njihove otkupne i prodajne cene, posmatranje mesečnog kretanja cena, godišnjeg nivoa proizvodnje, itd.

Ovo su primeri **jednodimenzionih** vremenskih serija, dok postoje i one vremenske serije koje su **višedimenzione**, npr. količina padavina na jednom mestu određenom njegovom geografskom dužinom i širinom, brzina vetra u prostoru i vremenu, itd.

Analiza vremenskih serija je statistička disciplina, međutim ona se razlikuje od klasične statističke analize. U klasičnoj teoriji verovatnoće i matematičke statistike najčešće se izučavaju nizovi nezavisnih slučajnih veličina. Vremenska serija je pak, matematički gledano, niz slučajnih veličina koje su **međusobno zavisne**, jer pri analizi vremenskih serija mora da se uzme u obzir njihov vremenski poredak.

Svaka vremenska serija dužine n može da se shvati kao realizacija niza n međusobno zavisnih slučajnih veličina. Često su bliski članovi vremenske serije međusobno sličniji nego dalji članovi te serije. Ovu međusobnu zavisnost (najčešće korelisanost) opservacija koristimo u analizi vremenskih serija u cilju formiranja **modela** vremenske serije.

Vremenska serija se može definisati i na sledeće načine:

Definicija 2.1.1: Vremenska serija predstavlja jednu realizaciju slučajnog procesa.

Definicija 2.1.2: Vremenska serija se može shvatiti kao realni slučajni proces $\{X_t, t \in T\}$ sa neprekidnim ili diskretnim slučajnim parametrom, pri čemu slučajne veličine X_t i X_s , $t \neq s$ nisu nekorelisane.

2.2. PODELA VREMENSKIH SERIJA

Vremenske serije možemo klasifikovati prema različitim kriterijumima. Podela vremenskih serija u zavisnosti od toga kako se registruju podaci (vrednosti) vremenske serije je na **prekidne** i **neprekidne**.

Prekidne vremenske serije su one kod kojih se opservacije beleže u istim vremenskim intervalima, npr. dnevno, mesečno, godišnje.

Neprekidne vremenske serije su one kod kojih se opservacije mogu registrovati u bilo kom vremenskom trenutku.

Podela vremenskih serija u zavisnosti od toga da li se statistička svojstva vremenske serije menjaju u toku vremena ili ne je na **stacionarne** i **nestacionarne**.

Stacionarne vremenske serije su one kod kojih je kretanje vremenske serije predvidivo tokom vremena. One ispoljavaju isti ili sličan obrazac ponašanja tokom vremena.

Nestacionarne vremenske serije su one kod kojih je kretanje vremenske serije nepredvidivo tokom vremena.

Stacionarnost je bitan pojam u analizi vremenskih serija i biće detaljnije opisan u narednim poglavljima.

2.3. KOMPONENTE VREMENSKIH SERIJA

Komponente vremenskih serija su **trend**, **sezonska**, **ciklična** i **slučajna komponenta**. Vremenska serija ne mora da sadrži sve četiri navedene komponente, ali mora da sadrži slučajnu komponentu. U nastavku teksta je objašnjena svaka od komponenti.

Trend predstavlja dugoročnu komponentu u kretanju vremenske serije.

Može biti rastući ili opadajući, u zavisnosti od toga da li vrednosti vremenske serije tokom vremena sistematski rastu ili opadaju.

Takođe, može biti deterministički ili stohastički, u zavisnosti od toga da li se promene vremenske serije tokom vremena mogu predvideti ili ne.

Identifikacija trenda predstavlja ozbiljan statistički problem. Kako bi se on dobro izabrao i analizirao, pretpostavlja se da se može dobro aproksimirati prostom matematičkom funkcijom koja može biti određena aritmetička sredina, polinom određenog stepena, ili neka složenija funkcija.

Sezonska komponenta predstavlja pravilnost u kretanju vremenske serije koja se pojavljuje u okviru jedne kalendarske godine. Sezonalnost je uvek fiksnog i poznatog perioda.

Vremenske serije koje imaju izraženu sezonsku komponentu obično se nazivaju *sezonske vremenske serije*.

Primer vremenskih serija koje imaju izražene sezonske varijacije su uvoz, izvoz, indeks industrijske proizvodnje, javni prihodi itd.

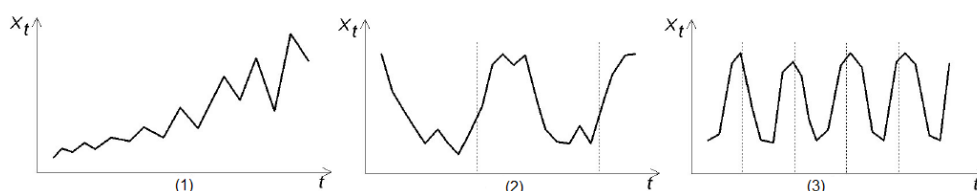
Prisustvo sezonske komponente utiče na to da postoji veći stepen korelacije između istih meseci različitih godina nego između različitih meseci iste godine.

Ciklična komponenta predstavlja pravilnost u kretanju vremenske serije koja se pojavljuje sa periodom različitim od jedne godine. Na primer, trajanje neke mode tokom izvesnog broja godina i njene zamene novom modom tokom sledećih godina, itd. U praksi, za komponentu trenda se pretpostavlja da uključuje i cikličnu komponentu.

Identifikacija ciklusa i periodičnosti je jedan od važnih zadataka analize jer omogućava dobru prognozu ponašanja budućih vrednosti serije.

Slučajna komponenta predstavlja promene u kretanju vremenske serije koje su slučajnog karaktera. Ona je stacionarni proces, koji će biti detaljnije objašnjen u daljem tekstu.

Na slici 2.3.1. prikazani su primeri vremenskih serija kod kojih je izražen trend, ciklična komponenta i sezonska komponenta, respektivno. Sa X_t je označena opservacija vremenske serije u trenutku t . Oznake i pojmovi u analizi vremenskih serija detaljno će biti opisani u poglavlju 2.5.



Slika 2.3.1. Vremenska serija sa izraženim trendom (1), cikličnim varijacijama (2) i sezonskim varijacijama (3) (slika je preuzeta iz literature [3])

Da bi se uklonile determinističke komponente iz vremenske serije (trend, sezonska i ciklična komponenta), vremenska serija se može razložiti na odvojene stacionarne i determinističke komponente. Taj metod se zove *metod dekompozicije* i detaljno će biti opisan u poglavlju 3.

2.4. CILJEVI ANALIZE VREMENSKIH SERIJA

Osnovni ciljevi u analizi vremenskih serija su:

Opisivanje vremenske serije (deskripcija) obezbeđuje informaciju o osnovnim svojstvima vremenske serije. Kako bi se izučile osnovne karakteristike vremenske serije koristi se grafički prikaz vremenske serije i njene sumarne statistike. To je vrlo korisno jer se često tokom analize pokaže da je npr. grafičkim prikazom došla do izražaja suštinska karakteristika posmatrane pojave, pa nema potrebe koristiti neke od složenijih statističkih metoda analize vremenske serije. Opisivanjem vremenske serije može se dobiti informacija da li je vremenska serija stacionarna ili ne, postoje li specifičnosti u kretanju vremenske serije, da li je vremenska serija normalno raspodeljena, da li ju je potrebno transformisati, itd.

Objašnjenje vremenske serije predstavlja izbor odgovarajućeg matematičkog modela koji opisuje kretanje vremenske serije na zadovoljavajući način. Ako postoji više vremenskih serija moguće je koristiti varijacije jedne vremenske serije u cilju objašnjenja varijacija u drugoj vremenskoj seriji. Generalizacijom modela jednodimenzionih vremenskih serija dobijaju se modeli višedimenzionih vremenskih serija, npr. od autoregresionih modela dobijaju se takozvani vektorski autoregresioni modeli.

Predviđanje vremenske serije odnosno prognoziranje budućeg stanja vremenske serije. Na osnovu prošlih opservacija vremenske serije identifikuje se i oceni model vremenske serije, a nakon toga se taj model koristi za prognoziranje budućeg stanja vremenske serije. Koriste se različiti statistički testovi i kriterijumi kojima se određuje valjanost korišćenog modela u poređenju sa drugim konkurentskim modelima.

U praksi se uvek polazi od opisivanja vremenske serije kako bi se najpre stekao pojam o osnovnim svojstvima vremenske serije. Kao što je rečeno, podaci se najpre predstavljaju grafički i posmatraju se sumarne statistike, a zatim se radi niz narednih opisanih koraka, koji će biti i praktično prikazani u poglavlju 5.

2.5. OSNOVNE OZNAKE I POJMOVI U ANALIZI VREMENSKIH SERIJA

U poglavlju 2.3. prilikom prikazivanja vremenskih serija grafički, sa X_t je označena opservacija vremenske serije u trenutku (periodu) t . Indeks t može označavati godinu, mesec, kvartal i slično.

U prethodnom poglavlju spomenuto je i prognoziranje vremenske serije, pa se uvodi oznaka $\hat{X}_n(h)$, koja predstavlja prognoziranu vrednost vremenske serije za h perioda unapred na osnovu prvih n podataka. To je ocena nepoznatog nivoa vremenske serije u trenutku $n + h$, tj. X_{n+h} . Period h za koji se prognozira, naziva se **horizont predviđanja (prognoziranja)**.

Između ostvarene i prognozirane vrednosti postoje razlike, odnosno odstupanja. To odstupanje se označava na sledeći način: $\hat{e}_n(h) = X_{n+h} - \hat{X}_n(h)$ i zove **greška prognoze** za h perioda unapred.

Uvode se i sledeći pojmovi koji će biti bitni za kasniju analizu vremenskih serija:

Operator docnje (kašnjenja) prvog reda L : $LX_t = X_{t-1}$. Upotrebom operatora kašnjenja nivo vremenske serije pomera se za jedan period unazad.

Operator diference (razlike) prvog reda Δ :

$$\Delta X_t = X_t - X_{t-1} = X_t - LX_t = (1 - L)X_t$$

Operator diference reda k , Δ_k :

$$\Delta_k X_t = X_t - X_{t-k}$$

Često se naziva i **operator sezonske diference** i primenjuje se u analizi kvartalnih i mesečnih podataka sa izraženom sezonskom komponentom.

Bitni pojmovi u analizi vremenskih serija su pojam **stacionarnosti** i **ergodičnosti**.

Analiza vremenskih serija nije lak zadatak jer je prilikom modeliranja vremenske serije na raspolaganju samo jedna realizacija na osnovu koje je potrebno zaključiti o karakteristikama celog procesa. Da bi se to postiglo mora da se uvedu ograničenja na slučajni proces, a tiču se vremenske homogenosti i memorije procesa. Slučajni proces koji ispunjava oba ograničenja zove se *stacionarni i ergodični proces*. Uvodi se i operativna definicija stacionarnosti koja je bazirana na momentima procesa- definicija *slabe stacionarnosti* odnosno *stacionarnosti drugog reda*. Modeli vremenskih serija se uglavnom odnose na ovu definiciju stacionarnosti. Stacionarnost je bitna jer vremenska nepromenljivost (stacionarnost) svojstava serije omogućava predviđanje (modeliranje) dinamike serije u budućnosti. Sledi pojašnjenje pojma i definicije.

Stacionarnost je svojstvo vremenske serije čije se kretanje tokom vremena odvija po ustaljenom obrascu. Postoje dva koncepta stacionarnosti:

- (1) *Stroga stacionarnost*
- (2) *Slaba stacionarnost*

Definicija 2.5.1: (Stroga stacionarnost vremenske serije) Za vremensku seriju X_t kaže se da je strogo stacionarna ako za bilo koja dva prirodna broja n i k i bilo koju n - torku prirodnih brojeva (t_1, t_2, \dots, t_n) slučajni nizovi $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ i $(X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k})$ imaju istu raspodelu verovatnoće.

Dakle, vremenska serija je strogo stacionarna ako se njena svojstva ne menjaju transliranjem u vremenu.

Uvode se i sledeći pojmovi:

Srednja vrednost procesa μ_t je

$$\mu_t = E(X_t)$$

Disperzija procesa je

$$\sigma_t^2 = E(X_t - \mu_t)^2$$

Kovarijaciona funkcija između X_{t_1} i X_{t_2} u dva vremenska trenutka t_1 i t_2 je

$$\gamma(t_1, t_2) = E(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})$$

Korelaciona funkcija između X_{t_1} i X_{t_2} je

$$\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sqrt{\sigma_{t_1}^2} \sqrt{\sigma_{t_2}^2}}$$

Definicija 2.5.2: (Slaba stacionarnost vremenske serije) Vremenska serija X_t je slabo stacionarna ukoliko zadovoljava sledeće uslove:

- (1) $E(X_t) = \mu = \text{const}, t = 1, 2, \dots$
- (2) $\sigma^2 = E(X_t - \mu)^2 = \text{const}, t = 1, 2, \dots$
- (3) $\gamma(t, t - k) = E(X_t - \mu)(X_{t-k} - \mu) = \gamma(k) = \gamma_k, t = 1, 2, \dots, k = 1, 2, \dots$

Navedeni uslovi sugerišu da se srednja vrednost i disperzija slabo stacionarne vremenske serije ne menjaju tokom vremena. Kovarijaciona funkcija između svaka dva člana ovakve vremenske serije je funkcija vremenskog rastojanja (kašnjenja) između njih.

Ako je vremenska serija strogo stacionarna onda je ona i slabo stacionarna jedino ako poseduje konačnu disperziju, a obrnuto ne važi.

Uslovi stroge i slabe stacionarnosti su ekvivalentni ako je zajednička raspodela slučajnog procesa normalna i takav proces se zove **Gausov slučajni proces (niz)**.

Definicija 2.5.3: Najjednostavniji stacionarni slučajni proces naziva se **beli šum**, označava se sa $e_t, t = 1, 2, \dots$ i ima sledeća svojstva:

- (1) $E(e_t) = 0, t = 1, 2, \dots$
- (2) $\sigma^2 = E(e_t)^2 = \text{const}, t = 1, 2, \dots$
- (3) $\gamma(t, t - k) = E(e_t e_{t-k}) = 0, t = 1, 2, \dots, k = 1, 2, \dots$

Ako se navedenim uslovima doda i uslov da su članovi niza nezavisne slučajne promenljive, čija je zajednička raspodela normalna, onda se ovaj slučajni proces zove **Gausov beli šum**.

Kako je već gore pomenuto, da bi se na osnovu jedne realizacije procesa zaključilo o karakteristikama celog procesa potrebno je pored ograničenja vremenske homogenosti uvesti i ograničenje u pogledu memorije procesa.

Pre nego što se definiše *ergodičnost* slučajnih procesa, uvodi se pojam *konvergencije u verovatnoći*.

Definicija 2.5.4: Niz slučajnih veličina (X_n) **stohastički konvergira (konvergira u verovatnoći)** ka slučajnoj veličini X , ako važi:

$$(\forall \varepsilon > 0) P(|X_n - X| \geq \varepsilon) \rightarrow 0, n \rightarrow \infty.$$

Za konvergenciju u verovatnoći koristi se oznaka $X_n \xrightarrow{P} X$.

Sledi definicija *ergodičnosti*.

Definicija 2.5.5: Za slabo stacionarnu vremensku seriju X_t sa $E(X_t) = \mu$ kaže se da je **ergodična** u odnosu na srednju vrednost ako važi sledeći uslov:

$$\bar{X} \xrightarrow{P} \mu, T \rightarrow \infty$$

gde je $\bar{X} = \frac{1}{T}(X_1 + X_2 + \dots + X_T) = \frac{1}{T} \sum_{i=1}^T X_i$.

Dakle, vremenska serija je ergodična u odnosu na srednju vrednost ako aritmetička sredina datog skupa podataka konvergira u verovatnoći ka srednjoj vrednosti vremenske serije.

Ergodičnost se može posmatrati kao uslov koji osigurava da će memorija procesa merena sa γ_k slabiti uprosečavanjem tokom vremena i uveden je kako bi se omogućilo modeliranje vremenske zavisnosti slučajnog procesa korišćenjem konačnog skupa parametara.

Dovoljan uslov je i sledeći:

$$\sum_{k=-\infty}^{+\infty} \gamma_k < \infty$$

da bi stacionarna vremenska serija bila ergodična u odnosu na srednju vrednost.

Ergodičnost je ispunjena za većinu stacionarnih procesa koji se sreću u praksi. Ako je proces ergodičan znači da vremenski prosek za jednu realizaciju procesa konvergira ka proseku celog ansambla realizacija.

Ograničenja na vremensku heterogenost i memoriju procesa obezbeđuju izvođenje asimptotskih rezultata bitnih za statističko zaključivanje u okviru analize vremenskih serija.

2.6. AUTOKOVARIJACIONA I AUTOKORELACIONA FUNKCIJA

Autokovarijaciona i autokorelaciona funkcija omogućavaju sagledavanje korelacione strukture vremenske serije što omogućava da se izabere odgovarajući model stacionarne vremenske serije. U prethodnom poglavlju definisan je pojam slabo stacionarne vremenske serije koji će se ovde koristiti.

Definicija 2.6.1: Neka je X_t slabo stacionarna i ergodična vremenska serija, čija je srednja vrednost μ . **Autokovarijacioni koeficijent na doznji, (sa kašnjenjem k), γ_k** za vremensku seriju X_t definiše se na sledeći način:

$$\gamma_k = \text{cov}(X_t, X_{t-k}) = E(X_t - \mu)(X_{t-k} - \mu), k = 1, 2, \dots$$

On predstavlja kovarijaciju između vrednosti vremenske serije koje se nalaze na rastojanju k .

Definicija 2.6.2: Niz $\gamma_1, \gamma_2, \dots$ naziva se **autokovarijaciona funkcija**.

Za $k = 0$, autokovarijacioni koeficijent se svodi na disperziju procesa X_t : $\gamma_0 = \text{var}(X_t) = E(X_t - \mu)^2$.

Međusobna zavisnost između članova vremenske serije odražava se i na vrednost **koeficijenta autokorelacije**. Autokorelacioni koeficijent prvog reda je mera zavisnosti između uzastopnih članova serije, dok je autokorelacioni koeficijent k -tog reda mera zavisnosti između svakog k -tog člana.

Definicija 2.6.3: **Autokorelacioni koeficijent na doznji k , ρ_k** predstavlja koeficijent korelacije između X_t i X_{t-k} i dat je sa:

$$\rho_k = \frac{\text{cov}(X_t, X_{t-k})}{\sqrt{\text{var}(X_t)\text{var}(X_{t-k})}} = \frac{\text{cov}(X_t, X_{t-k})}{\text{var}(X_t)} = \frac{\gamma_k}{\gamma_0}, k = 1, 2, \dots$$

Definicija 2.6.4: Skup autokorelacionih koeficijenata ρ_1, ρ_2, \dots koji su uređeni u odnosu na vreme naziva se **autokorelaciona funkcija**.

Definicija 2.6.5: Grafički prikaz vrednosti autokorelacionih koeficijenata u odnosu na doznje 1, 2, ... zove se **korelogram**.

U nastavku se definišu ocene autokovarijacione i autokorelacione funkcije.

Neka je niz X_1, X_2, \dots, X_T slabo stacionarna i ergodična vremenska serija obima T .

Definicija 2.6.6: Kao ocenu srednje vrednosti μ koristi se aritmetička sredina \bar{X} ,

$$\bar{X} = \frac{1}{T}(X_1 + X_2 + \dots + X_T) = \frac{1}{T}\sum_{i=1}^T X_i.$$

Definicija 2.6.7: U literaturi se najčešće predlažu sledeće dve ocene autokovarijacionih koeficijenata γ_k :

$$1. \hat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})$$

$$2. \hat{\gamma}_k = \frac{1}{T-k} \sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X}), k = 1, 2, \dots$$

Definicija 2.6.8: Ocena autokorelacionog koeficijenta ρ_k se definiše na sledeći način:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=k+1}^T (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2}, k = 1, 2, \dots$$

Definicija 2.6.9: Grafički prikaz ocena autokorelacionih koeficijenata u odnosu na dobnje 1,2,... naziva se **uzorački korelogram** i uobičajeno se označava sa *ACF* (engl. *autocorrelation function*).

Uzorački korelogram služi za prepoznavanje modela stacionarnih vremenskih serija i pomoću njega se može proceniti i da li vremenska serija ima trend, sezonsku komponentu, ili obe komponente. U poglavlju 5 ovo će biti praktično prikazano.

U opštem slučaju vremenskih serija očekuje se da međusobna zavisnost članova serije opada sa kašnjenjem k (pretpostavlja se i dalje stacionarnost i ergodičnost vremenske serije). Ako je vremenska serija sastavljena od međusobno nezavisnih članova, svi autokorelacioni koeficijenti su jednaki nuli. S obzirom da se na osnovu uzorka ne mogu dobiti koeficijenti korelacije jednaki tačno nuli, potrebno je odgovarajućim testom utvrditi da li se oni značajno razlikuju od nule. Jedan od takvih testova je Bartletov¹ test (metod) koji je detaljno opisan u nastavku.

Ocena autokorelacionog koeficijenta, $\hat{\rho}_k$, pod opštim uslovima je konzistentna ocena ρ_k , a na uzorcima malog obima ona je i pristrasna, pri čemu je pristrasnost reda $1/T$.

Ako se pretpostavi da su članovi vremenske serije međusobno nezavisne i jednako raspodeljene slučajne promenljive sa konačnom disperzijom, ocena $\hat{\rho}_k$ za $k \neq 0$ je asimptotski normalno raspodeljena sa parametrima 0 i $1/T$, tako da slučajna promenljiva $\hat{\rho}_k \sqrt{T}$ asimptotski ima normalnu standardizovanu raspodelu.

Ako je vremenska serija slabo stacionarna i ako se može izraziti u funkciji od normalno raspodeljene slučajne komponente, za $\hat{\rho}_k = 0, k > 1$, ocena $\hat{\rho}_k$ je asimptotski normalno raspodeljena sa disperzijom $\frac{1}{T} [1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_k^2]$. Dakle, oceni $\hat{\rho}_1$ odgovara disperzija $1/T$, $\hat{\rho}_2$ disperzija $\frac{1}{T} [1 + 2\rho_1^2]$ itd.

Dakle, postoje dva pristupa u određivanju disperzije ocene $\hat{\rho}_k$:

1) Gruba aproksimacija, podrazumeva da je korelacija između vremenske serije zanemarljivo mala ($\rho_1 = 0, \rho_2 = 0, \dots$). Tada je $var(\hat{\rho}_k) = \frac{1}{T}$.

2) Bartletov metod- vodi računa o korelaciji između članova vremenske serije. Tada je $var(\hat{\rho}_k) = \frac{1}{T} [1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_k^2], 1 < k$.

Ocena disperzije $\hat{\rho}_k, s^2(\hat{\rho}_k)$ je tada:

1) $\frac{1}{T}$ (Gruba aproksimacija)

2) $\frac{1}{T} [1 + 2\hat{\rho}_1^2 + 2\hat{\rho}_2^2 + \dots + 2\hat{\rho}_k^2]$ (Bartletova formula)

Jedan od osnovnih zadataka analize vremenskih serija je identifikacija dobnji na kojima postoji statistički značajna korelacija članova vremenske serije. Zato se proverava vrednost nulte hipoteze $H_0: \rho_k = 0$, protiv alternativne $H_1: \rho_k \neq 0, k = 1, 2, \dots$

Ako se koristi gruba aproksimacija postupak testiranja je sledeći:

¹ M. S. Bartlett (1910- 2002), engleski statističar.

Ukoliko je tačna nulta hipoteza, količnik $\frac{\hat{\rho}_k}{s(\hat{\rho}_k)}$ ima $N(0,1)$ raspodelu. Kada se $s(\hat{\rho}_k)$ zameni sa $\sqrt{1/T}$ dobija se da slučajna veličina $\hat{\rho}_k\sqrt{T}$ ima normalnu standardizovanu raspodelu. Odavde sledi:

$$P[-1.96 \leq \hat{\rho}_k\sqrt{T} \leq 1.96] = 0.95, \text{ tj. } P\left[-\frac{1.96}{\sqrt{T}} \leq \hat{\rho}_k \leq \frac{1.96}{\sqrt{T}}\right] = 0.95.$$

Nulta hipoteza o odsustvu autokorelacije na docnji k se ne može odbaciti ukoliko je ocena autokorelacionog koeficijenta na toj docnji k , $\hat{\rho}_k$ element intervala $(-1.96/\sqrt{T}, 1.96/\sqrt{T})$.

Nulta hipoteza se odbacuje na nivou značajnosti 5% i prihvata alternativna o postojanju autokorelacije na docnji k , ako ocena autokorelacionog koeficijenta $\hat{\rho}_k$ ne pripada intervalu $(-1.96/\sqrt{T}, 1.96/\sqrt{T})$.

Bartletov test je test slučajnosti, pa se hipoteze H_0 i H_1 mogu definisati i na sledeći način:

H_0 : Niz je slučajan (članovi vremenske serije su međusobno nezavisni)

H_1 : Niz nije slučajan (članovi vremenske serije su međusobno zavisni)

Detaljnije o testiranju slučajnosti biće opisano u narednom poglavlju.

2.7. TESTOVI SLUČAJNOSTI

Testiranje slučajnosti je od velikog značaja u statističkoj analizi podataka jer je dosta načina statističke analize zasnovano na pretpostavci da podaci obrazuju slučajan uzorak. U kontekstu analize vremenskih serija, testiranjem slučajnosti serije reziduala proverava se adekvatnost modela.

Cilj je dakle proveriti da li je vremenska serija slučajan uzorak, odnosno proveriti **hipotezu o slučajnosti**. To podrazumeva da se radi o realizacijama nekog niza slučajnih veličina $X_1, X_2, \dots, X_{n-1}, X_n$ koje su međusobno nezavisne i imaju istu raspodelu verovatnoća.

Polazna hipoteza označava se sa H_0 i zove **nulta hipoteza**. Ona označava slučajnost u gore navedenom smislu. Ispravnost te hipoteze znači da se postojeći podaci pri izučavanju mogu uzeti u proizvoljnom poretku.

Ako ova hipoteza ne važi, onda se postojeći vremenski redosled podataka ne sme menjati.

Izbor načina proveravanja nulte hipoteze zavisi od **alternativnih pretpostavki**. One se razlikuju od testa do testa i neke test statistike su više, a neke manje osetljive na određene alternativne hipoteze.

Neki od testova su: Test tačaka zaokreta, Test tačaka rasta, Test razlika rangova, Fon Nojmanov i Bartelsov test, Kendalov test, Ficov test, Goldov test...

U prethodnom poglavlju 2.6. je bio prikazan i test pomoću serijskih korelacija.

Za proveru da li postoji autokorelacija sa korakom 1 može se koristiti test Durban- Vatsona. Takođe se za proveru slučajnosti može koristiti i Boks- Pirsov i Ljung- Boksov test.

U ovom poglavlju korišćena je literatura [1], [2], [3], [15], [20].

3. METOD DEKOMPOZICIJE VREMENSKIH SERIJA

Postoje različiti pristupi analizi vremenskih serija i u skladu sa tim, različiti metodi. U ovoj sekciji biće opisan metod dekompozicije, koji će biti detaljnije razmatran u ovom radu.

Metod dekompozicije vremenskih serija polazi od pretpostavke da je vremenska serija sačinjena od četiri komponente: trenda, ciklične komponente, sezonske komponente i slučajne komponente.

Dekompozicija sa jedne strane predstavlja razlaganje vremenske serije na njene komponente, dok sa druge strane predstavlja mogućnost da se pojedine komponente po potrebi eliminišu iz vremenske serije. U literaturi se ovaj metod zove i *klasični metod statističke analize*.

U postupku primene metoda dekompozicije trend, ciklična i slučajna komponenta se ocene na osnovu raspoloživih podataka, a zatim se njihovom ekstrapolacijom formira prognoza budućih vrednosti vremenske serije. Iako se ova metoda može koristiti za predviđanje, njena osnovna primenljivost je bolje razumevanje vremenskih serija.

Postupci dekompozicije koriste se u vremenskim serijama za opis trenda i sezonskih faktora u vremenskoj seriji. Opširnija dekompozicija može takođe da uključi dugotrajne cikluse, efekte praznika, efekte dana u nedelji i tako dalje.

Jedan od glavnih ciljeva dekompozicije je ocena sezonskih efekata koji se mogu koristiti za kreiranje i prikazivanje sezonski prilagođenih vrednosti. Sezonsko prilagođavanje uklanja sezonske efekte tako da se trendovi mogu videti jasnije.

Mnoge poslovne i ekonomske vremenske serije održavaju i objavljuju vladine agencije kao što su Biro za popis i Biro za statistiku rada. Ove agencije koriste dekompoziciju vremenskih serija da bi kreirali desezonirane vremenske serije.

Razumevanje šta se zaista dešava sa vremenskim serijama često zavisi od korišćenja desezoniranih podataka.

Primer 1: Moglo bi biti od interesa da li se potrošnja električne energije povećava u nekom području. Pretpostavlja se da se zna da je potrošnja električne energije u septembru smanjena za 3% u odnosu na prethodni mesec. Potrebno je biti oprezan pri korišćenju takvih podataka, jer kad god je prisutan sezonski uticaj, takva poređenja mogu biti pogrešna ako podaci nisu desezonirani. Činjenica da je potrošnja električne energije smanjena za 3% od avgusta do septembra može biti samo sezonski efekat povezan sa smanjenjem upotrebe klima uređaja, a ne zbog dugoročnog pada upotrebe električne energije. Nakon prilagođavanja sezonskom učinku, moglo bi se na primer ustanoviti da se upotreba električne energije čak i povećala.

Primer 2: U mnogim regionima SAD nezaposlenost se smanjuje tokom leta usled povećane zaposlenosti u poljoprivrednim oblastima. Stoga pad stope nezaposlenosti u junu u odnosu na maj ne znači nužno da postoji trend smanjenja nezaposlenosti u zemlji. Da bi se videlo da li postoji pravi trend, trebalo bi se prilagoditi činjenici da je nezaposlenost u junu uvek manja nego u maju.

Mnoge druge vremenske serije, koje se odnose na statistiku nezaposlenosti, kućnu prodaju i maloprodaju, izložene su snažnim sezonskim uticajima. Važno je desezonirati takve podatke pre nego što se donese zaključak o postojanju bilo kakvog dugoročnog trenda.

Ukoliko se uklone i trend i sezonske varijacije, mogu se modelirati slučajne greške kao neki oblik stacionarnog procesa.

Da bi dekompozicija serije bila uspešno izvedena, pre svega je potrebno odrediti da li je u pitanju **aditivna serija** (kod koje je opservacija vremenske serije jednaka zbiru komponenti), **multiplikativna serija** (kod koje je opservacija vremenske serije jednaka proizvodu komponenti), ili neka "**srednja**" varijanta (mešoviti model vremenske serije). Opis ovakvih modela dat je u poglavlju 4.

3.1. OSNOVNI KORACI DEKOMPOZICIJE

1. Prvi korak je **ocena trenda**. Za to bi se mogla koristiti dva različita pristupa (sa varijacijama svakog). Jedan od pristupa je oceniti trend postupkom izravnavanja, poput metode pokretnih proseka. Uz ovaj pristup se ne koristi jednačina za opisivanje trenda. Drugi pristup je modeliranje trenda regresijskom jednačinom.

Sledi objašnjenje prvog postupka.

Neka je $\{X_t, t = 1, \dots, n\}$ vremenska serija i pretpostavlja se da u seriji postoje značajne fluktuacije.

Tada se izravnavanje može sprovesti formiranjem klizećih sredina na sledeći način:

$$\tilde{X}_2 = \frac{X_1 + X_2 + X_3}{3}, \tilde{X}_3 = \frac{X_2 + X_3 + X_4}{3}, \dots, \tilde{X}_{n-1} = \frac{X_{n-2} + X_{n-1} + X_n}{3}$$

Na taj način se dobija nova vremenska serija $\tilde{X}_2, \tilde{X}_3, \dots, \tilde{X}_{n-2}, \tilde{X}_{n-1}$.

Računanje se generalno može vršiti i pomoću $(2m + 1)$ tačaka:

$$\tilde{X}_t = \frac{1}{2m + 1} \sum_{j=-m}^m X_{t+j}, t = \overline{m + 1, n - m}$$

sa **jednakim ponderima** $c_j = \frac{1}{2m+1}$, ili sa **opštim ponderima**:

$$\tilde{X}_t = \sum_{j=-m}^m c_j X_{t+j}, t = \overline{m + 1, n - m}$$

koji zadovoljavaju uslov

$$\forall c_j > 0, \sum_{j=-m}^m c_j = 1.$$

Dakle, **vrednost trenda** u vremenskoj seriji u momentu t se ocenjuje vrednošću \tilde{X}_t .

2. Drugi korak je **uklanjanje trenda** iz vremenske serije. Za aditivnu dekompoziciju vrši se oduzimanje ocena trendova iz serije. Za multiplikativnu dekompoziciju vrši se deljenje serija ocenjenim vrednostima trenda.

3. Zatim, **sezonski faktori se ocenjuju** korišćenjem vremenske serije iz koje je uklonjen trend. Za mesečne podatke, to podrazumeva ocenu učinka za svaki mesec u godini. Za kvartalne podatke, to podrazumeva ocenu učinka za svako tromesečje. Najjednostavnija metoda za ocenu ovih efekata je da se uradi prosek de- trendiranih vrednosti za određenu sezonu. Na primer, da bi se postigao sezonski efekat za januar, uradi se prosek de- trendiranih vrednosti za svaki mesec januar u vremenskoj seriji, i tako dalje.

Broj sezonskih faktora jednak je frekvenciji vremenske serije (npr. mesečni podaci = 12 sezonskih faktora, kvartalni podaci = 4 itd.).

4. Sezonske efekte bi trebalo normalizovati.

- Za aditivni model, sezonski efekti se obično prilagođavaju tako da prosek s sezonskih komponenti bude 0 (ovo je ekvivalentno tome da njihov zbir bude jednak 0);

- Za multiplikativni model, s sezonskih efekata su prilagođeni tako da im je prosek 1 (ovo je ekvivalentno tome da je njihov zbir jednak s);

5. Sledeći korak je određivanje slučajne komponente.

Za aditivni model: slučajna komponenta = vremenska serija – (ocenjeni)trend – (ocenjena)sezonska komponenta.

Za multiplikativni model: slučajna komponenta = vremenska serija / ((ocenjeni)trend* (ocenjena)sezonska komponenta)

6. Analiziranje slučajne komponente. Koji god metod dekompozicije se koristio cilj je proizvesti **stacionarne** rezidualne.

7. Izbor modela koji odgovara stacionarnim rezidualima. Primer može biti neki od ARMA modela. Ovi modeli će biti detaljno objašnjeni u poglavlju koji sledi (4.2.).

8. Prognoziranje vremenske serije se može postići prognoziranjem reziduala i kombinujući sa prognozama trenda i sezonskim komponentama.

U sledećem poglavlju biće detaljnije opisani modeli vremenskih serija.

U ovom poglavlju korišćena je literatura [1], [3], [4], [5], [7], [13], [16].

4. MODELI VREMENSKIH SERIJA

4.1. OPIS MODELA

U sekciji 2.3. su opisane komponente vremenskih serija. Neka je X_t opservacija vremenske serije u vremenskom trenutku t kao što je definisano u sekciji 2.1. Uvode se sledeće oznake za komponente vremenske serije:

f_t – trend
 S_t – sezonska komponenta
 C_t – ciklična komponenta
 e_t – slučajna komponenta

4.1.1. ADITIVNI MODEL

Prema **aditivnom modelu**, varijacije pojave predstavljaju zbir trenda, sezonske, ciklične i slučajne komponente (ako se vremenska serija sastoji npr. od godišnjih podataka onda se isključuje sezonska komponenta):

$$X_t = f_t + S_t + C_t + e_t$$

Aditivni model pretpostavlja da se komponente vremenske serije menjaju nezavisno jedna od druge. Na primer, povećanje trenda neće uzrokovati povećanje sezonske komponente.

Ovaj model se koristi ako su sezonska, ciklična i slučajna komponenta nezavisne od ponašanja (kretanja) trenda, tj. ako su sezonski efekti isti iz godine u godinu. Dakle, on je koristan kada je sezonska varijacija relativno stabilna tokom vremena.

Primećuje se da u aditivnom modelu svaka komponenta ima istu dimenziju kao i originalna vremenska serija. Desezonirani (sezonski prilagođeni) podaci se mogu izraziti sa:

$$SA_t = X_t - S_t$$

Ograničenja za aditivni model su:

1. Pretpostavlja se da su sezonska i slučajna komponente nezavisne od trenda
2. Sezonska komponenta ostaje stabilna iz godine u godinu
3. Sezonska fluktuacija u proseku iznosi nula tokom godine

$$\sum_{j=1}^s S_{t+j} = 0$$

sumiranjem po j za $s = 12$ za mesečne vremenske serije ili $s = 4$ za kvartalne vremenske serije.

4.1.2. MULTIPLIKATIVNI MODEL

Po **multiplikativnom modelu** podaci vremenske serije se predstavljaju kao proizvod trenda, sezonske, ciklične i slučajne komponente:

$$X_t = f_t \cdot S_t \cdot C_t \cdot e_t$$

Multiplikativni model pretpostavlja da sve komponente međusobno zavise jedna od druge, pa stoga, ako raste trend, fluktuacije izazvane sezonom, ciklusom ili slučajnim kretanjem će takođe rasti.

Ovaj model se koristi ako se sa povećavanjem (smanjivanjem) nivoa pojave proporcionalno povećava (smanjuje) i uticaj sezonske komponente tokom vremena.

U multiplikativnom modelu trend ima istu dimenziju kao i originalna vremenska serija, međutim sezonska i slučajna komponenta su bezdimenzijski faktori centrirani oko jedinice. Desezonirani podaci mogu se izraziti sa:

$$SA_t = \frac{X_t}{S_t}$$

Multiplikativni model zahteva ograničenje

$$\sum_{j=1}^s S_{t+j} = s$$

sumiranjem po j za $s = 12$ za mesečne vremenske serije ili $s = 4$ za kvartalne vremenske serije.

Ovo ograničenje pomaže da se osigura da sezonski faktori ne obuhvate dugoročne cikluse u vremenskoj seriji. Na primer, ako sezonski faktori nisu centrirani oko jedinice, tj. ako je $\sum_{j=1}^s S_{t+j} \neq s$ gde je s broj perioda u godini, tada bi sezonski faktori obuhvatili neke od godišnjih promena u osnovnom nivou serije. Sa druge strane, ograničavanje ovih sezonskih faktora na sumu s tj. $\sum_{j=1}^s S_{t+j} = s$ osigurava da sezonski faktori imaju prosek veličine 1 i stoga ne daju zbirni doprinos nivou serije tokom godine.

Kada je jedan period sezonski veći (sezonski faktor veći od 1) u odnosu na ostale periode, mora postojati barem još jedan period koji to nadoknađuje tako što je sezonski nizak (sezonski faktor manji od 1), tako da je suma tokom 12 meseci (ili 4 kvartala) perioda i dalje s .

Multiplikativni model može takođe biti zapisan u aditivnoj formi korišćenjem logaritma:

$$\log X_t = \log f_t + \log S_t + \log C_t + \log e_t$$

Multiplikativni modeli ili log-aditivni modeli se ne mogu koristiti kada originalna vremenska serija sadrži nule ili negativne vrednosti. U ovom slučaju može se koristiti mešoviti model.

4.1.3. MEŠOVITI (KOMBINOVANI) MODEL

Ako se radi o **mešovitom modelu** podaci o pojavi mogu da budu rezultat različitih kombinacija dejstva trenda, sezonske, ciklične i slučajne komponente, na primer:

$$X_t = f_t \cdot S_t \cdot C_t + e_t$$

Jedan od primera mešovitog modela je i *pseudo-aditivni model*. Njegova formula je:

$$X_t = f_t \cdot C_t \cdot (S_t + e_t - 1)$$

On kombinuje elemente aditivnih i multiplikativnih veza i pretpostavlja da su sezonska i slučajna komponenta zavisne od trenda, ali nezavisne međusobno.

Uklanjanje multiplikativne veze između sezonske i slučajne komponente je posebno korisno ako serija sadrži jednu ili više nula vrednosti. To je zato što aditivna veza između sezonskih i slučajnih komponenti omogućava bilo kojoj od njih da apsorbuje nula vrednosti bez narušavanja ponašanja trenda vremenske serije.

Ovaj model se može zapisati i na sledeći način:

$$\begin{aligned} X_t &= f_t + f_t \cdot (S_t - 1) + f_t \cdot (e_t - 1) \\ &= f_t \cdot (S_t + e_t - 1) \end{aligned}$$

Kao i kod multiplikativnog modela, i u pseudo-aditivnom modelu trend ima iste dimenzije kao u originalnoj vremenskoj seriji, a sezonske i slučajne komponente nemaju dimenziju. Desezonirani podaci se mogu definisati na sledeći način:

$$SA_t = X_t - f_t \cdot (S_t - 1)$$

I kod pseudo-aditivnog modela, kao i kod multiplikativnog, sezonska i slučajna komponenta su centrirane oko jedinice. Zato mora da se oduzme 1 od S_t i e_t i omogući da su izrazi $f_t \cdot (S_t - 1)$ i $f_t \cdot (e_t - 1)$ centrirani oko nule.

Ovi izrazi se mogu tumačiti kao aditivna sezonska i aditivna slučajna komponenta, i pošto su centrirani oko nule, originalne vrednosti X_t biće centrirane oko vrednosti trenda f_t .

4.1.4. KAKO ODABRATI ODGOVARAJUĆI MODEL

Da bi se odabrao odgovarajući model vremenske serije potrebno je pre svega ispitati originalnu vremensku seriju, a zatim i isprobati niz modela i pronaći najpogodniji.

Nekada se odmah pomoću grafičkog prikaza vremenske serije može dobiti zaključak koji je model odgovarajući, a nekad je potrebna detaljnija analiza.

Kao što je i rečeno u poglavlju 4.1.1., aditivni model je koristan kada je sezonska varijacija relativno konstantna (stabilna) tokom vremena, dok je multiplikativni model koristan kada se sezonska varijacija npr. povećava tokom vremena (poglavlje 4.1.2.).

Prilikom analize ekonomskih pojava najčešće se koristi multiplikativni model po kome je uticaj faktora koji deluju na komponente relativan. Tačnije, trend je iskazan u mernim jedinicama kao i sama vremenska serija, a uticaji ostalih komponenti su dati u procentima od trenda (ili ostalih komponenti). Kod aditivnih modela sve komponente su iskazane u istim jedinicama kao i originalna vremenska serija. Multiplikativni modeli nisu pogodni ako su vrednosti nekih komponenti bliske nuli ili ako imaju negativan predznak. Tada je uglavnom najbolje koristiti neki mešoviti model.

Prilikom izbora modela koristan je metod dekompozicije koji je opisan u poglavlju 3. On daje prikaz svake komponente, te se tako može jasno videti trend, sezonski obrazac ako postoji, kao i da li je slučajna komponenta zaista slučajni obrazac.

Da bi se znalo koji model je odgovarajući za datu vremensku seriju potrebno je odrediti i autokorelacione i autokovarijacione funkcije, zato što one imaju specifične oblike za pojedinačne, različite modele. Može se koristiti grafički uzorački korelogram pomoću kojeg se ispituje struktura korelacije izvornih podataka, ali i slučajne komponente modela i na osnovu toga doneti zaključak o modelu.

Pomoću primera u poglavljima 5. i 6. koji slede biće praktično prikazan postupak odabira odgovarajućeg modela vremenskih serija.

4.2. MODELI AR, MA, ARMA, ARIMA

U poglavlju 3 pomenuto je da je u koracima dekompozicije vremenskih serija potrebno napraviti izbor modela koji odgovara stacionarnim rezidualima, te su stoga oni bitni i biće detaljno pojašnjeni u ovom poglavlju. Pre nego što se objasne modeli kojima se mogu opisati slabo stacionarne vremenske serije uvodi se pojam **linearnog procesa**.

Prema *Voldovoj² teoremi razlaganja* svaka slabo stacionarna vremenska serija se može predstaviti kao zbir jedne determinističke komponente i jedne stohastičke komponente koje su međusobno nekorelisane.

² H. O. A. Wold (1908- 1992), švedski matematičar

Deterministička komponenta predstavlja konstantu (srednju vrednost) i/ili funkciju linearnog trenda. Ona se može predvideti na osnovu informacija iz prošlosti. Sa druge strane, slučajna komponenta se ne može rekonstruisati prema sopstvenom kretanju iz prošlosti.

Teorema 4.2.1 (Voldova teorema razlaganja): Neka je sa e_t označen proces beli šum. Stohastička komponenta slabo stacionarne vremenske serije X_t može se predstaviti u formi linearnog procesa oblika:

$$X_t - \mu = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \dots = \sum_{i=0}^{\infty} \psi_i e_{t-i}, \psi_0 = 1,$$

pri čemu je $\mu = E(X_t)$ deterministička komponenta vremenske serije X_t , dok su (ψ_1, ψ_2, \dots) parametri pomoću kojih je opisana dinamička struktura vremenske serije i oni se uobičajeno nazivaju ψ -ponderi.

Upotrebom operatora docnje prvog reda L , reprezentacija linearnog procesa se može zapisati i na sledeći način:

$$X_t - \mu = e_t + \psi_1 L e_t + \psi_2 L^2 e_t \dots = (1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots) e_t = \psi(L) e_t$$

Postoje tri vrste modela kojima se mogu opisati slabo stacionarne vremenske serije:

1. Autoregresioni modeli (AR)
2. Modeli pokretnih proseka (MA)
3. Autoregresioni modeli pokretnih proseka (ARMA)

4.2.1. AUTOREGRESIONI MODELI (AR)

Definicija 4.2.1.1: Autoregresioni modeli vremenskih serija su oni modeli u kojima se zavisna promenljiva u trenutku t opisuje u funkciji od sopstvenih prethodnih vrednosti, tačnije, skup objašnjavajućih promenljivih su elementi iste vremenske serije samo u trenucima $t - 1, t - 2, \dots, t - p$.

Autoregresioni model reda p u oznaci $AR(p)$ definiše se na sledeći način:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t$$

gde su $\phi_1, \phi_2, \dots, \phi_p$ autoregresioni parametri, a e_t je proces beli šum.

Autoregresioni model prvog reda zapisuje se u oznaci $AR(1)$ i predstavlja se na sledeći način:

$$X_t = \phi_1 X_{t-1} + e_t$$

Vremenska serija opisana ovim modelom je stacionarna ako je $|\phi_1| < 1$.

Ovakva vremenska serija se može predstaviti kao specijalan slučaj linearnog procesa:

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + e_t \\ X_t &= \phi_1 L X_t + e_t \\ (1 - \phi_1 L) X_t &= e_t \\ X_t &= \frac{1}{(1 - \phi_1 L)} e_t \end{aligned}$$

Uz gore pomenut uslov $|\phi_1| < 1$, izraz $\frac{1}{(1-\phi_1 L)}$ predstavlja zbir članova geometrijskog reda $(1 + \phi_1 L + \phi_1^2 L^2 + \phi_1^3 L^3 + \dots)$ pa kad se ovaj izraz zameni u gornju jednačinu dobije se sledeće:

$$\begin{aligned} X_t &= (1 + \phi_1 L + \phi_1^2 L^2 + \phi_1^3 L^3 + \dots) e_t \\ &= e_t + \underbrace{\phi_1}_{\psi_1} e_{t-1} + \underbrace{\phi_1^2}_{\psi_2} e_{t-2} + \underbrace{\phi_1^3}_{\psi_3} e_{t-3} + \dots \end{aligned}$$

Dakle, autoregresioni model prvog reda je specijalan slučaj linearnog procesa pri uslovu $\psi_j = \phi_1^j, j = 1, 2, 3, \dots$ Iz ovoga sledi i da je $E(X_t) = 0$.

Slično je i za modele ostalih redova.

4.2.2. MODELI POKRETNIH PROSEKA (MA)

Definicija 4.2.2.1: Model pokretnih proseka reda q u oznaci $MA(q)$ definiše se na sledeći način:

$$X_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

gde je sa e_t označen proces beli šum, dok su $\theta_1, \theta_2, \dots, \theta_q$ parametri modela.

Nivo vremenske serije u trenutku t se opisuje u funkciji od članova procesa belog šuma u trenucima $t, t-1, \dots, t-q$.

S obzirom da je e_t beli šum, zaključujemo da je $E(X_t) = 0$.

I ovaj model je specijalan slučaj linearnog procesa:

$$X_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \dots$$

Gde je $\psi_1 = -\theta_1, \psi_2 = -\theta_2, \dots, \psi_q = -\theta_q, \psi_j = 0, j > q$.

Linearni proces se može označiti kao MA model beskonačnog reda, $MA(\infty)$.

4.2.3. AUTOREGRESIONI MODELI POKRETNIH PROSEKA (ARMA)

Definicija 4.2.3.1: Autoregresioni model pokretnih proseka u oznaci $ARMA(p, q)$, gde je p red autoregresione komponente i q red komponente pokretnih proseka definiše se na sledeći način u obliku opšte forme:

$$\underbrace{(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)}_{\Phi(L)} X_t = \underbrace{(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)}_{\Theta(L)} e_t$$

Proces e_t je beli šum, dok polinomi $\Phi(L)$ i $\Theta(L)$ opisuju respektivno, autoregresionu komponentu i komponentu pokretnih proseka vremenske serije X_t i pretpostavlja se da ne sadrže zajedničke faktore. Model $ARMA(p, q)$ obuhvata kao specijalne slučajeve prethodno razmatrane modele $AR(p)$ i $MA(q)$ na sledeći način: $ARMA(p, 0) = AR(p), ARMA(0, q) = MA(q)$.

4.2.4. AUTOREGRESIONI MODELI POKRETNIH PROSEKA ZA INTEGRISANE VREMENSKE SERIJE (ARIMA)

Definicija 4.2.4.1: Autoregresioni model pokretnih proseka za integrisane vremenske serije u oznaci $ARIMA(p, d, q)$, gde je p red autoregresione komponente, d nivo integrisanosti vremenske serije i q red komponente pokretnih proseka opisuje se sledećom formulom:

$$\underbrace{(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)}_{\Phi(L)} (1 - L)^d X_t = \theta_0 + \underbrace{(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)}_{\Theta(L)} e_t$$

Proces e_t je beli šum, dok polinomi $\Phi(L)$ i $\Theta(L)$ opisuju respektivno, autoregresionu komponentu i komponentu pokretnih proseka stacionarne vremenske serije $(1 - L)^d X_t$ i pretpostavlja se da ne sadrže zajedničke faktore.

Ovaj tip modela obuhvata kao specijalne slučajeve prethodno razmatrane modele na sledeći način:

$$ARIMA(p, 0, 0) = AR(p), \quad ARIMA(0, 0, q) = MA(q), \quad ARIMA(p, 0, q) = ARMA(p, q),$$

$ARIMA(0, 0, 0)$ je beli šum.

4.3. SEZONSKE VREMENSKE SERIJE

U sekciji 2.3. definisana je sezonska komponenta vremenske serije i u ovom poglavlju će biti razmatrani modeli koji uključuju tu komponentu. Vremenske serije koje sadrže sezonsku komponentu nazivaju se **i sezonske vremenske serije**.

Prilikom modeliranja sezonskih vremenskih serija značajan je operator sezonske diference Δ_s koji je opisan u sekciji 2.5.:

$$\Delta_s X_t = X_t - X_{t-s} = (1 - L^s) X_t$$

Ovim se dovode u vezu dve vrednosti vremenske serije koje se nalaze na rastojanju s . Trenutak koji kasni s perioda za tekućim naziva se *sezonska docnja*.

Sezonska priroda može biti deterministička ili stohastička, u zavisnosti od toga da li se sezonske varijacije menjaju tokom godina ili ne.

Ako je sezonska priroda deterministička, ona se modelira uključivanjem sezonskih veštačkih promenljivih, dok se stohastička sezonska komponenta opisuje posebnom vrstom ARIMA modela, koji se nazivaju **sezonski ARIMA modeli**.

Stohastičke varijacije sezonskog karaktera se mogu modelirati kroz dve komponente:

1. **Aditivna sezonska komponenta**- sezonske varijacije se dodaju ostalim varijacijama koje se modeliraju elementima ARIMA modela.

Model sa sezonskom komponentom aditivnog tipa je:

$$X_t = e_t - \theta_1 e_{t-1} - \Theta e_{t-s} = (1 - \theta_1 L - \Theta L^s) e_t$$

2. **Multiplikativna sezonska komponenta**- predstavlja interaktivno dejstvo svih varijacija, tj. potrebno je uključiti one komponente koje će biti proizvod standardnih i sezonskih komponenti.

Model sa sezonskom komponentom multiplikativnog tipa je:

$$X_t = e_t - \theta_1 e_{t-1} - \Theta e_{t-s} + \theta_1 \Theta e_{t-s-1} = (1 - \theta_1 L)(1 - \Theta L^s) e_t$$

Pri čemu su, u oba slučaja θ_1 i Θ parametri modela za koje važi $0 < |\theta_1| < 1$, $0 < |\Theta| < 1$ dok je sa e_t označen beli šum.

Na osnovu autokorelacione i autokovarijacione funkcije uočavaju se razlike između aditivne i multiplikativne vremenske serije.

Primer autokorelacione i autokovarijacione funkcije modela sa *aditivnom sezonskom komponentom*:

$$\gamma_k = E(X_t X_{t-k}) = E(e_t - \theta_1 e_{t-1} - \Theta e_{t-s})(e_{t-k} - \theta_1 e_{t-k-1} - \Theta e_{t-k-s})$$

$$\gamma_k = \begin{cases} (1 + \theta_1^2 + \Theta^2)\sigma^2, & k = 0 \\ -\theta_1\sigma^2, & k = 1 \\ \theta_1\Theta\sigma^2, & k = s - 1 \\ -\Theta\sigma^2, & k = s \\ 0, & \text{za ostale vrednosti } k \end{cases}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1, & k = 0 \\ \frac{-\theta_1}{1 + \theta_1^2 + \Theta^2}, & k = 1 \\ \frac{\theta_1\Theta}{1 + \theta_1^2 + \Theta^2}, & k = s - 1 \\ \frac{-\Theta}{1 + \theta_1^2 + \Theta^2}, & k = s \\ 0, & \text{za ostale vrednosti } k \end{cases}$$

Primer autokorelacione i autokovarijacione funkcije modela sa *multiplikativnom sezonskom komponentom*:

$$\gamma_k = E(X_t X_{t-k})E(e_t - \theta_1 e_{t-1} - \Theta e_{t-s} + \theta_1 \Theta e_{t-s-1})(e_{t-k} - \theta_1 e_{t-k-1} - \Theta e_{t-k-s} + \theta_1 \Theta e_{t-k-s-1})$$

$$\gamma_k = \begin{cases} (1 + \theta_1^2 + \Theta^2 + \theta_1^2 \Theta^2)\sigma^2, & k = 0 \\ (-\theta_1 - \theta_1 \Theta^2)\sigma^2, & k = 1 \\ \theta_1 \Theta \sigma^2, & k = s - 1 \\ (-\Theta - \theta_1^2 \Theta)\sigma^2, & k = s \\ \theta_1 \Theta \sigma^2, & k = s + 1 \\ 0, & \text{za ostale vrednosti } k \end{cases}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1, & k = 0 \\ \frac{-\theta_1 - \theta_1 \Theta^2}{1 + \theta_1^2 + \Theta^2 + \theta_1^2 \Theta^2}, & k = 1 \\ \frac{\theta_1 \Theta}{1 + \theta_1^2 + \Theta^2 + \theta_1^2 \Theta^2}, & k = s - 1 \\ \frac{-\Theta - \theta_1^2 \Theta}{1 + \theta_1^2 + \Theta^2 + \theta_1^2 \Theta^2}, & k = s \\ \frac{\theta_1 \Theta}{1 + \theta_1^2 + \Theta^2 + \theta_1^2 \Theta^2}, & k = s + 1 \\ 0, & \text{za ostale vrednosti } k \end{cases}$$

4.3.1. SEZONSKI ARIMA MODELI

Sezonski ARIMA model sa multiplikativnom sezonskom komponentom se opisuje sledećom formulom:

$$(1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_P L^{Ps})(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L^s)^D (1 - L)^d X_t = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_Q L^{Qs})(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) e_t$$

Pri čemu je s period sezone, $\Phi_1, \Phi_2, \dots, \Phi_P$ su parametri sezonske regresione komponente reda P , $\Theta_1, \Theta_2, \dots, \Theta_Q$ su parametri sezonske komponente pokretnih sredina reda Q , D nivo sezonske integrisanosti. Ostali parametri definisani su u poglavlju 4.2.

Ovaj model se može zapisati u oznaci $ARIMA(p, d, q) \times (P, D, Q)_s$.

4.3.2. PRIMERI SEZONSKIH ARIMA MODELA

U ovom poglavlju dati su neki primeri sezonskih ARIMA modela.

4.3.2.1. Sezonski autoregresioni model prvog reda

Ovaj model se zapisuje u oznaci $AR(1)_s$. Prethodno opisan model svodi se na ovaj model prvog reda za $P = 1, p = q = Q = d = D = 0$:

$$(1 - \Phi_1 L^s) X_t = e_t, |\Phi_1| < 1 \\ X_t = \Phi_1 X_{t-s} + e_t$$

Autokorelaciona funkcija je:

$$\rho_{ks} = \Phi_1^k, k = 1, 2, \dots$$

Koeficijenti autokorelacione funkcije sadrže nenulte vrednosti isključivo na sezonskim docnjama $s, 2s, 3s, \dots$ dok vrednosti opadaju po eksponencijalnoj putanji.

4.3.2.2. Sezonski integrisani model prvog reda

Ako je kod prethodnog modela $AR(1)_s$ parametar Φ_1 jednak jedan, onda se opšti model $ARIMA(p, d, q) \times (P, D, Q)_s$ svodi na model oblika $ARIMA(0, 0, 0) \times (0, 1, 0)_s$ gde je $D = 1, P = p = q = Q = 0$ i zapisuje na sledeći način:

$$(1 - L^s) X_t = e_t \\ X_t = X_{t-s} + e_t$$

Ovako definisana vremenska serija naziva se *slučajan hod sezonskog tipa*. Autokorelaciona funkcija vremenske serije postepeno opada tokom vremena po sezonskim docnjama $s, 2s, 3s, \dots$. Opadanje se ostvaruje od vrednosti koja je bliska vrednosti jedan. Ovaj vid nestacionarnosti može da se otkloni primenom sezonskog diferenciranja: $\Delta_s X_t = (1 - L^s) X_t = X_t - X_{t-s} = e_t$.

4.3.2.3. Sezonski model pokretnih proseka prvog reda

Ovaj model se zapisuje u oznaci $MA(1)_s$ i dobija se iz početnog modela za $Q = 1, P = p = q = D = d = 0$:

$$X_t = (1 - \Theta_1 L^s) e_t \\ X_t = e_t - \Theta_1 e_{t-s}$$

Autokorelaciona funkcija u modelu $MA(1)_s$ ima jednu nenultu vrednost na sezonskoj docnji s :

$$\rho_k = \begin{cases} 1, & k = 0 \\ -\frac{\theta_1}{(1 + \theta_1^2)}, & k = s \\ 0, & \text{za ostale vrednosti } k \end{cases}$$

4.3.2.4. „Vazduhoplovni model“

Model je dobio naziv po engleskom terminu „airline model“. Ova specifikacija korišćena je više puta u literaturi za modeliranje vremenske serije koja predstavlja broj prevezenih putnika u vazdušnom saobraćaju.

Specifikacija ovog modela dobija se polazeći od početnog modela za $D = d = q = Q = 1, P = p = 0$:

$$(1 - L^s)(1 - L)X_t = (1 - \theta_1 L)(1 - \theta_1 L^s) e_t, 0 < |\theta_1| < 1, 0 < |\theta_1| < 1$$

AR deo modela se sastoji od redovne i sezone razlike, dok MA deo uključuje dva parametra.

Ovo je integrisana sezonska vremenska serija prvog reda. Da bi se dobila stacionarna reprezentacija vremenske serije potrebno je primeniti operatore prve i sezone diference:

$$\Delta \Delta_s X_t = (1 - L)(1 - L^s)X_t = X_t - X_{t-1} - X_{t-s} + X_{t-s-1}$$

Tako dobijena vremenska serija predstavlja multiplikativni model pokretnih proseka i on je opisan u sekciji 4.3. Takođe, autokorelaciona funkcija je definisana u istoj sekciji i ona je ista za ovaj model.

Ovaj model ima široku primenu u modeliranju sezonskih vremenskih serija.

U ovom poglavlju korišćena je literatura [2], [4], [5], [7].

5. PRIMENA METODA DEKOMPOZICIJE I IZBOR MODELA

U ovom poglavlju biće prikazana primena metoda dekompozicije na raznim primerima vremenskih serija korišćenjem različitih mogućnosti programskog jezika **R**. Takođe će biti razmatrano koji je model odgovarajući za različite vremenske serije.

U poglavlju 3 je navedeno da metod dekompozicije polazi od pretpostavke da je vremenska serija sačinjena od četiri komponente: trenda, ciklične komponente, sezonske komponente i slučajne komponente.

Vremenska serija može biti i „nesezonska“, odnosno može da se sastoji samo od trenda i slučajne komponente i može se opisati aditivnim modelom. Dekompozicija može biti izvršena i za ovakvu vremenskom seriju, ocenjivanjem trenda i slučajne komponente. Međutim, u ovom radu će biti fokus na one vremenske serije koje imaju sezonsku komponentu i za njih će biti prikazan postupak dekompozicije.

Osnovna i najstarija funkcija za ovaj metod je **decompose()** koja je deo R baze.

Prednost ovog metoda je to što je lak i jednostavan za upotrebu. Njime se brzo dobija predstava o podacima i pruža početnu tačku za dalje istraživanje podataka.

Nedostaci: Prvih nekoliko opservacija daju N/A vrednosti, jer metod koristi pokretne proseke koji zahtevaju neke početne podatke.

Metod je takođe spor da bi se uočili brzi porasti u skupu podataka, a uz to model pretpostavlja da sezonska komponenta ostaje stabilna.

Ovo je pretpostavka koja postaje problematična kada imamo veoma dugu vremensku seriju sa većom šansom za sezonske promene.

Većina slabih tačaka osnovne verzije metoda pokriveno je alternativnim metodama.

Najbolje funkcije za ovaj metod nalaze se u R bazi i u paketima *Forecast* i *Seasonal*:

Jedan od alternativnih metoda je takozvani **stl** metod, i u R bazi takođe postoji funkcija **stl()**.

Još neki od alternativnih metoda su **SEATS** i **x11**. Ako se ovi metodi pokreću, potreban je paket *Seasonal* i funkcija **seas()**. Ova funkcija je vrlo svestrana i ima na raspolaganju nekoliko metoda tako da se mogu prilično dobro pokriti i uporediti svi uobičajeni metodi dekompozicije.

Postoje i dodatne funkcije iz paketa *Forecast* koje integrišu **decompose** i **stl** generisane objekte i omogućavaju prognoziranje podataka sa tim objektima.

5.1. PRIMER 1

Za ovaj primer koristi se **nottem** skup podataka koji se nalazi u R bazi podataka. On predstavlja prosečne mesečne temperature u Notingemu, od 1920. do 1939. godine. Tačnije, ovaj skup podataka je vremenska serija koja sadrži prosečne temperature vazduha u zamku Notingema izražene u Farenhajtima tokom 20 godina.

Prikazuju se prvo vrednosti skupa podataka, a zatim i sumarne statistike pomoću funkcije **summary()** i **sd()**. Funkcijom *summary()* se dobijaju podaci o minimalnoj vrednosti, medijani, srednjoj vrednosti itd. Funkcijom *sd()* računa se standardna devijacija.

Kako bi se dobio bolji uvid u podatke podaci se predstavljaju i grafički pomoću funkcije **plot()**.

```
> nottem
```

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1920	40.6	40.8	44.4	46.7	54.1	58.5	57.7	56.4	54.3	50.5	42.9	39.8
1921	44.2	39.8	45.1	47.0	54.1	58.7	66.3	59.9	57.0	54.2	39.7	42.8
...												
1938	42.1	41.2	47.3	46.6	52.4	59.0	59.6	60.4	57.0	50.7	47.8	39.2
1939	39.4	40.9	42.4	47.8	52.4	58.0	60.7	61.8	58.2	46.7	46.6	37.8

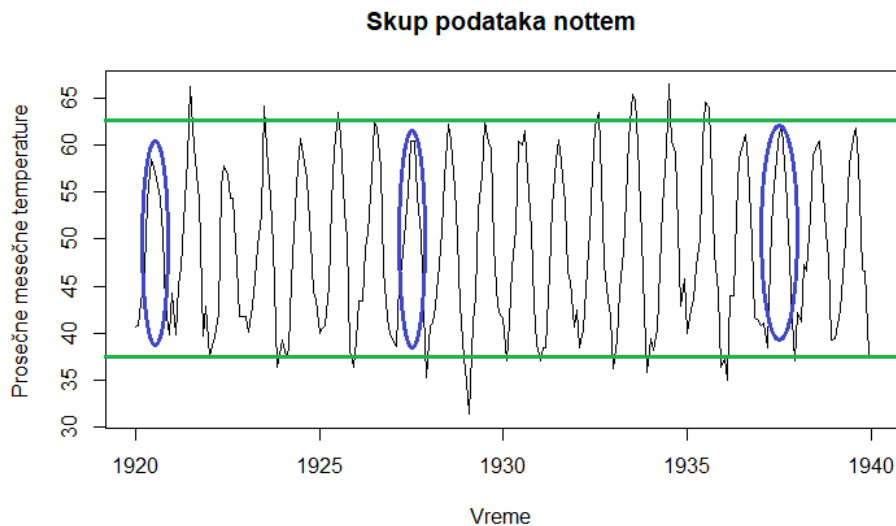
```
> summary(nottem)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
31.30	41.55	47.35	49.04	57.00	66.50

```
> sd(nottem)
```

```
[1] 8.572324
```

```
> plot(nottem, xlab="Vreme", ylab="Prosečne mesečne temperature ",
+       main="Skup podataka nottem")
```



Slika 5.1.1. Grafički prikaz nottem skupa podataka

Sa grafika se vidi da ovaj set podataka ima stabilnu sezonalnost i da nema trend. Udaljenost između uspona i padova u sezoni se ne povećava sa vremenom. Ako bi postojao trend u ovakvom sezonskom skupu podataka, moglo bi se videti kako se vrhovi sezona pomeraju gore ili dole tokom celog raspona vremenske serije. Kao što se ovde vidi, amplituda kretanja prema gore i dole u toku jedne sezone ostaje stabilna, a vrhovi ostaju približno u jednoj liniji.

Dakle, ovaj set podataka se može opisati **aditivnim modelom**.

Kad god se koristi funkcija **decompose()**, potrebno je proveriti da li skup podataka ima unapred postavljenu frekvenciju.

To se proverava pomoću funkcije **frequency()**, a takođe može da se prover i dužina skupa podataka funkcijom **length()**.

```
> frequency(nottem)
[1] 12
> length(nottem)
[1] 240
```

Frekvencija u ovom slučaju iznosi 12 za 12 meseci, dok je dužina 240, što predstavlja 20 godina (20x12). Sada može da se koristi ključna funkcija u proceduri. To je funkcija *decompose()* u R bazi, za koju takođe može da se odredi i tip modela.

Za aditivni model: *decompose(Ime vremenske serije, type = "additive")*
Za multiplikativni model: *decompose(Ime vremenske serije, type = "multiplicative")*

Važno je da se kao preliminarni model mora koristiti naredba **ts()** da bi se definisao sezonski raspon za seriju.

Za kvartalne podatke, to može biti: *Ime vremenske serije = ts(Ime vremenske serije, freq = 4)*

Za mesečne podatke, to može biti: *Ime vremenske serije = ts(Ime vremenske serije, freq = 12)*

Ovaj skup podataka je već vremenska serija koja ima frekvenciju 12, pa nije potrebno da se koristi naredba **ts()**.

Bira se type= "additive" kao parametar funkcije *decompose()* za aditivni tip modela.

```
> decompose(nottem, type = "additive")
```

```
$x
  Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
1920 40.6 40.8 44.4 46.7 54.1 58.5 57.7 56.4 54.3 50.5 42.9 39.8
1921 44.2 39.8 45.1 47.0 54.1 58.7 66.3 59.9 57.0 54.2 39.7 42.8
...
1939 39.4 40.9 42.4 47.8 52.4 58.0 60.7 61.8 58.2 46.7 46.6 37.8

$seasonal
  Jan Feb Mar Apr May Jun
Jul
1920 -9.3393640 -9.8998904 -6.9466009 -2.7573465 3.4533991 8.9865132 12.9
672149
1921 -9.3393640 -9.8998904 -6.9466009 -2.7573465 3.4533991 8.9865132 12.9
672149
...
1939 -9.3393640 -9.8998904 -6.9466009 -2.7573465 3.4533991 8.9865132 12.9
672149
  Aug Sep Oct Nov Dec
1920 11.4591009 7.4001096 0.6547149 -6.6176535 -9.3601974
1921 11.4591009 7.4001096 0.6547149 -6.6176535 -9.3601974
...
1939 11.4591009 7.4001096 0.6547149 -6.6176535 -9.3601974

$trend
  Jan Feb Mar Apr May Jun Jul Au
g Sep
1920 NA NA NA NA NA NA 49.04167 49.1500
0 49.13750
1921 49.56667 50.07083 50.32917 50.59583 50.61667 50.60833 50.45417 50.1291
7 49.85000
...
```

```

1939 49.67917 49.78333 49.89167 49.77500 49.55833 49.45000      NA      N
A      NA      Oct      Nov      Dec
1920 49.17917 49.19167 49.20000
1921 49.41250 49.27500 49.30417
...
1939      NA      NA      NA

$random
      Jan      Feb      Mar      Apr      May
Jun
1920      NA      NA      NA      NA      NA
NA
1921 3.972697368 -0.370942982 1.717434211 -0.838486842 0.029934211 -0.89
4846491
...
1939 -0.939802632 1.016557018 -0.545065789 0.782346491 -0.611732456 -0.43
6513158
      Jul      Aug      Sep      Oct      Nov
Dec
1920 -4.308881579 -4.209100877 -2.237609649 0.666118421 0.325986842 -0.03
9802632
1921 2.878618421 -1.688267544 -0.250109649 4.132785088 -2.957346491 2.85
6030702
...
1939      NA      NA      NA      NA      NA
NA

$figure
[1] -9.3393640 -9.8998904 -6.9466009 -2.7573465 3.4533991 8.9865132 12.9
672149 11.4591009
[9] 7.4001096 0.6547149 -6.6176535 -9.3601974

$type
[1] "additive"

attr(,"class")
[1] "decomposed.ts"

```

Ova funkcija daje originalne podatke neobrađenog skupa podataka *nottem* kao element **X**, zatim vrednosti za trend, sezonsku i slučajnu komponentu za svaku vremensku tačku, osim nekih nepoznatih vrednosti (N/A) na početku i na kraju koje su potrebne za izračunavanje celog modela.

Rezultati dekompozicije se mogu sačuvati u imenovani objekat i na taj način se pristupa elementima dekompozicije. Oni se dobiju tako što se unese znak dolar \$, a zatim ime elementa.

```

> decompnnottem<-decompose(nottem, type = "additive")
> decompnnottem$trend

```

```

$trend
      Jan      Feb      Mar      Apr      May      Jun      Jul      Au
g      Sep
1920 0 49.13750      NA      NA      NA      NA      NA 49.04167 49.1500
1921 49.56667 50.07083 50.32917 50.59583 50.61667 50.60833 50.45417 50.1291
7 49.85000
...
1939 49.67917 49.78333 49.89167 49.77500 49.55833 49.45000      NA      N
A      NA      Oct      Nov      Dec
1920 49.17917 49.19167 49.20000
1921 49.41250 49.27500 49.30417
...
1939      NA      NA      NA

```

Kako su izračunate vrednosti trenda

U poglavlju 3.1. detaljno je opisan metod pokretnih proseka. Osnovna metoda pokretnih proseka koristi neparan broj tačaka- opservacija i na osnovu njih ocenjuje trend u "srednjoj" tački. Ako se ocenjivanje vrši na osnovu parnog broja tačaka, tu ne postoji "srednja" tačka u kojoj se ocenjuje trend, pa se zato uzme prosek tačaka sa početka i kraja, te se za to onda sračuna prosek.

Ova metoda je takozvana metoda „centriranog“ pokretnog proseka. Ako su u pitanju kvartalni podaci, onda je raspon 4 (jer postoje četiri četvrtine godišnje). Ovako bi se računao centrirani pokretni prosek za vreme $t = 3$.

Izračunati prosek posmatranih vrednosti podataka u vremenima od 1 do 4:

$$\frac{1}{4}(x_1 + x_2 + x_3 + x_4)$$

Izračunati prosek vrednosti u vremenima od 2 do 5:

$$\frac{1}{4}(x_2 + x_3 + x_4 + x_5)$$

Zatim izračunati prosek ova dva proseka:

$$\frac{1}{2}\left(\frac{1}{4}(x_1 + x_2 + x_3 + x_4) + \frac{1}{4}(x_2 + x_3 + x_4 + x_5)\right) = \frac{1}{8}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_3 + \frac{1}{4}x_4 + \frac{1}{8}x_5$$

Generalno, centrirani pokretni prosek gladak za vreme t (sa 4 kvartala) je

$$\frac{1}{8}x_{t-2} + \frac{1}{4}x_{t-1} + \frac{1}{4}x_t + \frac{1}{4}x_{t+1} + \frac{1}{8}x_{t+2}$$

Za mesečne podatke „centrirani“ pokretni prosek „glatkiji“ za vreme t će biti:

$$\frac{1}{24}x_{t-6} + \left(\sum_{j=-5}^5 \frac{1}{12}x_{t+j}\right) + \frac{1}{24}x_{t+6}$$

Zaista, ako bi se posmatrale vrednosti elementa **X**, koje su originalni podaci i primenila ova formula za trend, dobile bi se vrednosti elementa **trend** koje su bile malopre prikazane.

Sezonske vrednosti

Sezonska komponenta je periodična. Vrednosti sezonskog efekta ponavljaju se svake godine (u svakom redu) u **\$seasonal** objektu iz numeričkog izlaza.

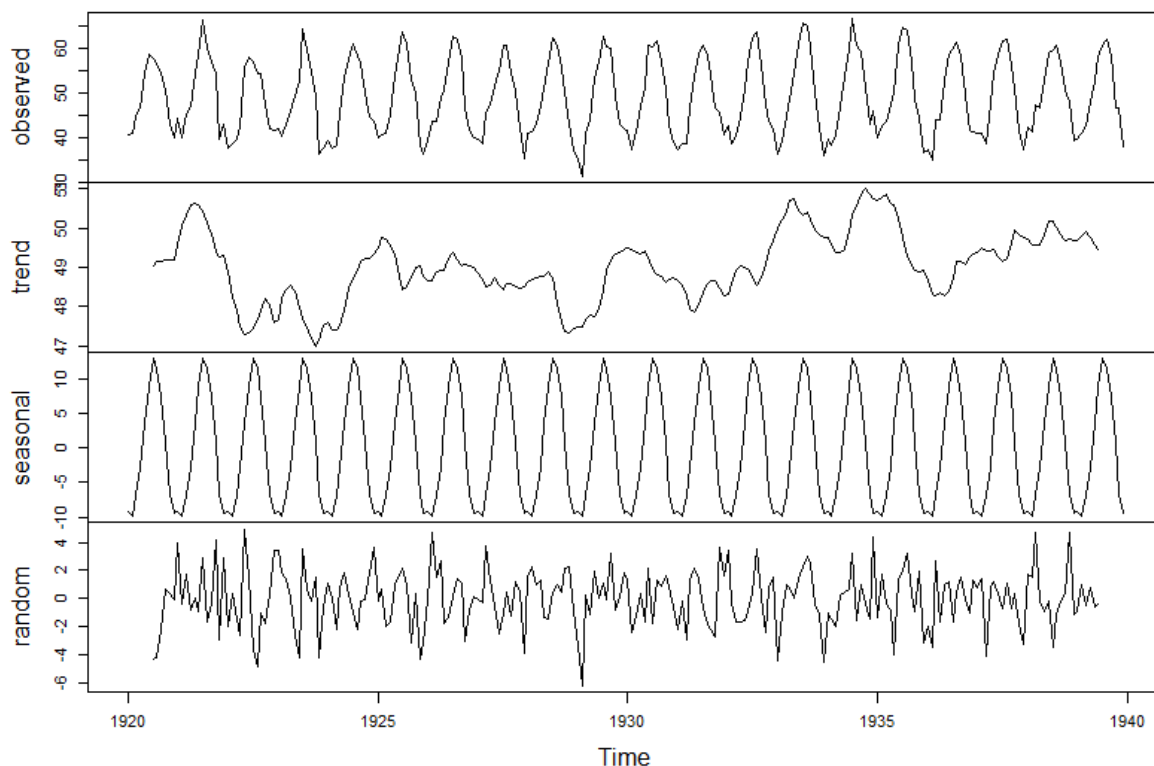
Elementi **\$figure** su efekti za četiri četvrtine.

Sezonske vrednosti koriste se za sezonsko prilagođavanje budućih vrednosti.

Ako se cela funkcija *decompose()* ubaci u funkciju **plot()** dobijaju se svi podaci u pogodnom grafičkom obliku koji prikazuje linijske grafike svake od komponenti, uključujući originalni skup podataka na vrhu i slučajnu komponentu na dnu.

```
> plot(decompose(nottem, type = "additive"))
```

Decomposition of additive time series



Slika 5.1.2. Grafički prikaz dekompozicije nottem skupa podataka korišćenjem aditivnog modela

Ovde se mogu videti trend i sezonska komponenta. Linija trenda ide „gore-dole“ tokom 20 godina, a neki vrhovi postoje oko 1921. i 1935. godine. Ali, generalno, srednja vrednost ostaje prilično stabilna i nema jasnog smera linije trenda. Zato se može zaključiti da ne postoji trend u skupu podataka.

Sezonski deo sa druge strane je sasvim jasno prepoznat i ostaje stabilan tokom čitavog vremenskog niza što je priroda ove vrste dekompozicije. Sezonska komponenta ostaje potpuno stabilna u čitavoj vremenskoj seriji.

Dakle, ovakav grafik je koristan kad god se radi sa podacima vremenskih serija koji su sezonski, a dobra stvar je u tome što *plot()* funkcija automatski detektuje klasu objekta i proizvodi tačan tip grafika.

Alternativni metod- funkcija *stl()*

R funkcija *stl()* vrši aditivnu dekompoziciju u kojoj se koristi izravnavanje sa manjim stepenom da bi se ocenio trend i (potencijalno) sezonski efekti. Postoji nekoliko parametara koji se mogu prilagoditi, ali su i default-ni parametri prilično dobri.

Sintaksa je sledeća:

stl(Ime vremenske serije, "periodic") ili ***stl(Ime vremenske serije, s.window= "periodic")***

Parametar „periodic“ u osnovi uzrokuje da se sezonski efekti ocenjuju na uobičajen način, kao prosečne vrednosti de-trendiranih vrednosti. Alternativa ovome je parametar *s.window* (sezonski prozor) koji se koristi da bi se izračunao sezonski deo. Može da se stavi da je vrednost ovog parametra jednaka nekom neparnom broju kašnjenja (najčešće 7 ili više). On koristi postupke izravnavanja sa manjim stepenom za ocenu sezonskih efekata na osnovu jednakosti broj godina= vrednost parametra *s.window*. Kada se to uradi, sezonski efekti će se menjati tokom pomeranja kroz vremensku seriju.

Zahvaljujući prirodni funkcije i argumentu *s.window*, izbegavaju se N/A vrednosti na početku i na kraju rezultata.

Prikazuju se rezultati prilikom pokretanja funkcije *stl()* za nottem podatke.

```
> stl(nottem, s.window = "periodic")
```

```
Call:
stl(x = nottem, s.window = "periodic")

Components
      seasonal      trend      remainder
Jan 1920 -9.3471980  49.68067  0.266525379
Feb 1920 -9.8552496  49.54552  1.109728805
Mar 1920 -6.8533008  49.41037  1.842931803
Apr 1920 -2.7634710  49.32862  0.134848770
May 1920  3.5013569  49.24688  1.351767558
Jun 1920  8.9833032  49.21027  0.306425938
Jul 1920 12.8452501  49.17367 -4.318916345
Aug 1920 11.4763813  49.13389 -4.210271506
Sep 1920  7.4475114  49.09411 -2.241625601
Oct 1920  0.4736899  49.15198  0.874331161
Nov 1920 -6.4301309  49.20984  0.120287167
Dec 1920 -9.4781423  49.49282 -0.214674402
Jan 1921 -9.3471980  49.77579  3.771408356
Feb 1921 -9.8552496  50.08132 -0.426073148
Mar 1921 -6.8533008  50.38686  1.566444922
Apr 1921 -2.7634710  50.49363 -0.730156884
May 1921  3.5013569  50.60040 -0.001756869
Jun 1921  8.9833032  50.50241 -0.785710552
Jul 1921 12.8452501  50.40441  3.050335100
Aug 1921 11.4763813  50.15065 -1.727027471
Sep 1921  7.4475114  49.89688 -0.344388977
Oct 1921  0.4736899  49.62951  4.096796830
Nov 1921 -6.4301309  49.36215 -3.232018119
Dec 1921 -9.4781423  49.04873  3.229417050
...
Jan 1939 -9.3471980  49.81506 -1.067860710
Feb 1939 -9.8552496  49.78783  0.967421826
Mar 1939 -6.8533008  49.76060 -0.507296065
Apr 1939 -2.7634710  49.68448  0.878994770
May 1939  3.5013569  49.60836 -0.709712574
Jun 1939  8.9833032  49.51780 -0.501104910
Jul 1939 12.8452501  49.42725 -1.572497911
Aug 1939 11.4763813  49.33655  0.987068518
Sep 1939  7.4475114  49.24585  1.506636011
Oct 1939  0.4736899  49.16331 -2.936997128
Nov 1939 -6.4301309  49.08076  3.949368976
Dec 1939 -9.4781423  49.00510 -1.726954804
```

I kao što se može videti vrednosti svih opservacija su ovde.

To je velika prednost u odnosu na standardni rad funkcije *decompose()* gde se zbog izvršenog računa dobiju N/A vrednosti na početku ili na kraju serije.

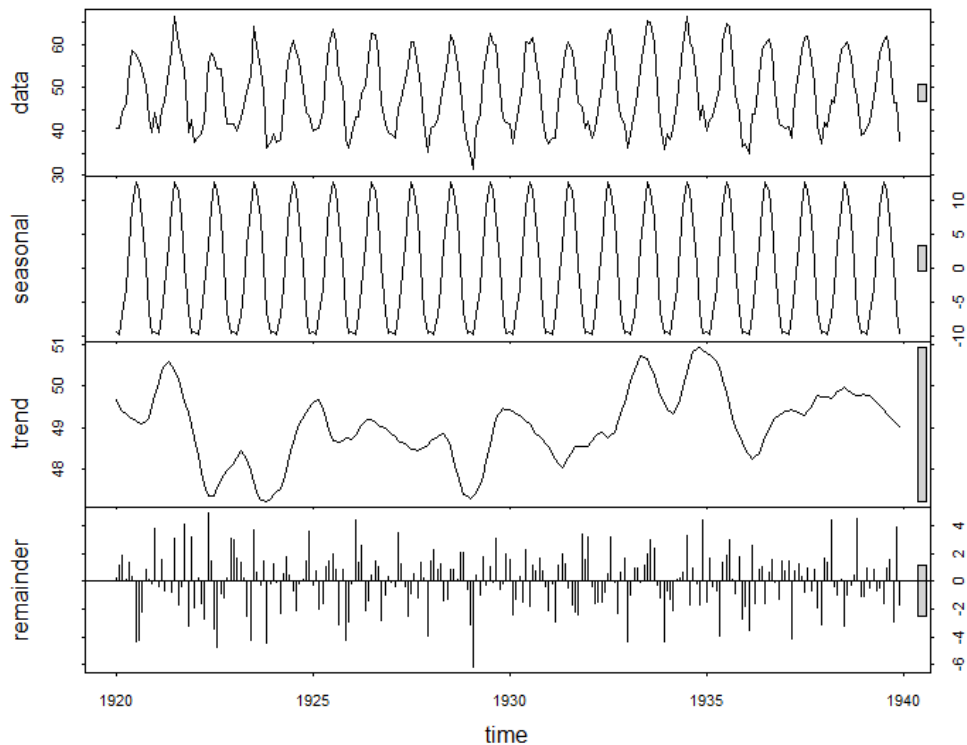
Aditivni sezonski efekti su -9.3471980, -9.8552496, -6.8533008, -2.7634710, 3.5013569, 8.9833032, 12.8452501, 11.4763813, 7.4475114, 0.4736899, -6.4301309, -9.4781423.

Oni se ne razlikuju mnogo od onoga što se dobilo aditivnom dokompozicijom. Te sezonske vrednosti su bile:

-9.3393640 , -9.8998904, -6.9466009, -2.7573465, 3.4533991, 8.9865132, 12.9672149, 11.4591009, 7.4001096, 0.6547149 , -6.6176535, -9.3601974

Ova funkcija se može prikazati grafički.

```
> plot(stl(nottem, s.window = "periodic"))
```



Slika 5.1.3. Grafički prikaz aditivne dekompozicije *nottem* podataka korišćenjem funkcije *stl()*

Nakon primene metoda dekompozicije sezoneke vremenske serije, rezultat se može koristiti na različite načine.

Sezonsko prilagođavanje

Iz *nottem* skupa podataka može se eliminisati sezonski efekat. Potrebno je oduzeti sezonsku komponentu od skupa podataka.

Prethodno je kreiran objekat *decompnottem* na sledeći način:

```
decompnottem<-decompose(nottem, type = "additive").
```

Elementima dekompozicije ovog objekta se može pristupiti koristeći znak dolar \$ i ime elementa. Ovde će biti potreban sezonski element.

Može se i proveriti klasa objekta *decompnottem* koristeći funkciju **class()**.

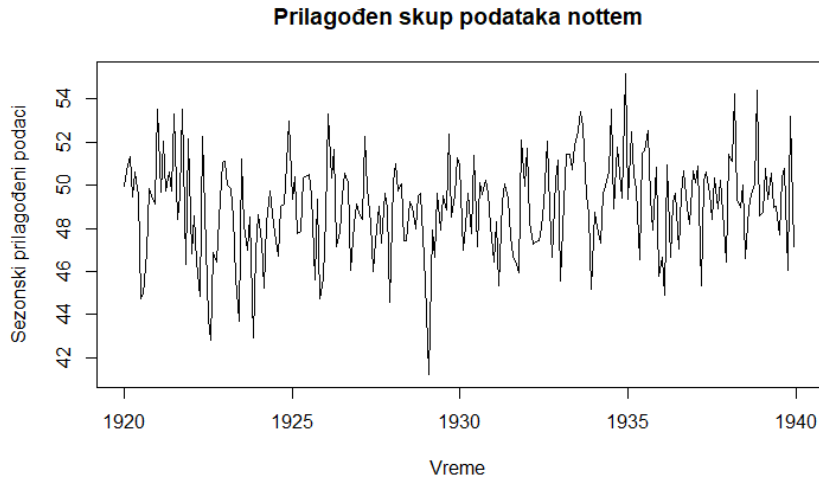
```
> decompnottem<-decompose(nottem, type = "additive")
> class(decompnottem)
[1] "decomposed.ts"
```

Nakon toga može se oduzeti sezonski element skupa podataka da bi se postavio sezonski prilagođeni element. Prilagođeni skup podataka naziva se ovde *nottemadjusted* što predstavlja *nottem* skup podataka minus sezonski deo prethodno kreiranog objekta *decompnottem*.

```
> nottemadjusted<-nottem- decompnottem$seasonal
```

Sada se može grafički prikazati objekat *nottemadjusted*.


```
> > plot(nottemadjusted, xlab="Vreme", ylab="Sezonski prilagođeni podaci",
+       main="Prilagođen skup podataka nottem")
```

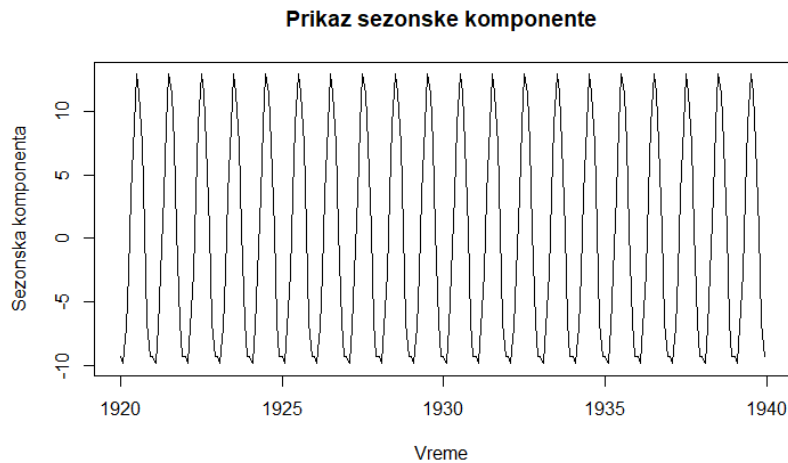


Slika 5.1.4. Grafički prikaz prilagođenog skupa podataka nottem

Na osnovu ovog prikaza bi se moglo zaključiti da su u seriji ostale samo slučajne varijacije. To je upravo ono što bi se očekivalo od ovakvog skupa podataka. Ne postoji trend, pa ovaj sezonski deo daje originalnu strukturu skupa podataka.

Zbog toga se radi poređenja grafički predstavlja sezonska komponenta.

```
> plot(decomptottem$seasonal, xlab="Vreme", ylab="Sezonska komponenta",
+       main="Prikaz sezonske komponente")
```



Slika 5.1.5. Grafički prikaz sezonske komponente skupa podataka nottem

I kao što se moglo očekivati, postoji jasan obrazac u sezonskom delu. Usponi i padovi su jasno strukturirani tokom godine.

Određivanje prognoze

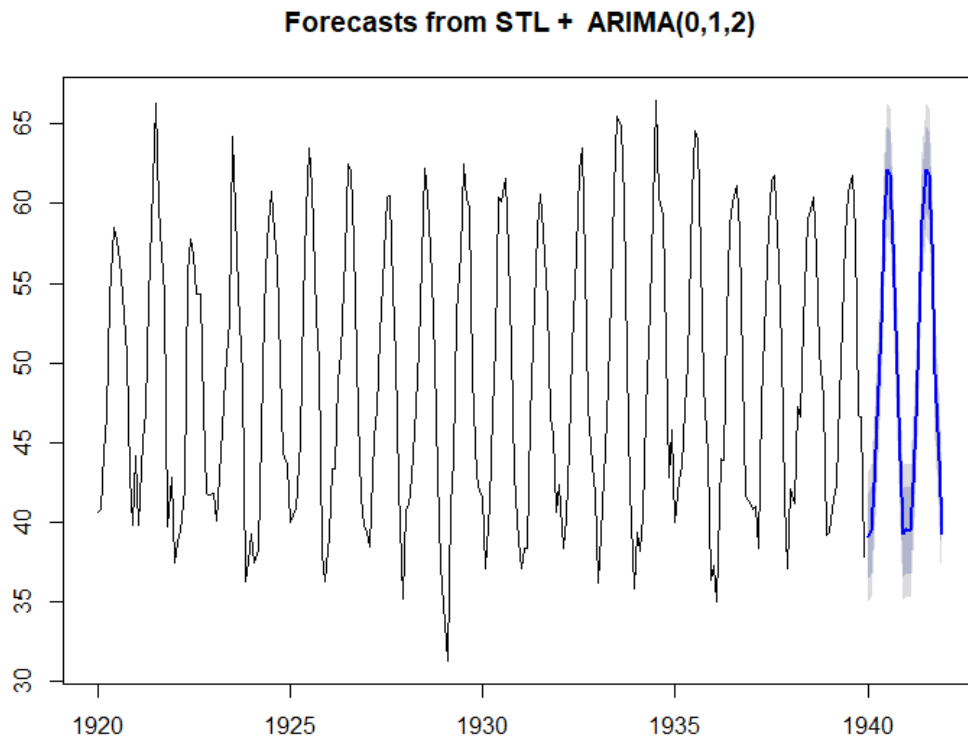
Dekomponovana vremenska serija se može koristiti da bi se kreirala prognoza.

Jedan od načina je da se koristi *stl* dekompozicija u kombinaciji sa **stlf()** funkcijom iz paketa *Forecast*.

Ova funkcija je dodata funkciji za *stl* koja nudi i prognoziranje.

Direktno se predstavlja prognoza pomoću finkcije `plot()`. U okviru funkcije `stlf()` može se postaviti metod za prognoziranje. Može se odabrati *ets*, *naive* ili *rw drift*. U ovom primeru koristi se metod ARIMA.

```
> install.packages("forecast")
> library(forecast)
> plot(stlf(nottem, method="arima"))
```



Slika 5.1.6. Grafički prikaz prognoziranja *nottem* podataka pomoću metoda ARIMA

Korelacija unutar i među vremenskim serijama

Koncepti kovarijancije i korelacije veoma su važni u analizi vremenskih serija. Konkretno, može se ispitati struktura korelacije izvornih podataka ili slučajnih grešaka iz modela dekompozicije kako bi pomogla da se identifikuju odgovarajući modeli vremenskih serija.

Uzorački korelogram (ACF)

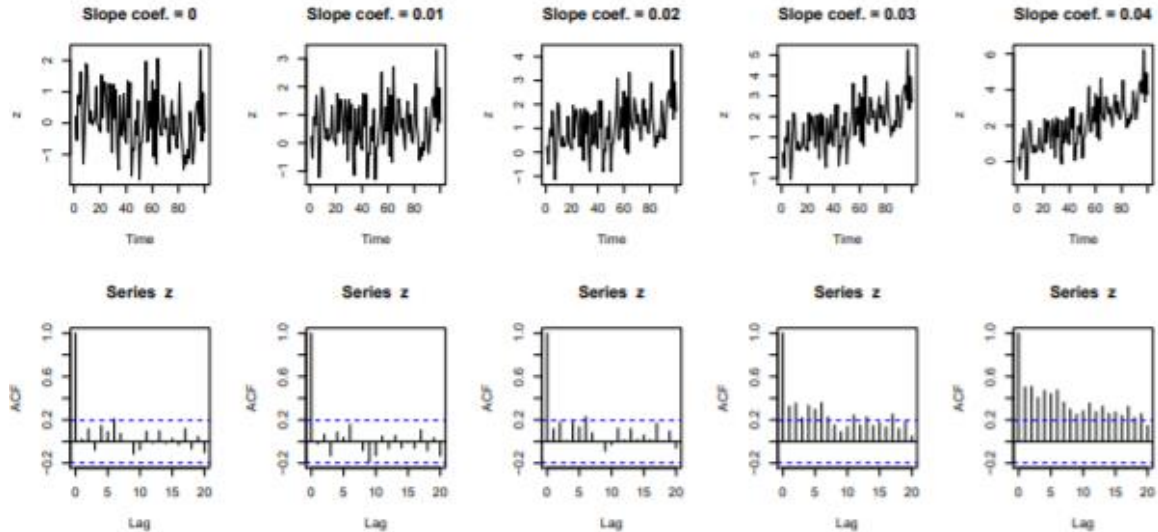
Uzorački korelogram služi za prepoznavanje modela stacionarnih vremenskih serija.

Pomoću uzoračkog korelograma može se proceniti i da li vremenska serija ima trend, sezonsku komponentu, ili obe komponente. U narednim koracima prikazaće se ACF za neke determinističke vremenske serije, koji će pomoći da se identifikuju trendovi i sezonski efekti u stohastičkim vremenskim serijama i objasne u modelima vremenskih serija. Prikazivanjem npr. ACF-a za sezonsku komponentu može se zaključiti da li je ona relativno stabilna tokom vremena, što nam pomaže u daljem odlučivanju koji bi model bio pogodniji.

U R-u se koristi funkcija **`acf()`** za predstavljanje grafičkog uzoračkog korelograma. Pozivanje same funkcije automatski će proizvesti korelogram tj. grafik autokorelacije u odnosu na vremensko kašnjenje. Argument *lag.max* omogućava da se podesi broj pozitivnih i negativnih kašnjenja. Argument

type predstavlja tip acf -a za računanje. Dozvoljene vrednosti su "correlation" (default-na vrednost), "covariance" ili "partial".

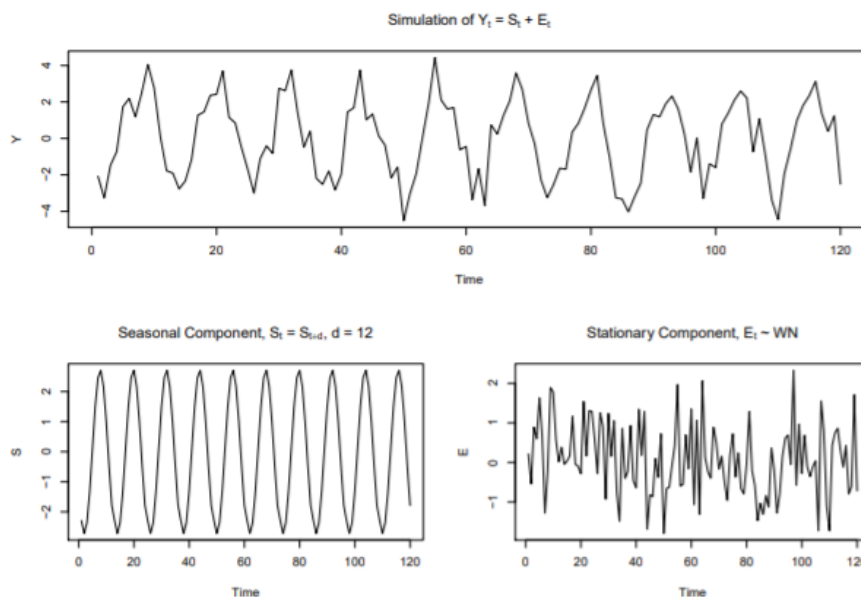
Neka je trend predstavljen jednačinom: $X_t = \alpha \cdot t + \varepsilon_t$. Na sledećoj slici prikazani su razni primeri grafika vremenskih serija pri čemu se trend povećava, odnosno povećava se koeficijent nagiba, α . Ovaj koeficijent pokazuje srednji apsolutni porast, odnosno prosečnu promenu vremenske serije (u zavisnosti od znaka, rast ili opadanje) u sukcesivnim vremenskim intervalima, u obuhvaćenom periodu. Takođe su prikazani i odgovarajući uzorački korelogrami.



Slika 5.1.7. Grafički prikaz trendova i njegovih uzoračkih korelograma, slope coef. predstavlja koeficijent nagiba (slika preuzeta iz literature [16])

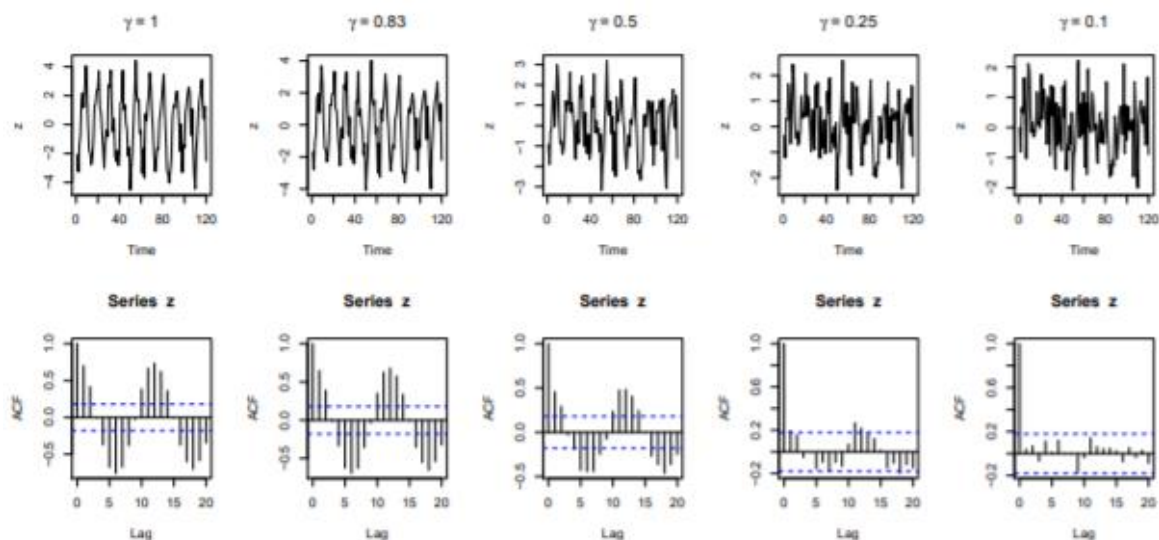
Ovo su nestacionarne vremenske serije sa konstantnom disperzijom i nekonstantnom sredinom. Što je trend izraženiji, ACF sporije opada.

Kao što je rečeno ACF se može koristiti i da bi se utvrdilo da li je u vremenskoj seriji prisutna sezonalnost. Na narednoj slici prikazana je vremenska serija sa sledećom formulom: $X_t = \gamma \cdot S_t + \varepsilon_t$



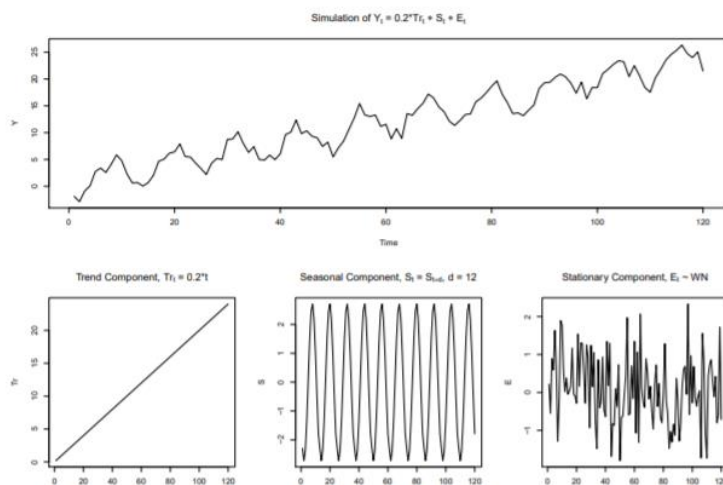
Slika 5.1.8. Grafički prikaz primera vremenske serije sa sezonskom komponentom (slika preuzeta iz literature [16])

Na sledećoj slici prikazani su neki primeri vremenski serija sa izraženijom sezonalnošću i njihov uzorački korelogram.

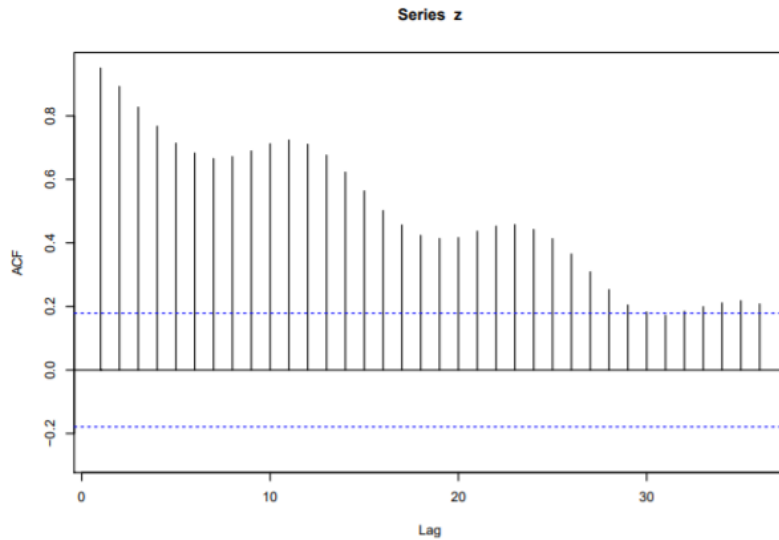


Slika 5.1.9. Grafički prikaz vremenskih serija sa izraženom sezonalnošću i njihov uzorački korelogram (slika preuzeta iz literature [16])

Sa slike se vidi da što je veća amplituda sezonskih fluktuacija, oscilacije su izraženije u ACF. Na sledećim slikama prikazana je vremenska serija koja ima i trend i sezonsku komponentu i njen uzorački korelogram.



Slika 5.1.10. Grafički prikaz vremenske serije sa trendom i sezonskom komponentom (slika preuzeta iz literature [16])

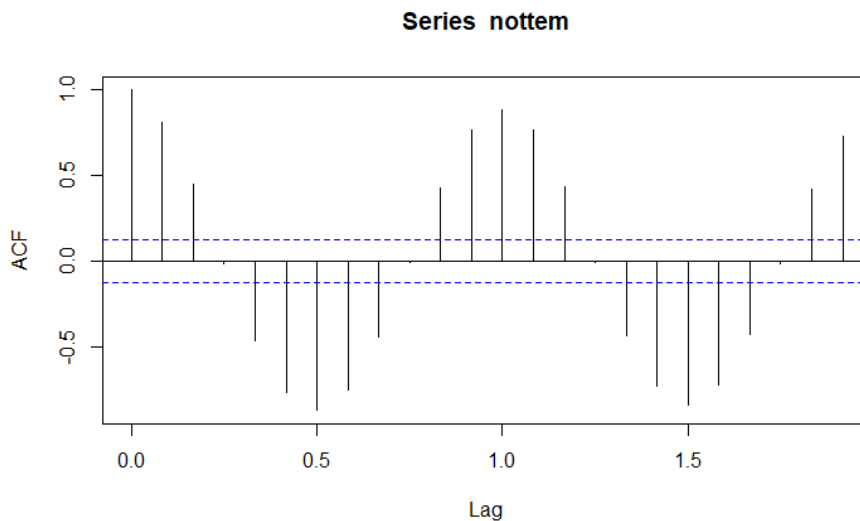


Slika 5.1.11. Grafički uzorački korelogram vremenske serije sa trendom i sezonskom komponentom (slika preuzeta iz literature [16])

ACF na ovoj slici pokazuje i spor pad i oscilacije.

Dakle, ACF je moćan alat u analizi vremenskih serija za identifikovanje važnih karakteristika podataka. Sada se može pogledati ACF za skup podataka *nottem*.

```
> par(mfrow = c(1, 1))
> acf(nottem)
```



Slika 5.1.12. Grafički uzorački korelogram skupa podataka *nottem*

Iz prikaza grafičkog uzoračkog korelograma može se zaključiti da skup podataka *nottem* ima sezonsku komponentu koja je stabilna tokom vremena i da ne postoji trend.

ACF se može koristiti i za ispitivanje da li je vremenska serija slučajna ili ne. Ovde se može ispitati ACF slučajne komponente i onda na osnovu toga zaključiti da li je izabran pogodan model.

Prilikom dekompozicije korišćenjem funkcije *decompose()* dobija se slučajna komponenta, međutim ona ima N/A vrednosti. S obzirom da za korišćenje funkcije *acf()* vrednosti ne smeju biti N/A, koristiće se slučajna komponenta dobijena alternativnim metodom *stl()*.

Vrednosti dobijene funkcijom *stl()* biće smeštene u objekat *nottemstl*.

```
> nottemstl<-stl(nottem, s.window = "periodic")
> nottemstl$time.series
```

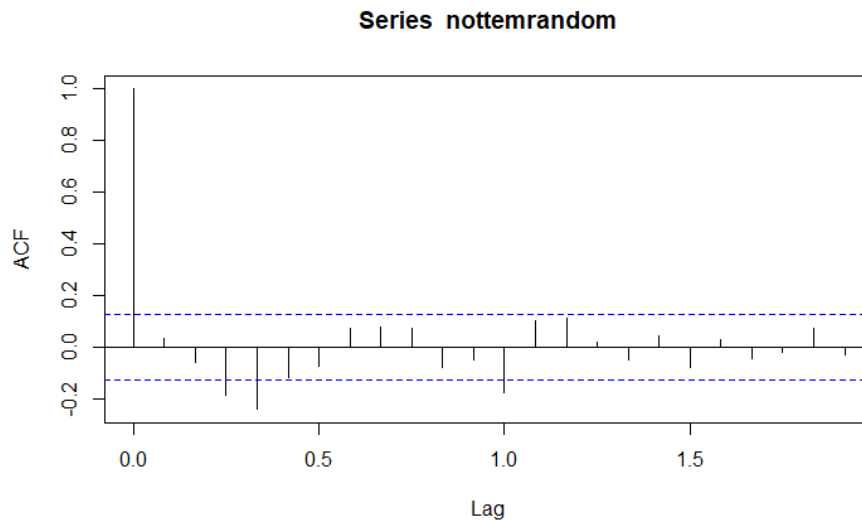
```
seasonal trend remainder
Jan 1920 -9.3471980 49.68067 0.266525379
Feb 1920 -9.8552496 49.54552 1.109728805
...
```

Da bi se pristupilo slučajnoj komponenti, odnosno objektu *remainder*, koristi se sledeća sintaksa:

```
> nottemrandom<-nottemstl$time.series[,"remainder"]
```

Zatim se prikazuje grafički uzorački korelogram.

```
> acf(nottemrandom)
```



Slika 5.1.13. Grafički uzorački korelogram slučajne komponente podataka *nottem*

Kašnjenja 3,4 i 12 su značajna (van intervala poverenja). Da bi vremenska serija bila slučajna očekuje se da nijedno kašnjenje nije značajno osim kašnjenja 0, što ovde nije zadovoljeno.

Ali slučajnost se može testirati i pomoću Boks- Pirsovog testa za ispitivanje nulte hipoteze o nezavisnosti u datoj vremenskoj seriji, kao što je opisano u poglavlju 2.7.

Sintaksa u R-u: **Box.test(x, lag = 1)**

gde je x vremenska serija.

U narednom koraku biće primenjena na slučajnu komponentu *nottemrandom*

```
> Box.test(nottemrandom, lag = 1)
```

```
Box-Pierce test
data: nottemrandom
X-squared = 0.28861, df = 1, p-value = 0.5911
```

Primenom ovog testa dobijena p-vrednost testa iznosi 0.5911, što je veće od nivoa značajnosti koji u podrazumevanom slučaju iznosi 0.05, pa se prihvata nulta hipoteza o slučajnosti.

5.2. PRIMER 2

Za ovaj primer koristi se **JohnsonJohnson** skup podataka koji se nalazi u R bazi podataka. On predstavlja tromesečnu zaradu (u dolarima) po deonicama kompanije Johnson & Johnson u periodu od 1960. do 1980. godine. Ovaj skup podataka je vremenska serija.

Sledi prikaz podataka, sumarne statistike, provera klase podataka i grafičko predstavljanje.

```
> JohnsonJohnson
```

	Qtr1	Qtr2	Qtr3	Qtr4
1960	0.71	0.63	0.85	0.44
1961	0.61	0.69	0.92	0.55
...				
1979	14.04	12.96	14.85	9.99
1980	16.20	14.67	16.02	11.61

```
> summary(JohnsonJohnson)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.440	1.248	3.510	4.800	7.133	16.200

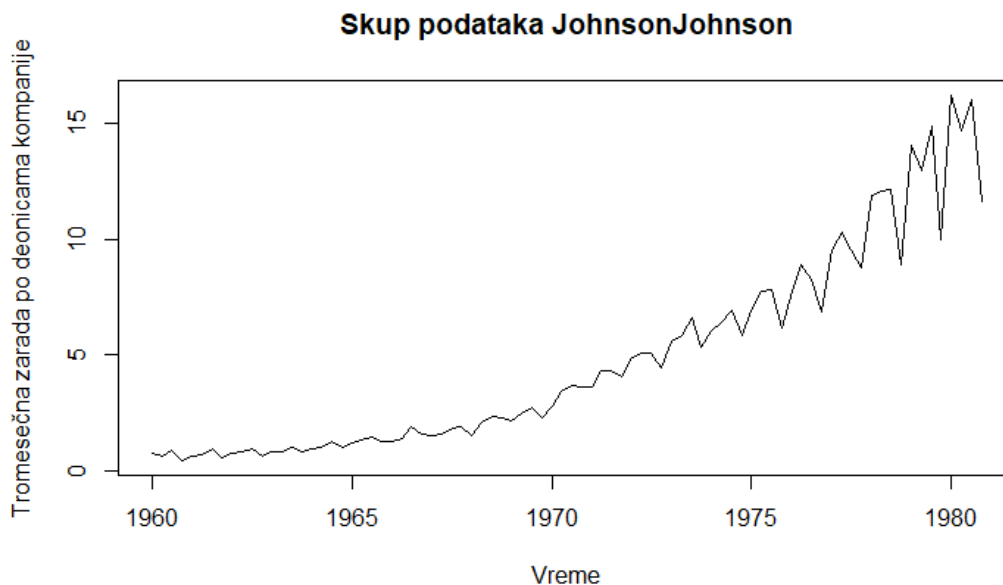
```
> sd(JohnsonJohnson)
```

```
[1] 4.309991
```

```
> class(JohnsonJohnson)
```

```
[1] "ts"
```

```
> plot(JohnsonJohnson, xlab="Vreme", ylab="Tromesečna zarada po deonicama kompanije",
+       main="Skup podataka JohnsonJohnson")
```



Slika 5.2.1. Grafički prikaz skupa podataka JohnsonJohnson

Zarada je eksponencijalno rasla tokom perioda uzorkovanja i imala je snažnu sezonalnost. Štaviše, promenljivost zarade se vremenom povećavala.

Sa grafika se takođe vidi i postojanje trenda. Dakle, može se koristiti multiplikativna dekompozicija, odnosno ovaj set podataka može se opisati **multiplikativnim modelom**.

Proverava se frekvencija i dužina skupa podataka.

```
> frequency(JohnsonJohnson)
[1] 4
> length(JohnsonJohnson)
[1] 84
```

Ciklični obrazac se ponavlja svake godine tako da periodičnost serije iznosi 4, za 4 kvartala (tromesečja), dok je dužina 84, što predstavlja 21 godinu (21x4).

Koristi se funkcija `decompose()` pri čemu se za tip upisuje **multiplicative**.

```
> Johnsondecomp<-decompose(JohnsonJohnson, type = "multiplicative")
> Johnsondecomp
```

```
$x
      Qtr1  Qtr2  Qtr3  Qtr4
1960  0.71  0.63  0.85  0.44
1961  0.61  0.69  0.92  0.55
...
1980 16.20 14.67 16.02 11.61

$seasonal
      Qtr1      Qtr2      Qtr3      Qtr4
1960 0.9930006 1.0329845 1.1140535 0.8599614
1961 0.9930006 1.0329845 1.1140535 0.8599614
...
1980 0.9930006 1.0329845 1.1140535 0.8599614

$trend
      Qtr1      Qtr2      Qtr3      Qtr4
1960      NA      NA  0.64500  0.64000
1961  0.65625  0.67875  0.70625  0.73000
...
1980 14.07375 14.42250      NA      NA

$random
      Qtr1      Qtr2      Qtr3      Qtr4
1960      NA      NA  1.1829139  0.7994545
1961  0.9360758  0.9841141  1.1692929  0.8761145
...
1980  1.1591928  0.9846815      NA      NA

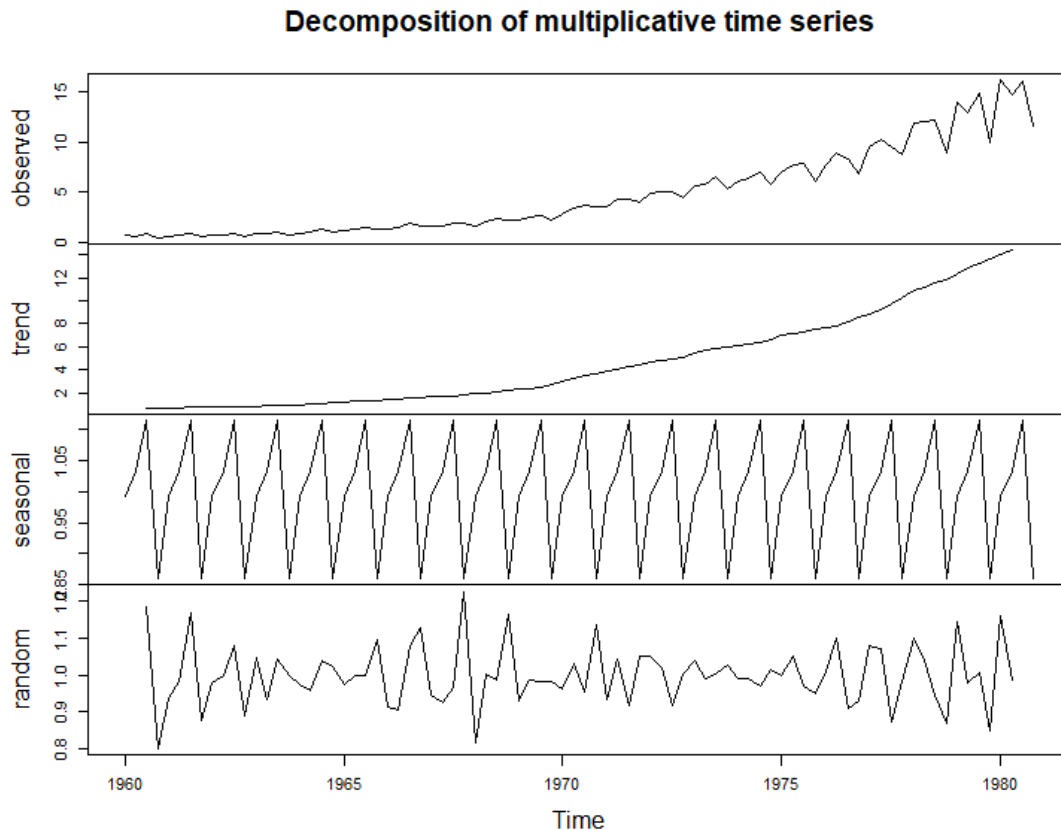
$figure
[1] 0.9930006 1.0329845 1.1140535 0.8599614

$type
[1] "multiplicative"

attr(,"class")
[1] "decomposed.ts"
```

Pomoću funkcije **plot()** se podaci prikazuju grafički. Dobiće se linijski grafici svake od komponenti, uključujući originalni skup podataka na vrhu i slučajnu komponentu na dnu, kao u prethodnom primeru.


```
> plot(Johnsondecomp)
```



Slika 5.2.2. Grafički prikaz multiplikativne dekompozicije podataka JohnsonJohnson

I ovde se vidi postojanje trenda i izražena sezonska komponenta.

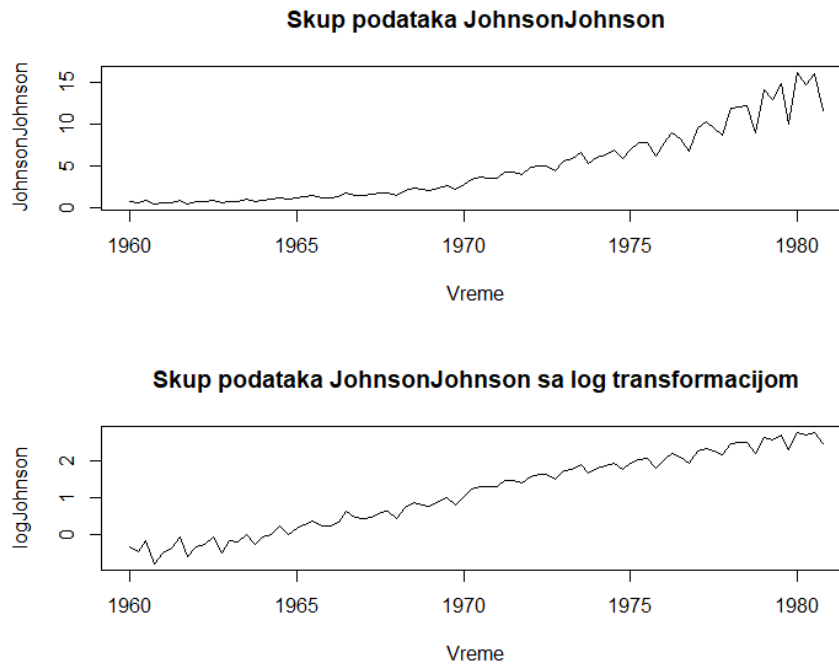
Može se uraditi logaritamska transformacija nad podacima. Ona se obično koristi u analizi finansijskih i ekonomskih vremenskih serija. U ovom slučaju, svi podaci su pozitivni, tako da nije potrebno prilagoditi podatke pre transformacije.

```
> logJohnson<-log(JohnsonJohnson)
> logJohnson
```

	Qtr1	Qtr2	Qtr3	Qtr4
1960	-0.34249031	-0.46203546	-0.16251893	-0.82098055
1961	-0.49429632	-0.37106368	-0.08338161	-0.59783700
...				
1979	2.64191040	2.56186769	2.69799987	2.30158459
1980	2.78501124	2.68580459	2.77383794	2.45186680

Sledi grafički prikaz podataka bez i sa logaritamskom transformacijom. Oba grafika su prikazana na istoj slici, jedan ispod drugog.

```
> par(mfrow=c(2,1))
> plot(JohnsonJohnson,xlab="Vreme", main="Skup podataka JohnsonJohnson")
> plot(logJohnson,xlab="Vreme", main="Skup podataka JohnsonJohnson sa log t
ransformacijom")
```



Slika 5.2.3. Grafički prikaz podataka JohnsonJohnson (slika gore) i podataka JohnsonJohnson sa logaritamskom transformacijom (slika dole)

Logaritamska transformacija se koristi kako bi se eliminisao eksponencijalni rast vremenske serije. I zaista, drugi grafik potvrđuje da je rast linearni na logaritamskoj skali. Drugi razlog je taj što se transformacija koristi da stabilizuje varijabilnost serije. Na prvom grafiku se vidi rastući obrazac varijabilnosti kojeg nema na drugom grafiku.

Korišćenje diferenciranja za uklanjanje trenda ili sezonskih efekata

Pored dekompozicije, za uklanjanje trenda ili sezonskih efekata iz vremenske serije može se koristiti diferenciranje. Diferenca prvog reda $\Delta: \Delta X_t = X_t - X_{t-1}$ je objašnjena u poglavlju 2.5. Operator diference reda k , Δ_k takođe je opisan u poglavlju 2.5.:

$$\Delta_k X_t = X_t - X_{t-k}$$

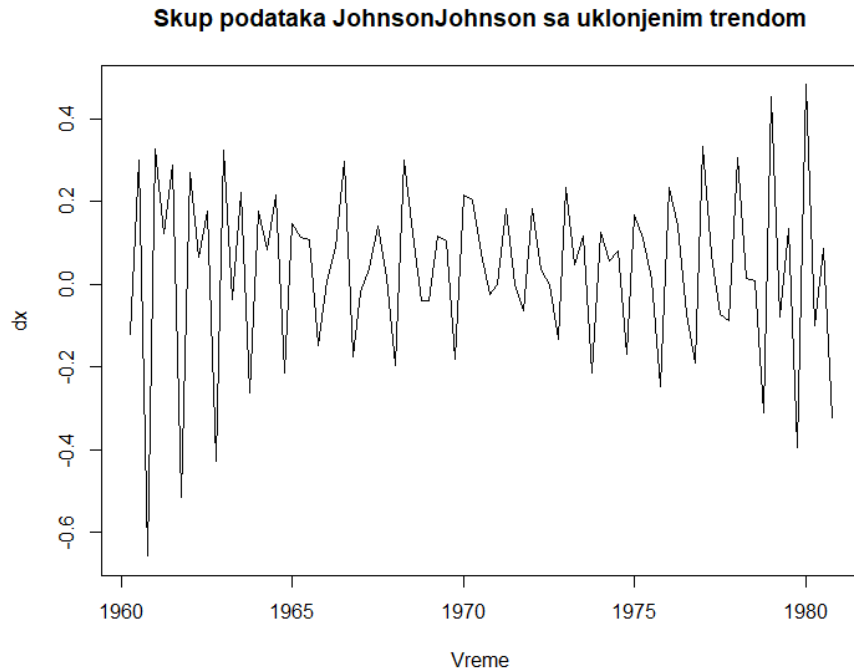
Ovde će se sa X_t označiti vremenska serija koja je logaritamski transformisana.

U R-u se koristi funkcija **diff()** za diferenciranje vremenske serije. Ona ima tri argumenta: x (podaci), lag (kašnjenje za koje se diferencira) i $differences$ (red diferenciranja, odnosno k u prethodnoj jednačini).

Na primer, diferenciranje prvog reda vremenske serije ukloniće linearni trend ($differences = 1$); dvostruko diferenciranje ukloniće kvadratni trend ($differences = 2$). Pored toga, diferenciranje prvog reda vremenske serije u kašnjenju jednakom periodu ukloniće sezonski trend (npr. postaviti $lag = 12$ za mesečne podatke).

Koristi se funkcija `diff()` za uklanjanje trenda iz vremenske serije X_t . Sa prethodnog grafika se videlo da je u pitanju linearni trend, pa se podešava parametar $differences = 1$.

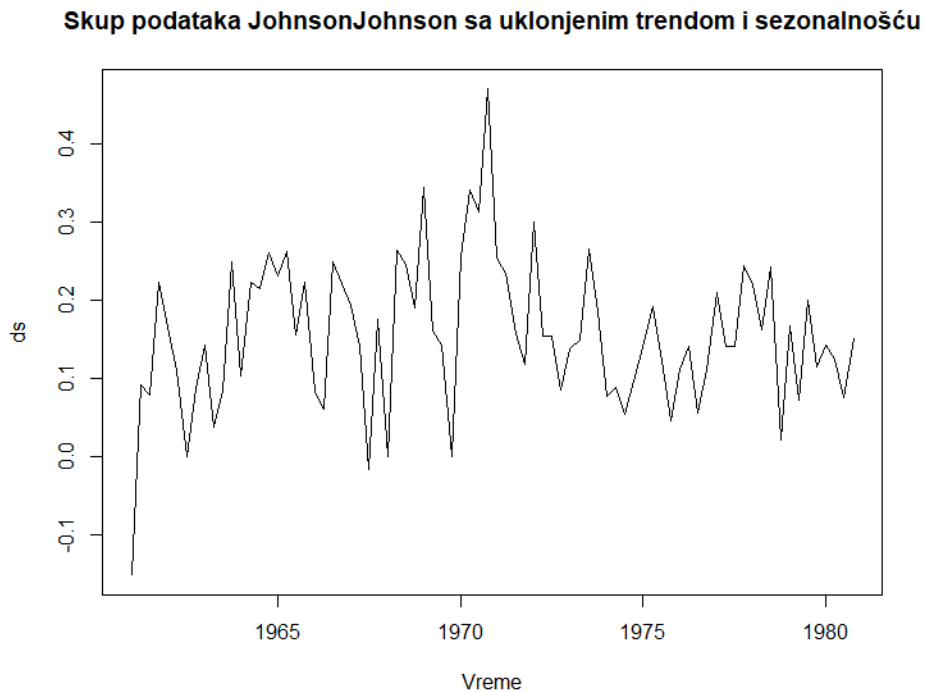
```
> par(mfrow=c(1,1))
> x<-logJohnson
> dx <- diff(x, differences = 1)
> plot(dx, xlab="Vreme",main="Skup podataka JohnsonJohnson sa uklonjenim tr
endom")
```



Slika 5.2.4. Grafički prikaz podataka JohnsonJohnson sa logaritamskom transformacijom sa uklonjenim trendom

Sa grafika se vidi da je uklonjen trend, ali i dalje postoje sezonski obrasci. Sledeći korak je diferenciranje vremenske serije za kašnjenje (lag) 4 jer su u pitanju kvartalni podaci.

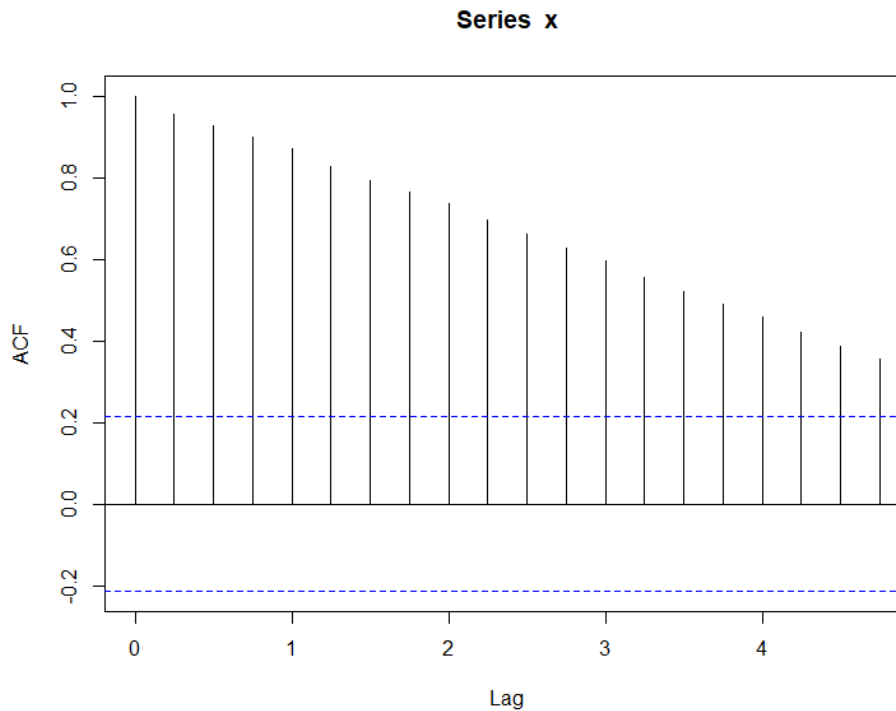
```
> ds <- diff(x, lag=4, differences = 1)
> plot(ds)
```



Slika 5.2.5. Grafički prikaz podataka JohnsonJohnson sa logaritamskom transformacijom sa uklonjenim trendom i sezonalnošću

Dobije se vremenska serija koja se čini slučajnim bez očiglednog trenda ili sezonskih komponenti. U nastavku je predstavljen grafički uzorak korelogram (ACF) vremenske serije X_t .

```
> acf(x)
```



Slika 5.2.6. Grafički uzorački korelogram (ACF) podataka JohnsonJohnson sa logaritamskom transformacijom

Korelogram je linearno opadajuća funkcija tokom vremena i trend je jako izražen, jer ACF sporo opada. ACF u kašnjenju 0 po default-u je jednak 1 (tj. korelacija vremenske serije sa sobom).

X-osa ima vrednosti od 0 do 4 za kašnjenja, što je uzrokovano time što R koristi indeks godine kao kašnjenje.

Vodoravne plave linije su približno 95% interval poverenja.

Sa grafika se vidi i da kvartalna zarada akcija ima snažne serijske korelacije.

Uobičajena metoda za rukovanje ovako snažnim serijskim korelacijama je razmatranje difference prvog reda ΔX_t . Primeniće se i operator sezonske difference na podatke ΔX_t :

$$\Delta_4(\Delta X_t) = \Delta X_t - \Delta X_{t-4}$$

Gde je

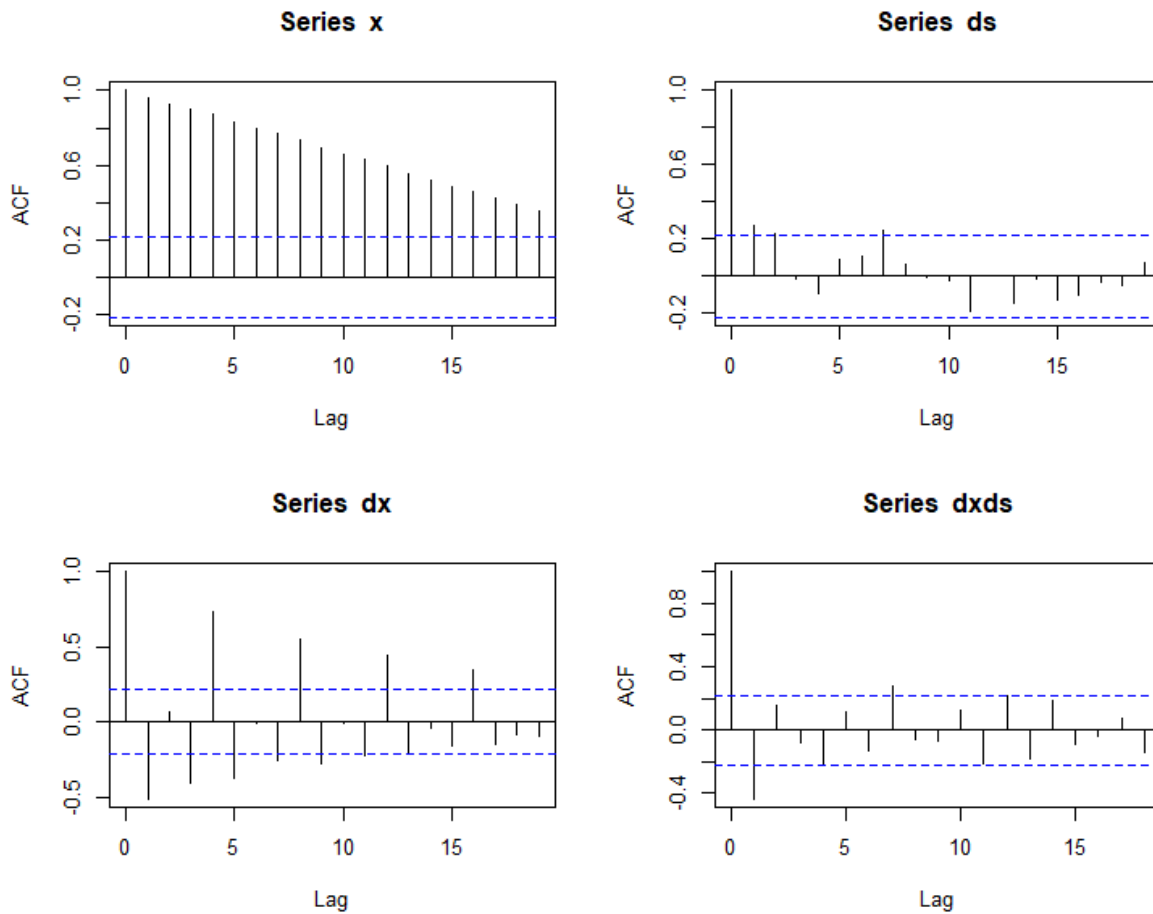
$$\Delta_4 X_t = X_t - X_{t-4}$$

operator difference reda 4.

Ideja je da se prikaže na jednom grafiku ACF od X_t , ΔX_t , $\Delta_4 X_t$ i $\Delta_4(\Delta X_t)$.

Kako bi se prikazala kašnjenja na x-osi bez indeksiranja potrebno je predstaviti podatke X_t kao vektor, zatim izračunati vrednost dx, ds i dxds i prikazati sve ACF na jednom grafiku (gde je $s = 4$).

```
> x<-c(logJohnson)
> dx <- diff(x, differences = 1)
> ds <- diff(x, lag=4, differences = 1)
> dxds <- diff(dx, lag=4, differences = 1)
> par(mfrow=c(2,2))
> acf(x)
> acf(ds)
> acf(dx)
> acf(dxds)
```



Slika 5.2.7. Grafički uzorački korelogram (ACF) podataka X_t , ΔX_t , $\Delta_4 X_t$ i $\Delta_4(\Delta X_t)$

Ako se pogleda grafikon u donjem levom uglu vidi se da je ACF jak kada je kašnjenje broj koji je deljiv brojem 4.

Na grafikonu u donjem desnom uglu prikazan je ACF od $\Delta_4(\Delta X_t)$ koji ima značajan negativni ACF u kašnjenju 1 i marginalnu negativnu korelaciju u kašnjenju 4.

Ponašanje uzorka ACF prikazano na graficima je uobičajeno među sezonskim vremenskim serijama. Ono dovodi do razvoja sledećeg posebnog sezonskog modela vremenskih serija:

$$(1 - L^s)(1 - L)X_t = (1 - \theta_1 L)(1 - \theta_1 L^s) e_t$$

Ovo je takozvani „Vazduhoplovni model“ koji je opisan u sekciji 4.3.2.4. Dakle, ovo jeste primer sezonskog ARIMA modela sa multiplikativnom sezonskom komponentom.

Njegova oznaka je $ARIMA(0,1,1) \times (0,1,1)_4$.

5.3. PRIMER 3

U ovom primeru koristi se set podataka *AirPassengers* koji postoji u R bazi. On predstavlja podatke o mesečnim brojevima putnika internacionalne aviokompanije u periodu od 1949. do 1960. godine i radi se o vremenskoj seriji. Sledi prikaz podatka, sumarne statistike i grafički prikaz.

```
> AirPassengers
```

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1949	112	118	132	129	121	135	148	148	136	119	104	118
1950	115	126	141	135	125	149	170	170	158	133	114	140
...												
1959	360	342	406	396	420	472	548	559	463	407	362	405
1960	417	391	419	461	472	535	622	606	508	461	390	432

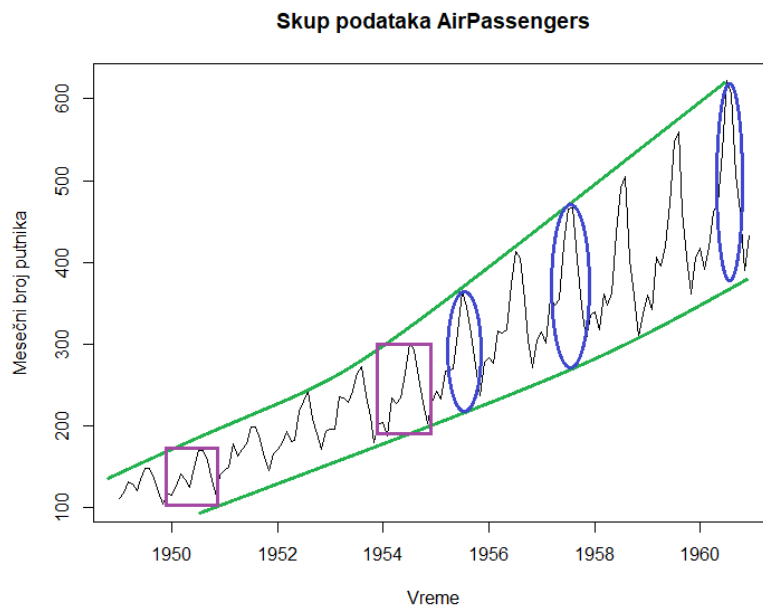
```
> summary(AirPassengers)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
104.0	180.0	265.5	280.3	360.5	622.0

```
> sd(AirPassengers)
```

```
[1] 119.9663
```

```
> par(mfrow=c(1,1))
> plot(AirPassengers, xlab="Vreme", ylab="Mesečni broj putnika",
+      main="Skup podataka AirPassengers")
```



Slika 5.3.1. Grafički prikaz skupa podataka *AirPassengers*

Ovaj skup podataka ima trend koji je skoro izvestan. Ako se pogledaju sezonski usponi i padovi, oni se neprestano povećavaju, svaki sezonski maksimum veći je nego godinu ranije, a slično je i sa najnižim nivoima, iako nije tako jasno vidljivo. Dakle, može se zaključiti da postoji trend i sezonski obrazac. Sledi provera frekvencije i dužine podataka.

```
> frequency(AirPassengers)
```

```
[1] 12
```

```
> length(AirPassengers)
```

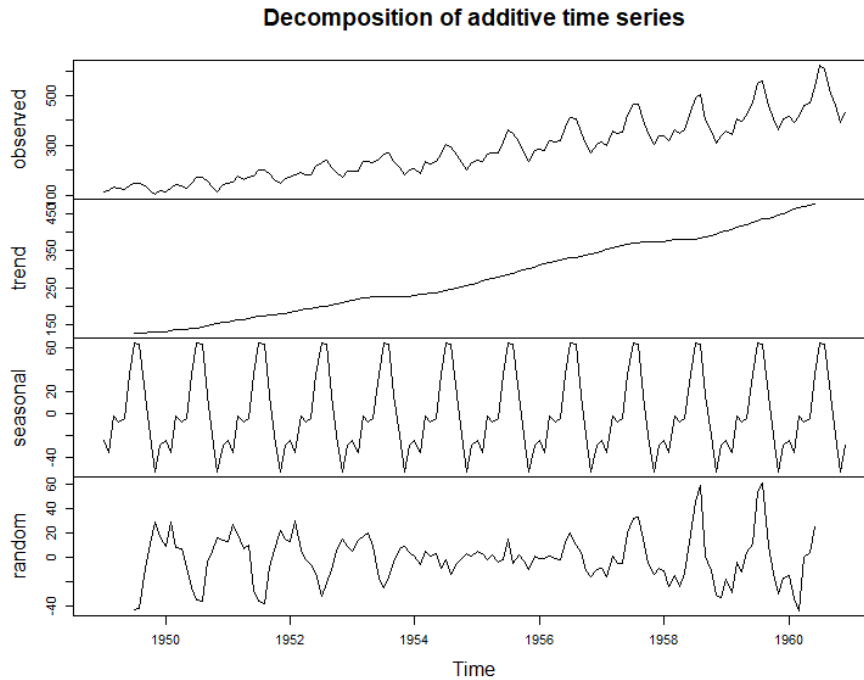
```
[1] 144
```

Postoji vremenski ciklus od 12 meseci u vremenskoj seriji, a obrazac na grafiku to takođe pokazuje. Sezonska amplituda se prilično povećava barem na poslednjem delu grafika.

Udaljenost između sezonskog uspona i pada postaje veća. Ovo bi mogao biti nagoveštaj da bi multiplikativni model mogao biti najpogodniji, ali pošto ovo nije konstantno u čitavoj vremenskoj seriji, pomoću funkcija R-a će se odlučiti koji je model najbolji.

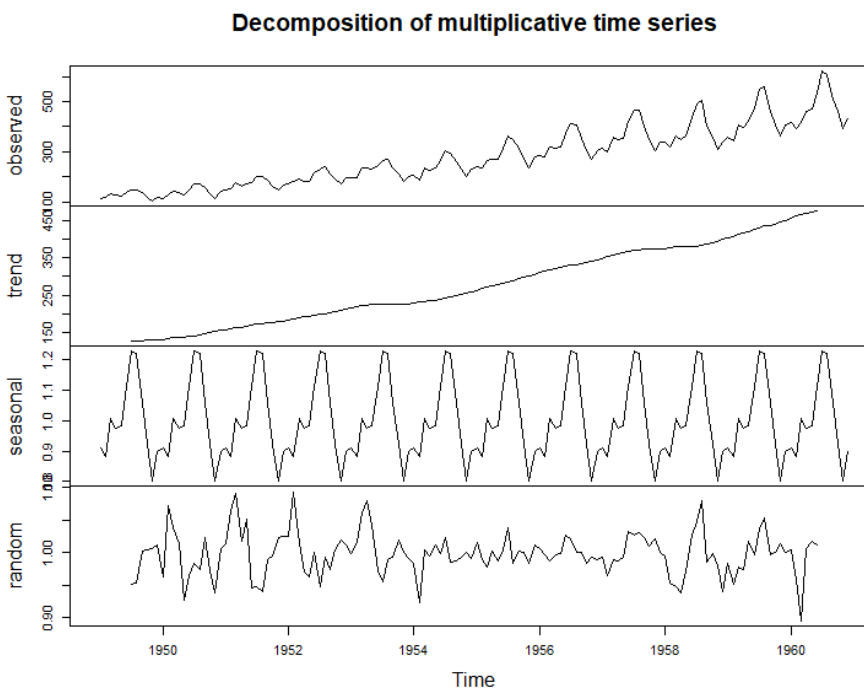
Primeniće se i aditivni i multiplikativni tip dekompozicije, a zatim će se kreirati grafici i uporediti rezultati.

```
> mymodel1<-decompose(AirPassengers, type = "additive")
> mymodel2<-decompose(AirPassengers, type = "multiplicative")
> plot(mymodel1)
```



Slika 5.3.2. Grafički prikaz aditivne dekompozicije skupa podataka AirPassengers

```
> plot(mymodel2)
```



Slika 5.3.3. Grafički prikaz multiplikativne dekompozicije skupa podataka AirPassengers

Aditivni model: Važna informacija koja se dobija je linija trenda. Očigledno se neprestano povećava tokom celog grafika i to je upravo onako kako se pretpostavljalo na osnovu originalnog skupa podataka. Takođe se mogu videti sezonski obrasci koji uvek dostižu vrhunac tokom letnjeg vremena.

Ono što je posebno zanimljivo je poslednja linija – slučajna komponenta. Izgleda kao da je ostao neki obrazac, i ovo se ne čini potpuno slučajnim. Ostao je obrazac koji nekako podseća na sezonski obrazac. Srednji deo grafika od oko 1954. do 1959. godine se čini slučajnim (nasumičnim), ali ostatak je u određenom intervalu izabrao najviše i najniže vrednosti. To nije dobar znak i pokazuje da bi se model mogao poboljšati.

Da bi model bio dobar, potrebno je da svi obrasci budu u modelu, samo bi slučajna komponenta ili beli šum trebalo da budu u ostatku (deo *random* na grafiku).

Multiplikativni model: Ovde su trend i sezonska komponenta ponovo prisutni. Dakle, slični su kao ranije, ali ne baš identični.

Međutim slučajni deo je značajno drugačiji nego kod aditivne dekompozicije. A ako je ovaj deo drugačiji, to se odnosi i na ostale dve komponente.

Pseudo slučajni interval je sada manje jasan. Cela linija izgleda „više slučajno“ iako i ovde postoji neki obrazac koji je prisutan barem na početku do oko 1953. godine.

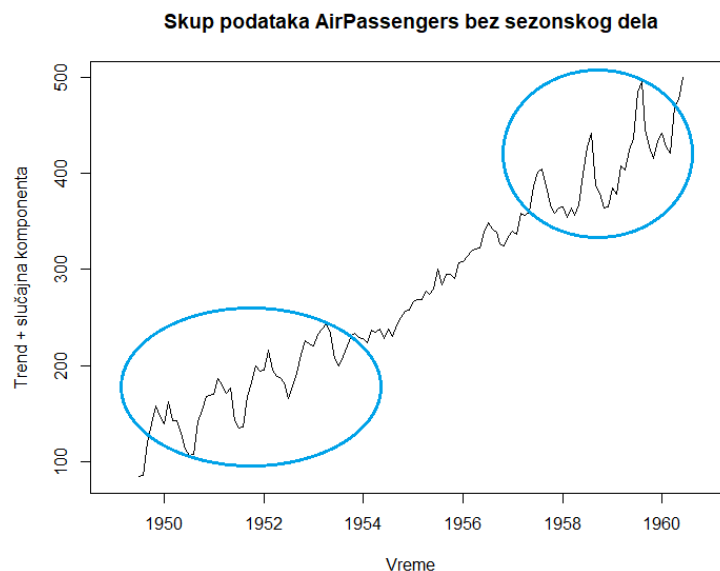
Ipak, model bi se mogao poboljšati i verovatno bi mu mogla odgovarati čak i druga metoda koja omogućava kombinaciju aditivnih i multiplikativnih komponenti, dakle neki mešoviti model.

Sezonske vrednosti

Da bi se dobio skup podataka bez sezonskog dela, potreban je sezonski prilagođen grafik aditivnog modela.

Postoji nekoliko načina kako da se ovo dobije: može se sumirati trend i slučajna komponentna modela ili oduzeti sezonska komponenta od prvobitnog skupa podataka, dobiće se isti rezultat.

```
> plot(mymodel1$trend + mymodel1$random, xlab="Vreme", ylab="Trend + slučajna komponenta",
+       main="Skup podataka AirPassengers bez sezonskog dela")
```



Slika 5.3.4. Grafički prikaz skup podataka AirPassengers bez sezonskog dela

Postoji trend kao što i treba da bude i zapravo je mnogo jasniji nego u originalnim podacima. Ali takođe se vidi da postoje pseudo sezonski pikovi na liniji trenda i oni su ravnomerno raspoređeni, što znači da postoji obrazac barem na početku i na kraju skupa podataka. Srednji deo oko 1956. izgleda dovoljno slučajno (nasumično).

Ovaj grafik pokazuje da će barem sezonskom delu modela biti potrebno poboljšanje. Zapravo, ovi podaci zahtevaju složeniji model koji ima mogućnost kombinovanja aditivnih i multiplikativnih karakteristika.

Može se proveriti da li bi odgovarao neki od ARIMA modela.

ARIMA model

U R paketu *forecast* nalazi se funkcija **auto.arima()** koja služi za pronalaženje i prikazivanje ARIMA modela koji najbolje odgovara jednodimenzionoj vremenskoj seriji. Default-ni argumenti su dizajnirani za brzu procenu modela za mnoge vremenske serije.

Funkcija **Arima()** u istom paketu služi za kreiranje ARIMA modela sa zadatim parametrima. Neki od argumenata funkcije su: **Arima(y, order = c(0, 0, 0), seasonal = c(0, 0, 0),...)** gde je:

y- vremenska serija klase *ts*,

order- specifikacija nesezonskog dela ARIMA modela, tj. komponente (p, d, q) opisane u poglavlju 4.2. i 4.3.

seasonal- specifikacija sezonskog dela ARIMA modela, plus period (koji podrazumevano ima frekvenciju $frequency(y)$). Ovo bi trebalo da bude lista sa komponentama *order* i *period*, pri čemu je *order* vektor sa komponentama (P, D, Q) koje su opisane u poglavlju 4.3. Periodu odgovara oznaka s iz istog poglavlja.

Funkcija `auto.arima()` se može primeniti na podatke *AirPassengers* i pogledati predloženi model.

```
> library(forecast)
> model<-auto.arima(AirPassengers)
> model
```

```
Series: AirPassengers
ARIMA(2,1,1)(0,1,0)[12]

Coefficients:
      ar1      ar2      ma1
      0.5960  0.2143 -0.9819
s.e.  0.0888  0.0880  0.0292

sigma^2 estimated as 132.3:  log likelihood=-504.92
AIC=1017.85  AICc=1018.17  BIC=1029.35
```

Predložen je najbolji model za ove podatke, a to je sezonski ARIMA model $ARIMA(2,1,1) \times (0,1,0)_{12}$. Mogu se pogledati sumarne karakteristike modela, a zatim i uzorački korelogram reziduala.

```
> summary(model)
```

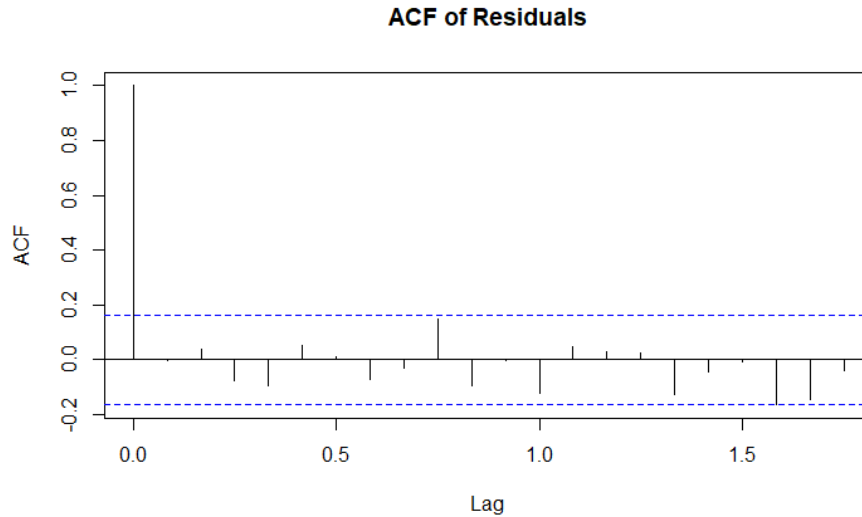
```
Series: AirPassengers
ARIMA(2,1,1)(0,1,0)[12]

Coefficients:
      ar1      ar2      ma1
      0.5960  0.2143 -0.9819
s.e.  0.0888  0.0880  0.0292

sigma^2 estimated as 132.3:  log likelihood=-504.92
AIC=1017.85  AICc=1018.17  BIC=1029.35

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 1.3423 10.84619  7.86754  0.420698  2.800458  0.245628 -0.00124847
```

```
> acf(resid(model), main = "ACF of Residuals")
```



Slika 5.3.5. Grafički uzorački korelogram reziduala modela

Sa grafika se vidi da nijedno kašnjenje nije značajno osim kašnjenja 0, tako da su ovde reziduali slučajni i može se zaključiti da je model dobar.

Može se i probati neki drugi ARIMA model, npr. $ARIMA(0,1,1) \times (0,1,0)_{12}$.

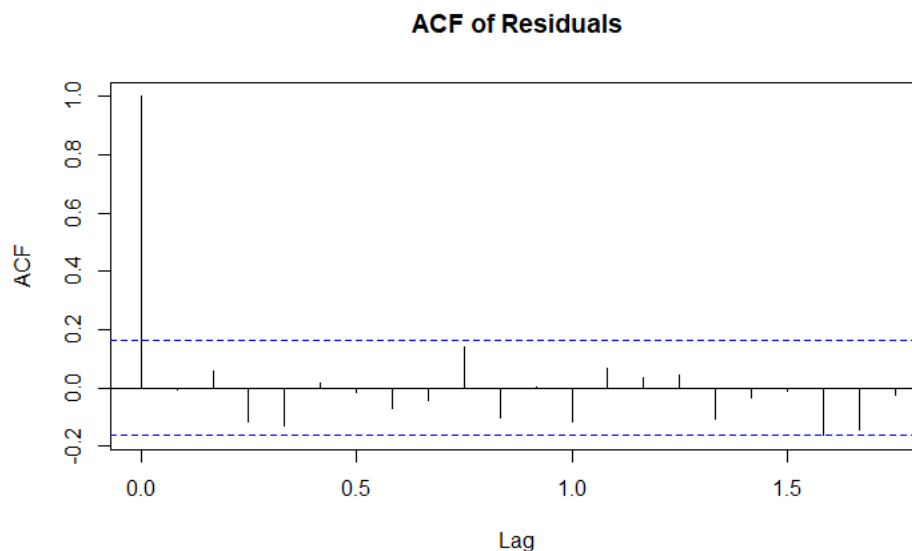
```
> model2<-Arima(AirPassengers, order = c(0,1,1), seasonal=list(order = c(0,
1,0),period=12))
> model2
```

```
Series: AirPassengers
ARIMA(0,1,1)(0,1,0)[12]
```

```
Coefficients:
      ma1
      -0.3184
s.e.    0.0877
```

```
sigma^2 estimated as 138.3: log likelihood=-508.32
AIC=1020.64 AICc=1020.73 BIC=1026.39
```

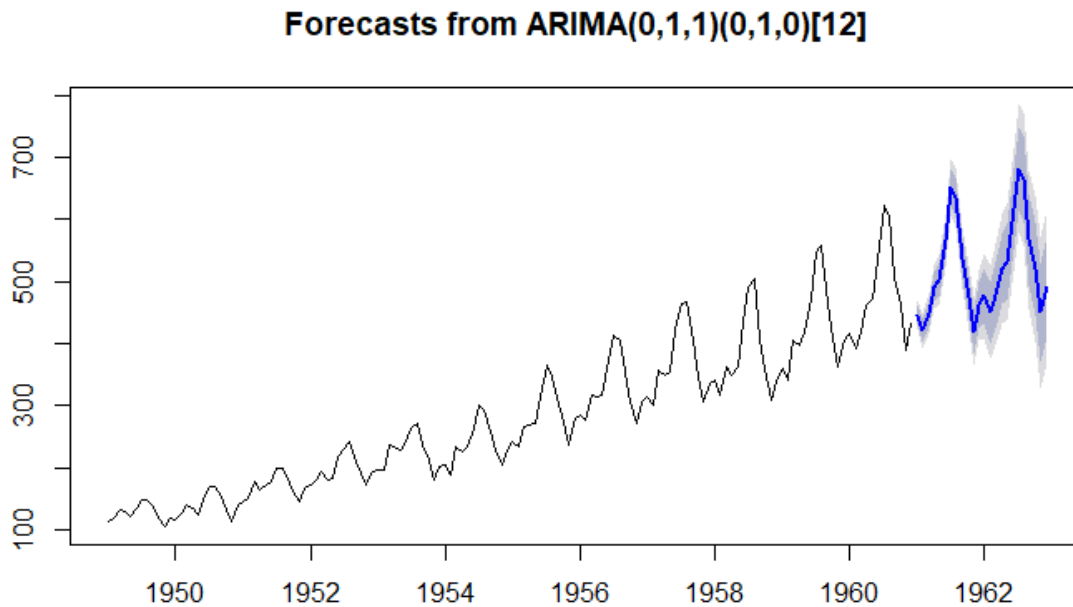
```
> acf(resid(model2), main = "ACF of Residuals")
```



Slika 5.3.6. Grafički uzorački korelogram reziduala modela2

I u ovom slučaju nijedno kašnjenje nije značajno osim kašnjenja 0, tako da su i ovde reziduali slučajni pa se takođe može zaključiti da je model dobar. Može se kreirati prognoza za dve godine korišćenjem funkcije `forecast()`.

```
> plot(forecast(model12,24))
```



Slika 5.3.7. Grafički prikaz prognoziranih podataka korišćenjem ARIMA modela

5.4. PRIMER 4

U ovom primeru koristi se skup podataka o maloprodaji putem e-trgovine, preuzet sa sajta: <https://www.kaggle.com/>. Prodaja putem e-trgovine podrazumeva prodaju robe i usluga gde kupac daje narudžbinu ili se cena i uslovi prodaje dogovaraju putem Interneta, mobilnih uređaja (M-commerce), ektraneta, elektronske mreže za razmenu podataka (EDI), elektronske pošte ili drugim uporedivim mrežnim sistemom. Plaćanje se može ili ne mora izvršiti putem Interneta.

U pitanju su kvartalni podaci, a vrednosti su izražene u milionima dolara. Podaci su preuzeti u vidu .csv fajla, pa ih je potrebno prvo učitati, a zatim uraditi određenu transformaciju. Koristi se R naredba *read.csv*.

```
> ecomnsa <- read.csv("C:/Users/Sonja.Jovanovic/Documents/matf/master/baze/
ECOMNSA.csv",
+                   header = TRUE)
> ecomnsa
```

	realtime_start	realtime_end	value	date
1	2019-10-26	2019-10-26	5241	1999-10-01
2	2019-10-26	2019-10-26	5553	2000-01-01
...				
78	2019-10-26	2019-10-26	129015	2019-01-01
79	2019-10-26	2019-10-26	139671	2019-04-01

Značajne vrednosti su u koloni **value**, tako da je potrebno kreirati odgovarajuću vremensku seriju gde je period kvartal. Takođe, mora se voditi računa da podaci počinju od poslednjeg kvartala 1999. godine.

```
> ecomn <- ecomnsa$value
> ecomn
```

[1]	5241	5553	6059	6892	9104	7923	7816	7737	10784	9621
10076	10760									
...										
[73]	145230	115602	122934	124214	160894	129015	139671			

```
> ecomntimeseries <- ts(ecomn, start=c(1999,4), frequency = 4)
> ecomntimeseries
```

	Qtr1	Qtr2	Qtr3	Qtr4
1999				5241
2000	5553	6059	6892	9104
2001	7923	7816	7737	10784
...				
2018	115602	122934	124214	160894
2019	129015	139671		

Prikazuju se sumarne statistike podataka, a zatim i grafički prikaz podataka.

```
> summary(ecomntimeseries)
```

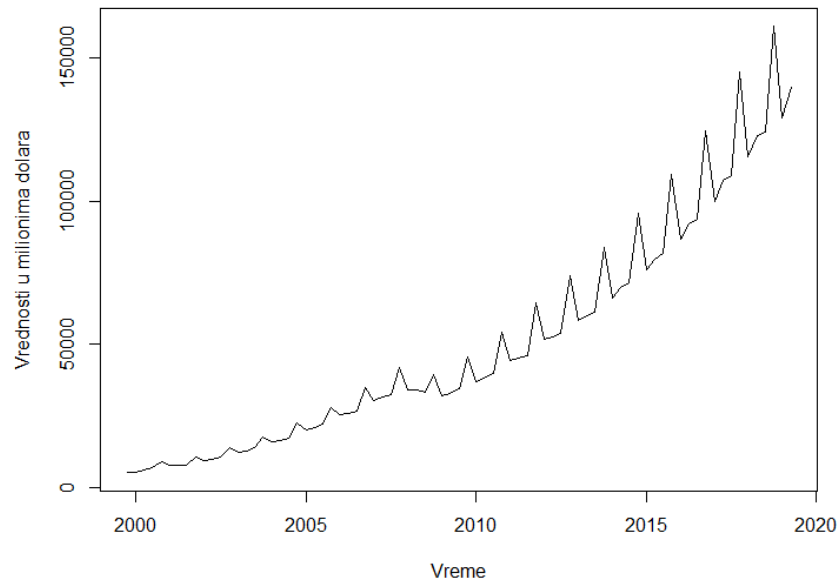
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
5241	19029	38467	51091	74873	160894

```
> sd(ecomntimeseries)
```

```
[1] 39529.31
```

```
> plot.ts(ecomntimeseries, xlab="Vreme", ylab="Vrednosti u milionima dolara"
+         main="Skup podataka o maloprodaji")
```

Skup podataka o maloprodaji



Slika 5.4.1. Grafički prikaz skup podataka o maloprodaji

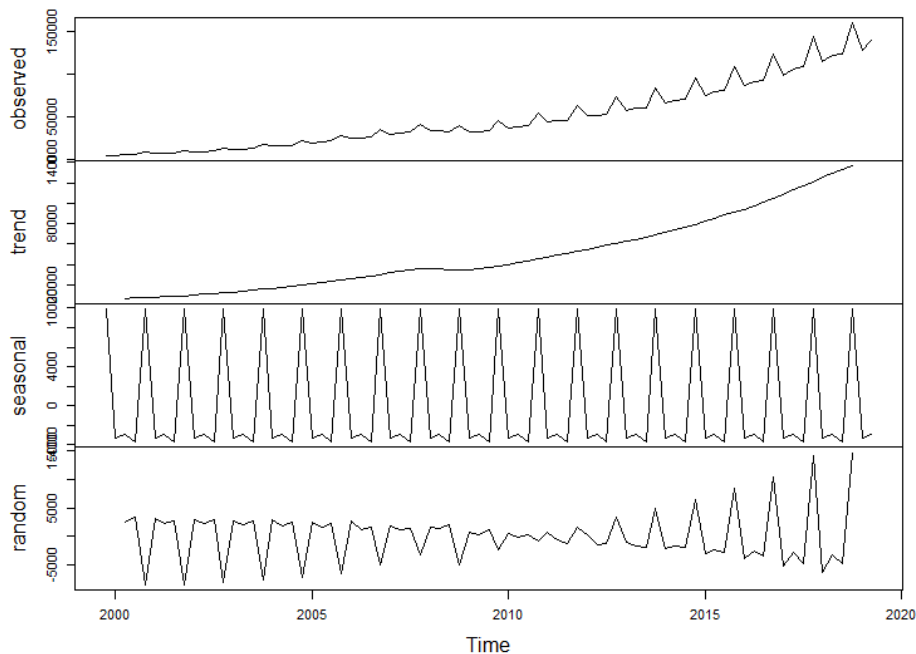
Sa grafika se vidi postojanje trenda i izražena sezonalnost. Prodaja se povećavala tokom vremena. Ako se pogledaju sezonski usponi i padovi, primećuje se da se oni povećavaju tokom vremena. Svaki sezonski maksimum je veći nego godinu dana ranije i slično za svaki sezonski minimum.

Deluje da se ovaj set podataka može opisati **multiplikativnim modelom**.

Ali, najbolje da se primene oba modela dekompozicije pa da se proceni koji je bolji.

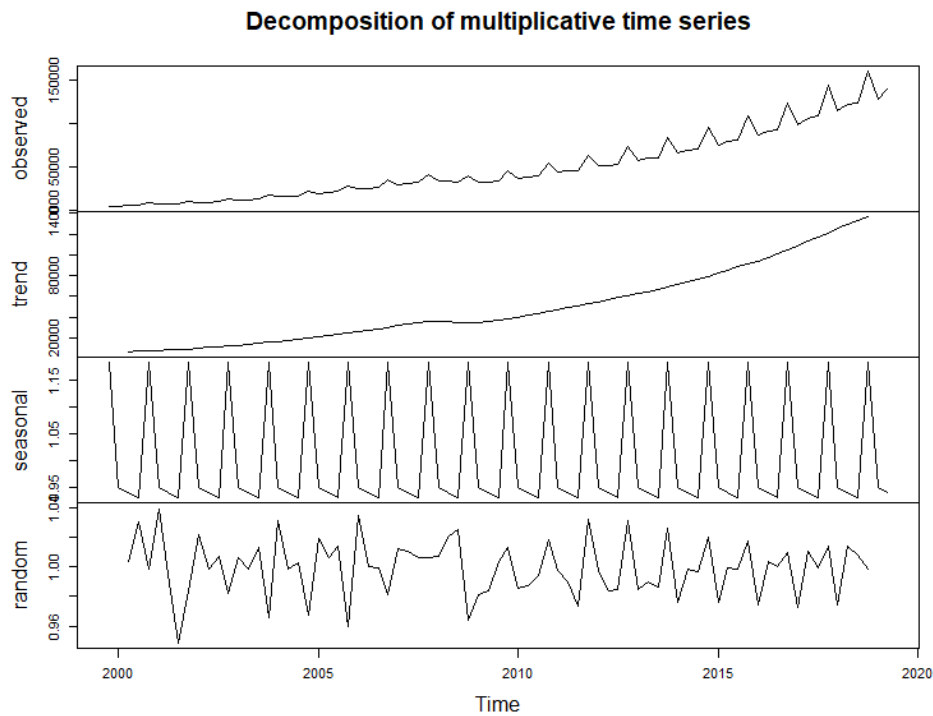
```
> mymodel1<-decompose(ecomntimeseries, type = "additive")
> mymodel2<-decompose(ecomntimeseries, type = "multiplicative")
> plot(mymodel1)
```

Decomposition of additive time series



Slika 5.4.2. Grafički prikaz aditivne dekompozicije skupa podataka o maloprodaji

```
> plot(mymode12)
```



Slika 5.4.3. Grafički prikaz multiplikativne dekompozicije skupa podataka o maloprodaji

Na prvom grafiku gde je prikazana dekompozicija aditivnog modela vidi se linija trenda koja se neprestano povećava tokom celog grafika. Takođe se mogu videti jasni sezonski obrasci.

Što se tiče slučajne komponente, izgleda kao da je ostao neki obrazac, pa se ona ne čini potpuno slučajnom. Obrazac podseća na sezonski obrazac.

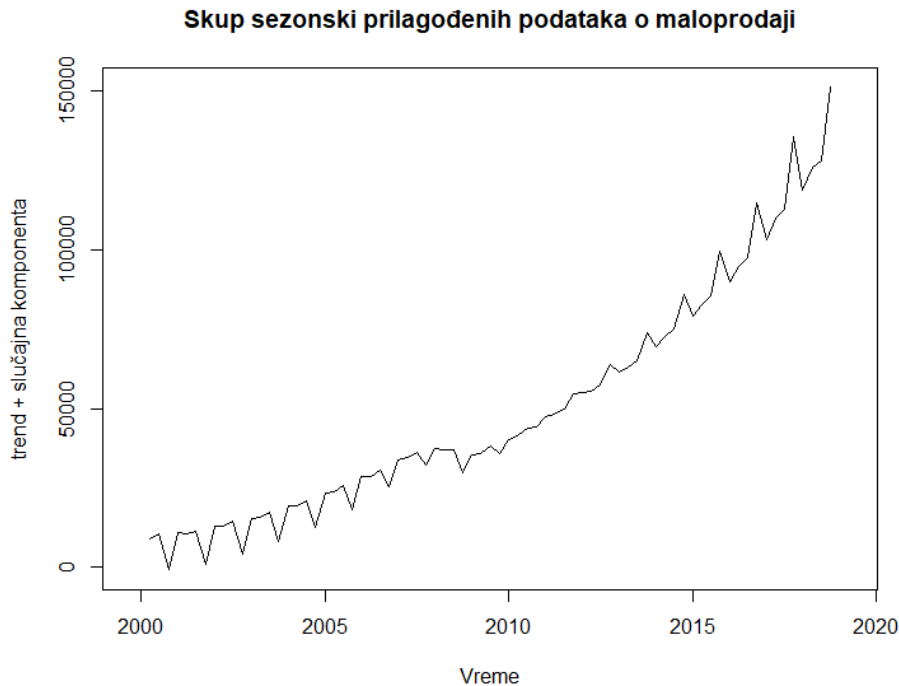
Samo se jedan deo grafika oko 2010. godine čini slučajnim. Ovo nije dobar znak kao u prethodnom primeru i pokazuje da bi se model mogao poboljšati.

Ako se pogleda grafik dekompozicije multiplikativnog modela vidi se da su isto prisutni trend i sezonska komponenta kao i na prethodnom grafiku. Međutim slučajni deo se razlikuje.

Slučajna komponenta sada deluje da nema neki jasan obrazac, osim možda za poslednje godine.

I u ovom slučaju prikazuje se sezonski prilagođen grafik aditivnog modela.

```
> plot(mymode11$trend + mymode11$random,xlab="vreme", ylab="trend + slučajna komponenta",
+      main="Škup sezonski prilagođenih podataka o maloprodaji")
```



Slika 5.4.4. Grafički prikaz sezonski prilagođenih podataka

Kao i u prethodnom primeru, i ovde se vidi jasan obrazac na početku i na kraju skupa podataka, dok srednji deo deluje slučajno, pa je zaključak da podaci zahtevaju složeniji model koji ima mogućnost kombinovanja aditivnih i multiplikativnih karakteristika, dakle neki mešoviti model.

I u ovom primeru se pomoću funkcije *auto.arima()* može pogledati predloženi model.

```
> library(forecast)
> mymodel3<-auto.arima(ecomntimeseries)
> mymodel3
```

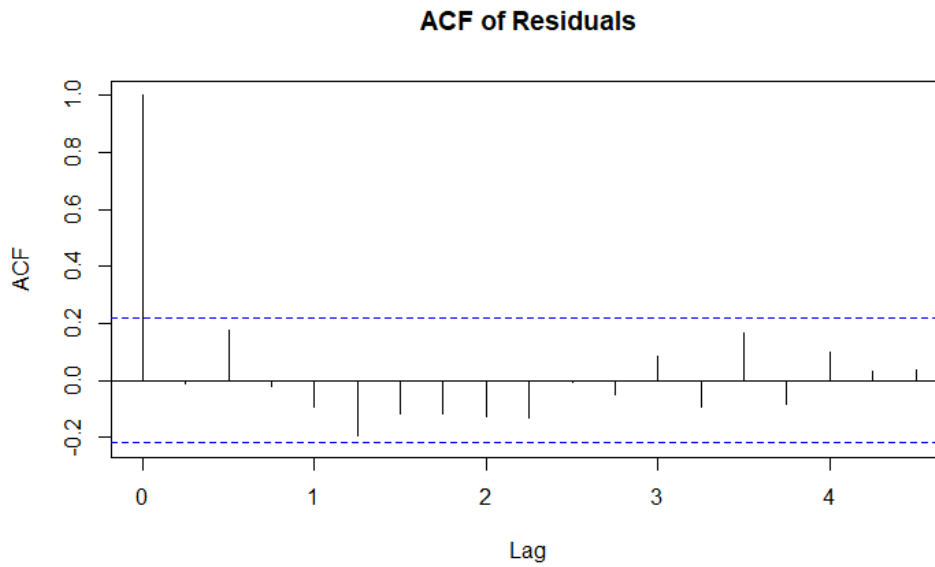
```
Series: ecomntimeseries
ARIMA(0,1,0)(2,1,0)[4]

Coefficients:
      sar1      sar2
      0.2645  0.4431
s.e.    0.1045  0.1146

sigma^2 estimated as 2288011:  log likelihood=-647.17
AIC=1300.34  AICc=1300.68  BIC=1307.25
```

Predložen je najbolji model za ove podatke, a to je sezonski ARIMA model $ARIMA(0,1,0) \times (2,1,0)_4$. Može se pogledati uzorački korelogram reziduala.

```
> acf(resid(mymodel3), main = "ACF of Residuals")
```



Slika 5.4.5. Grafički uzorački korelogram reziduala modela3

Sa grafika se vidi da nijedno kašnjenje nije značajno osim kašnjenja 0, tako da su ovde reziduali slučajni i može se zaključiti da je predloženi model dobar.

U ovom poglavlju korišćena je literatura [6], [7], [8], [10], [11], [12], [13], [16], [17], [18], [19], a takođe je korišćen kod i primeri iz kursa [9].

6. SIMULACIJA MODELA

Dosadašnji primeri u radu su se sastojali od analize konkretnih vremenskih serija i procene da li je bolji aditivni, multiplikativni, ili neki mešoviti model. Sada ćemo poći obrnutim putem i simulirati različite vrste modela i predstaviti ih grafički kako bi se videlo kako oni otprilike izgledaju. Time se postiže da se dobije ideja o tome kako bi određeni model izgledao i koje bi karakteristike imao, što omogućava da se prepozna za kakve podatke je taj model adekvatan.

6.1. SIMULACIJA ADITIVNIH MODELA

Za simulaciju aditivnog modela potrebna je prvo simulacija belog šuma.

Simulacija belog šuma u R-u

Funkcija **arima.sim()** može da se koristi za simulaciju podataka iz različitih modela vremenskih serija. Na osnovu modela koji se primenjuje, odrede se odgovarajuće vrednosti za p , d i q za model ARIMA (p , d , q). Opšti oblik funkcije `arima.sim()`:

arima.sim (model, n)

Za parametar *model* stavlja se AR, redosled diferencijacije i MA komponente, dok je n dužina izlaznih serija.

Za model **belog šuma**, svi p , d i q u modelu ARIMA su 0. Dakle, ARIMA (0,0,0) je model belog šuma. Da bi se generisalo 80 opservacija, postavlja se vrednost argumenta n na 80.

```
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
```

Ovim će se kreirati objekat vremenske serije koji predstavlja model belog šuma.

Može se izračunati srednja vrednost i standardna devijacija ove vremenske serije i primetiti da će ona imati srednju vrednost blizu 0, a standardnu devijaciju blizu 1.

```
> mean(e)
[1] -0.1201799
> var(e)
[1] 0.6717574
> sd(e)
[1] 0.8196081
```

Pomoću funkcije `arima.sim()` može se generisati vremenska serija slična belom šumu, ali sa eksplicitno zadatom srednjom vrednošću i standardnom devijacijom. U sledećem primeru simuliran je model vremenske serije sa srednjom vrednošću 1 i standardnom devijacijom 0.2.

```
> e2 <- arima.sim(model = list(order = c(0, 0, 0)), n = 80, mean=1, sd=0.2)
```

Drugi način simulacije belog šuma, tačnije Gausovog belog šuma je korišćenjem funkcije **rnorm()**. Ova funkcija generiše vektor normalno raspodeljenih slučajnih brojeva.

Opšti oblik funkcije je: **rnorm(n, mean = 0, sd = 1)**

gde n predstavlja broj opservacija, *mean* je vektor srednjih vrednosti, dok je *sd* vektor standardnih devijacija.

U narednom koraku simulira se Gausov šum za 80 opservacija.

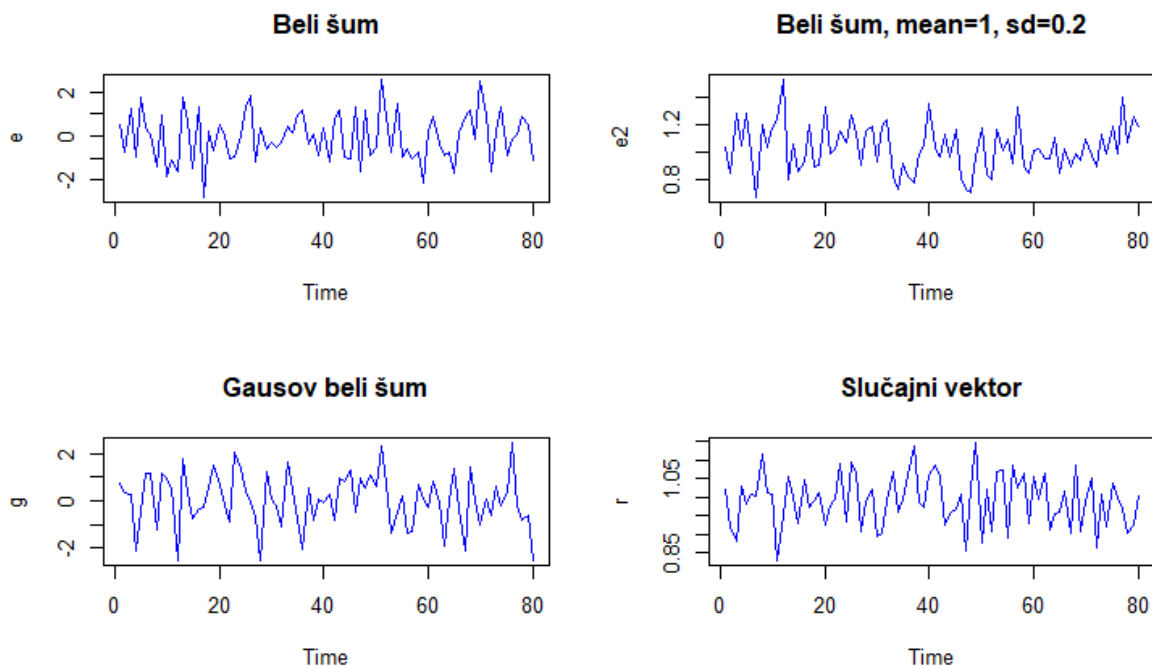
```
> g <- rnorm(80, mean=0, sd=1)
```

Pomoću funkcije `rnorm()` može se kreirati i slučajni vektor sa eksplicitno zadatom srednjom vrednošću i standardnom devijacijom. U sledećem primeru kreira se vektor sa 80 opservacija gde je srednja vrednost 1, a standardna devijacija 0.08.

```
> r <- rnorm(80, mean=1, sd=0.08)
```

Sve kreirane vremenske serije se mogu grafički prikazati pomoću funkcije `plot.ts()` i radi poređenja predstavljaju se na istom grafiku.

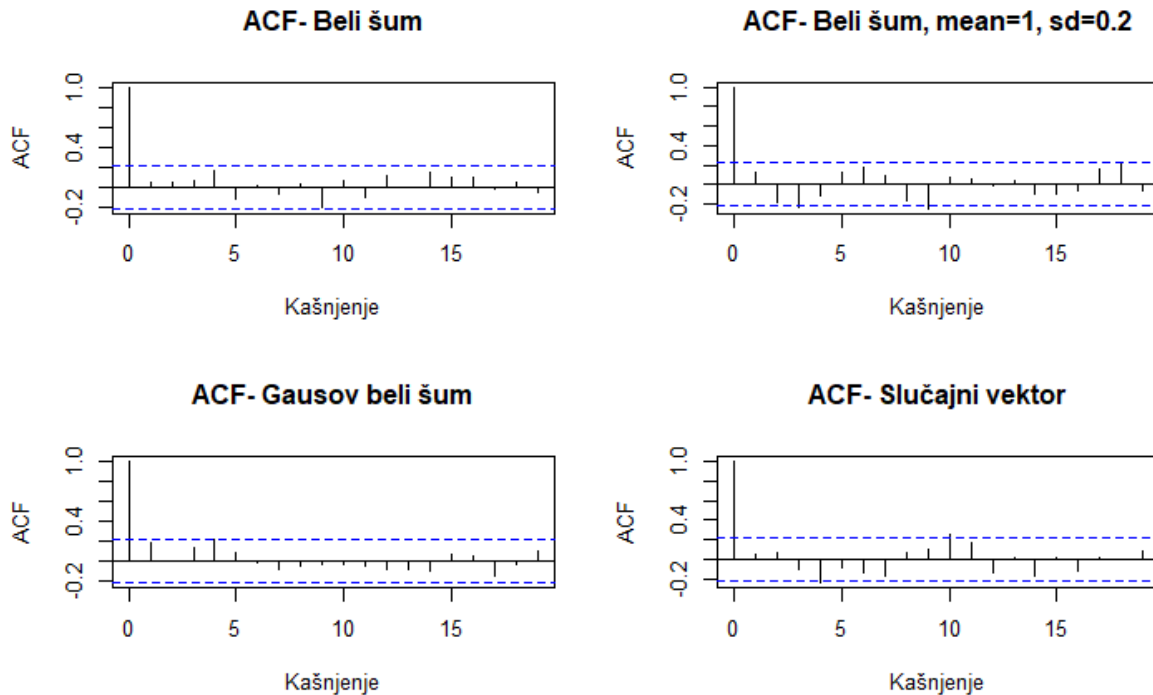
```
> par(mfrow=c(2,2))
> plot.ts(e, col="blue", main="Beli šum")
> plot.ts(e2, col="blue", main="Beli šum, mean=1, sd=0.2" )
> plot.ts(g, col="blue", main="Gausov beli šum")
> plot.ts(r, col="blue", main="Slučajni vektor")
```



Slika 6.1.1. Grafički prikaz Belog šuma (slika gore levo), Belog šuma sa zadatom srednjom vrednošću i standardnom devijacijom (slika gore desno), Gausovog belog šuma (slika dole levo) i Slučajnog vektora (slika dole desno)

U primerima u poglavlju 5 prikazani su grafički uzorački korelogrami slučajne komponente vremenskih serija. Grafički uzorački korelogram belog šuma je takav da nijedno kašnjenje nije značajno osim kašnjenja 0 i to je pokazatelj da je vremenska serija slučajna. U nastavku je prikazan grafički uzorački korelogram belog šuma, belog šuma sa srednjom vrednošću 1 i standardnom devijacijom 0.2, Gausovog belog šuma i slučajnog vektora.

```
> acf(e, main="ACF- Beli šum", xlab="Kašnjenje")
> acf(e2, main="ACF- Beli šum, mean=1, sd=0.2", xlab="Kašnjenje")
> acf(g, main="ACF- Gausov beli šum", xlab="Kašnjenje")
> acf(r, main="ACF- Slučajni vektor", xlab="Kašnjenje")
```



Slika 6.1.2. Grafički uzorački korelogram Belog šuma (slika gore levo), Belog šuma sa zadanom srednjom vrednošću i standardnom devijacijom (slika gore desno), Gausovog belog šuma (slika dole levo) i Slučajnog vektora (slika dole desno)

Dodavanje sezonske komponente i trenda

Potrebno je sada kreirati sezonsku komponentu. Kreira se npr. vektor proizvoljnih sezonskih vrednosti, gde je sezona kvartal. Koristi se funkcija `rep()` i proizvoljan vektor od 4 vrednosti pošto je u pitanju kvartal i ponavlja se 20 puta kako bi vektor `s` bio dužine 80. Međutim, s obzirom da se ovde radi o aditivnom modelu, mora se voditi računa o tome da zbir komponenti bude 0.

```
> s<-rep(c(-6,-2,2,6),times=20)
```

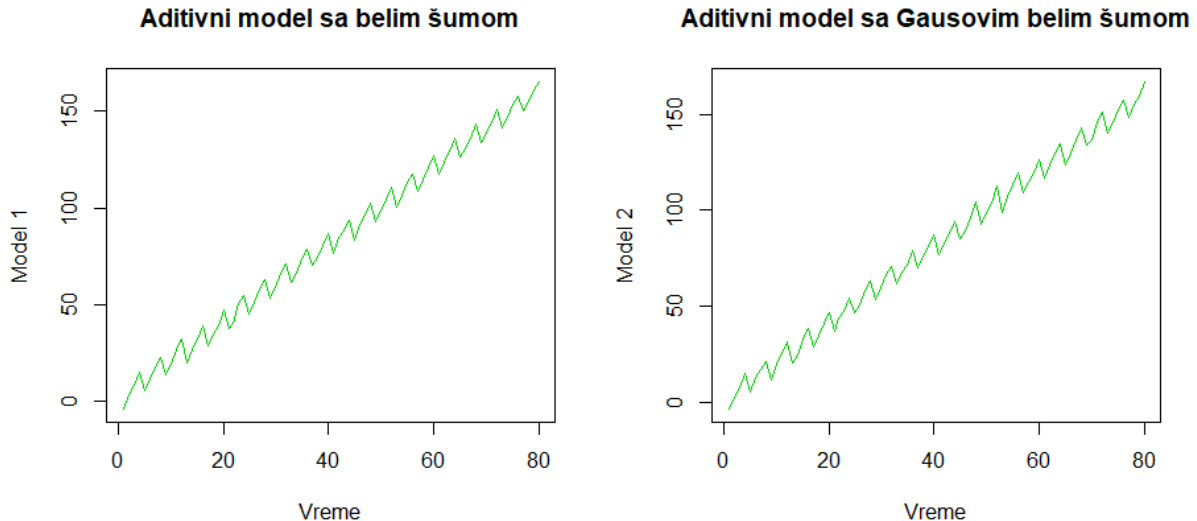
Zatim se kreira proizvoljan linearni trend.

```
> xt <- 1 : 80
> t<- 1 + 2 *xt
```

Kreiranje modela i simulacija

S obzirom da su kreirane sve komponente, one se mogu sabrati i tako dobiti aditivni model. Za prvi model koristi se prethodno kreiran beli šum `e`, a za drugi Gausov beli šum `g`. Oba modela predstavljaju se grafički.

```
> A1<-t+s+e
> A2<-t+s+g
> par(mfrow=c(1,2))
> plot.ts(A1, col=3,xlab="Vreme",ylab="Model 1", main="Aditivni model sa belim šumom")
> plot.ts(A2, col=3,xlab="Vreme",ylab="Model 2", main="Aditivni model sa Gausovim belim šumom")
```



Slika 6.1.3. Grafički prikaz aditivnog modela sa belim šumom (slika levo) i sa Gausovim belim šumom (slika desno)

Dobijaju se dva slična modela. Sa oba grafika se jasno vidi stabilna sezonalnost, prisustvo trenda i može se lako zaključiti da je u pitanju aditivni model. Sledi i prikaz sumarnih statistika, srednje vrednosti i standardne devijacije za oba modela i može se videti da su brojevi slični za oba modela.

```
> summary(A1)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-4.002	39.841	82.085	81.880	121.834	165.266

```
> sd(A1)
```

```
[1] 46.92473
```

```
> summary(A2)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-3.346	43.147	82.174	81.858	120.816	166.994

```
> sd(A2)
```

```
[1] 46.82987
```

Zanimljivo bi bilo pogledati kako bi izgledao model kada bi se simulirao više puta. Trend i slučajna komponenta ostaju isti uvek, ali se izbor slučajnih vrednosti kod belog šuma menja. Prvo se kreira funkcija koja pravi aditivni model.

```
> aditivni_model1 <- function(){
+
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   xt <- 1 : 80
+   t <- 1 + 2 *xt
+   s <- rep(c(-6, -2, 2, 6), times=20)
+   A <- t+s+e
+ }
```

Ovde je prvo korišćen beli šum **e**. Zatim se funkcija poziva npr. 1000 puta i vrednosti se smeštaju u matricu.

```
> v_simulacija <- replicate(1000, aditivni_model1())
```

Vrednosti modela upisuju se u kolone i s obzirom da se ponavlja 1000 puta, matrica ima 1000 kolona i 80 redova koliki je broj opservacija. Narednom petljom kreira se model sa prosečnom vrednošću kolona i smešta se u vektor **X**.

```

> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
>
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

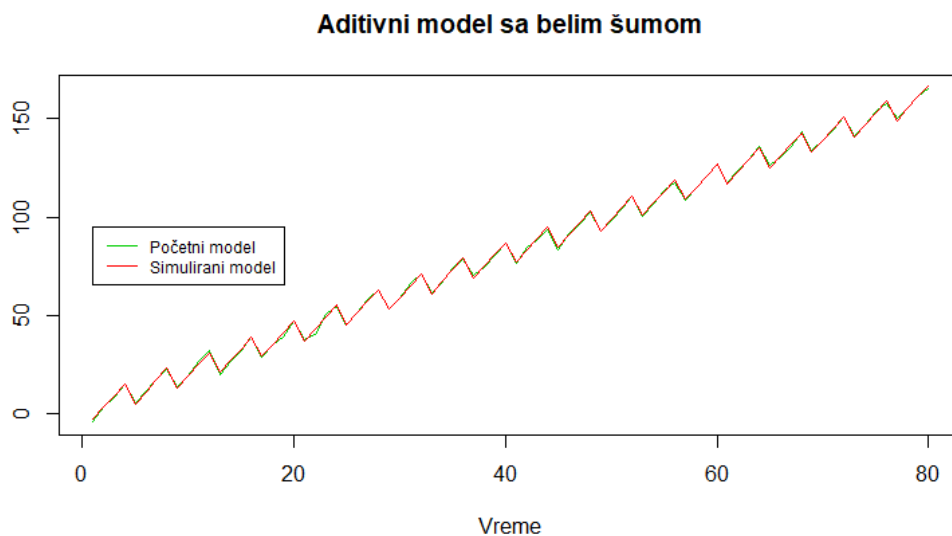
```

Oba modela se mogu predstaviti na jednom grafiku.

```

> par(mfrow=c(1,1))
> plot.ts(A1, col=3,xlab="Vreme",ylab=" ", main="Aditivni model sa belim šu
mom")
> lines(X, col=2)
> legend(1, 95, legend=c("Početni model", "Simulirani model"),
+       col=c(3, 2), lty=1:1, cex=0.75)

```



Slika 6.1.4. Grafički prikaz početnog i simuliranog aditivnog modela sa belim šumom

Simulacijom se dobije skoro identičan model kao početni.

Sada se može simulirati model sa Gausovim belim šumom, **g**.

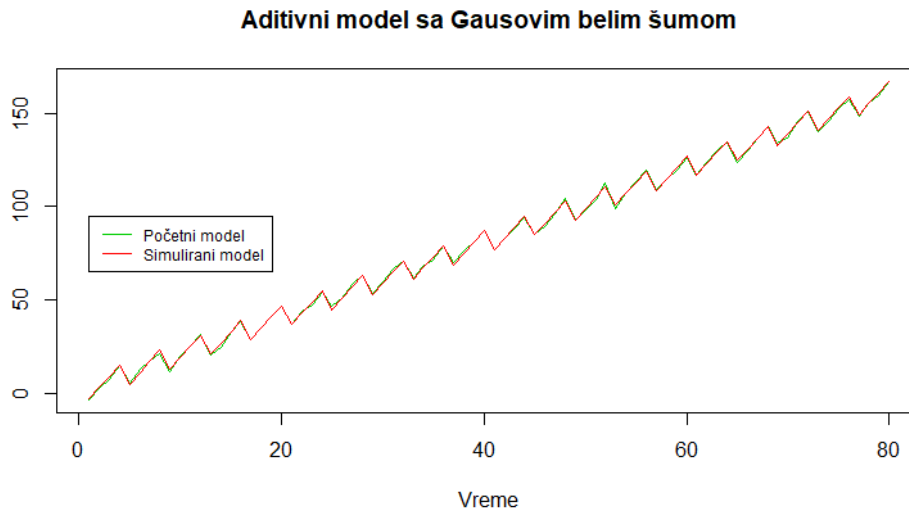
```

> aditivni_model2 <- function(){
+   g <- rnorm(80,mean=0,sd=1)
+   xt <- 1 : 80
+   t <- 1 + 2 *xt
+   s<-rep(c(-6,-2,2,6),times=20)
+   A<-t+s+g
+ }
> v_simulacija <- replicate(1000, aditivni_model2())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
>
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

```

Simulirani model i početni model se mogu predstaviti na istom grafiku.

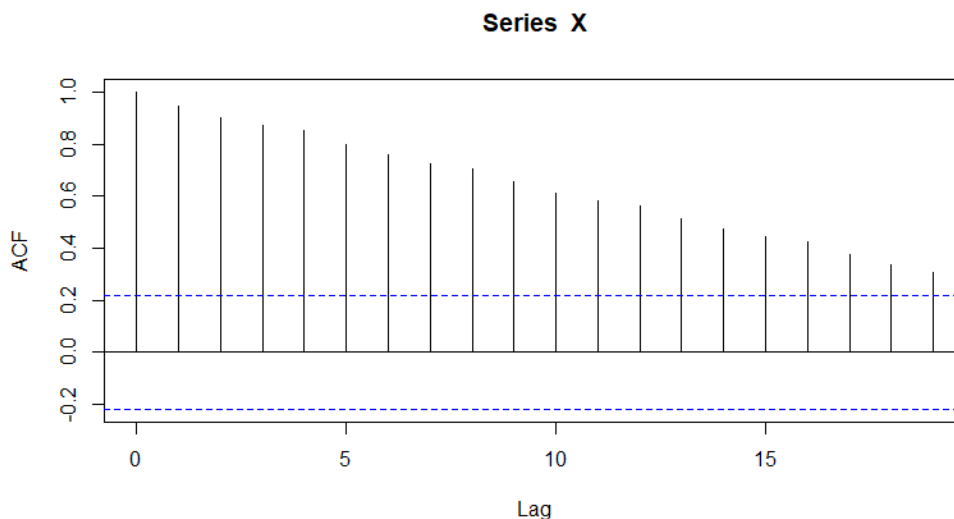
```
> plot.ts(A2, col=3,xlab="Vreme",ylab=" ", main="Aditivni model sa Gausovim
belim šumom")
> lines(x, col=2)
> legend(1, 95, legend=c("Početni model", "Simulirani model"),
+       col=c(3, 2), lty=1:1, cex=0.75)
```



Slika 6.1.5. Grafički prikaz početnog i simuliranog aditivnog modela sa Gausovim belim šumom

Simulirani model je sličan kao početni, a takođe, sličan je i modelu koji koristi običan beli šum, e . Može se prikazati i grafički uzorački korelogram dobijenog simuliranog modela X .

```
> acf(X)
```



Slika 6.1.6. Grafički uzorački korelogram simuliranog aditivnog modela

Sa grafičkog uzoračkog korelograma se jasno vidi da postoji linearni trend, kao i da je stabilna sezonalnost.

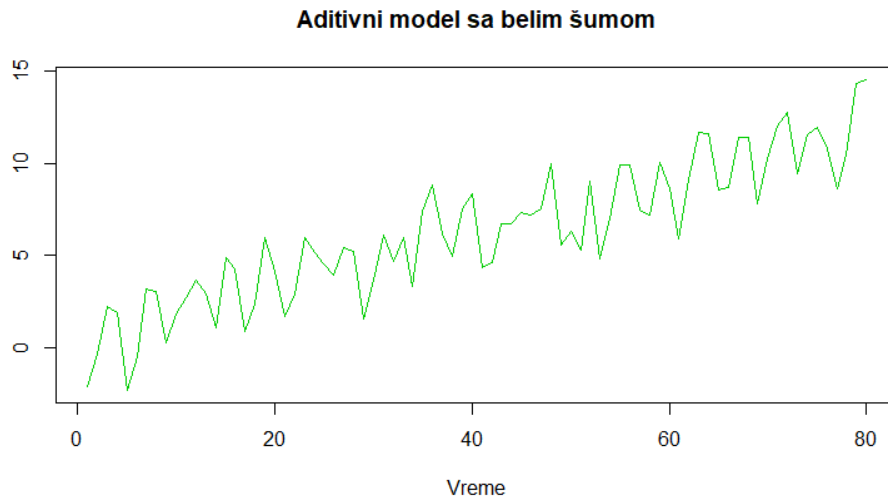
Simulacijom se nekada i ne dobiju rezultati kakvi se očekuju. Naime, simuliranjem aditivnog modela neće se uvek dobiti ovako jasno izražena stabilna sezonalnost, što se može videti u sledećem primeru.

```
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(-0.95,-1,0.75,1.2),times=20)
```

```
> t<-seq(0.5, 12, length.out=80)
> A<-t+s+e
```

Uzete su iste vrednosti za beli šum e , sezonska komponenta je sa proizvoljnim vrednostima, ali opet u zbiru daju 0, dok su vrednosti trenda manje i uzete proizvoljno korišćenjem funkcije `seq()`. Model je predstavljen grafički.

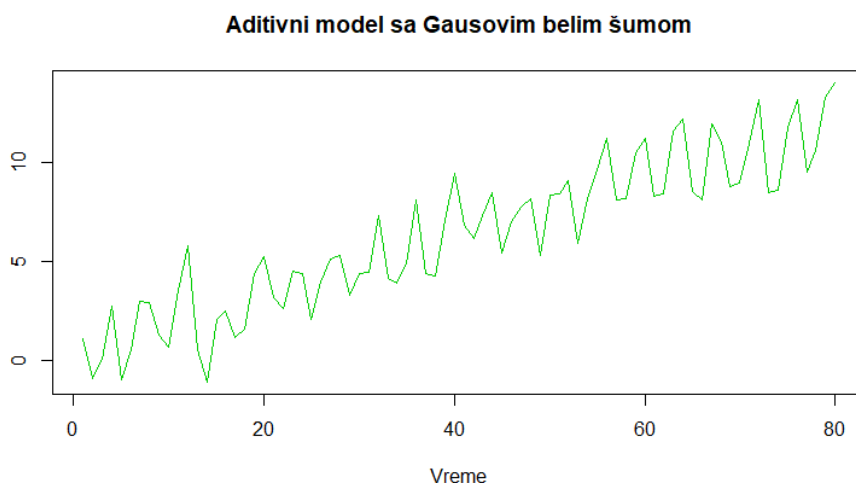
```
> plot.ts(A, col=3, xlab="vreme",ylab=" ",main="Aditivni model sa belim šumom")
```



Slika 6.1.7. Grafički prikaz aditivnog modela sa belim šumom

Sa grafika se ne vidi jasno izražena sezonalnost, dok se vidi da postoji trend. Korišćenjem Gausovog belog šuma, dobije se i ovde sličan model.

```
> g <- rnorm(80,mean=0,sd=1)
> A<-t+s+g
> plot.ts(A, col=3,xlab="vreme",ylab=" ", main="Aditivni model sa Gausovim belim šumom")
```



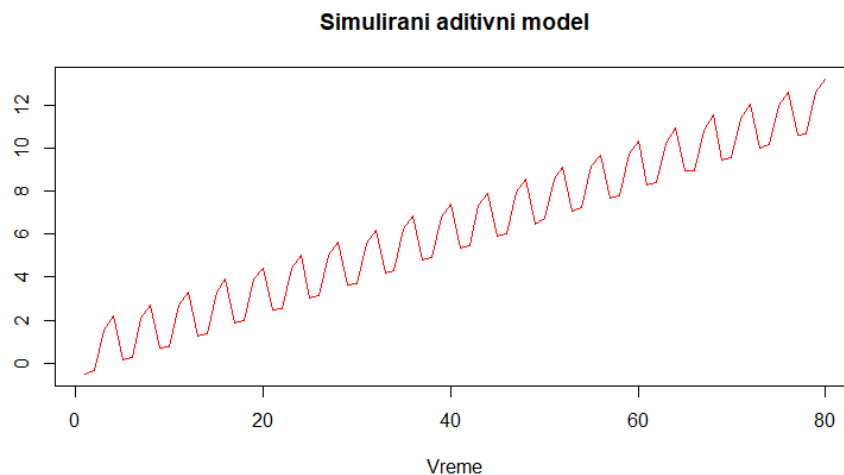
Slika 6.1.8. Grafički prikaz aditivnog modela sa Gausovim belim šumom

Može se uraditi simulacija i ovog modela i možda prosečan model bude dao jasnije izražene sezone varijacije. Koristi se beli šum e .

```

> aditivni_model1 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   s<-rep(c(-0.95,-1,0.75,1.2),times=20)
+   t<-seq(0.5, 12, length.out=80)
+   A<-t+s+e
+ }
> v_simulacija <- replicate(1000, aditivni_model1())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
> plot.ts(X,col=2,xlab="Vreme",ylab=" ", main="Simulirani aditivni model")

```



Slika 6.1.9. Grafički prikaz simuliranog aditivnog modela sa belim šumom

Sezonski obrazac se sada jasno vidi.

Ako se simulira model sa istim trendom i belim šumom, a sezonska komponenta ima npr. samo vrednosti -5 i 5, dobiće se model sličan prvom sa još pravilnije izraženom sezonskom komponentom. U nastavku je primer.

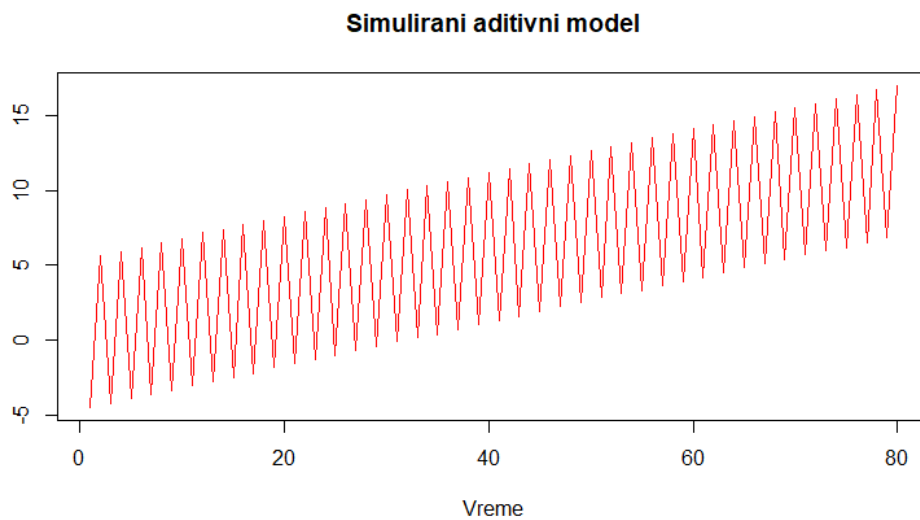
```

> aditivni_model1 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   t<-seq(0.5, 12, length.out=80)
+   s<-rep(c(-5,5),times=40)
+   A<-t+s+e
+ }
> v_simulacija <- replicate(1000, aditivni_model1())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

```



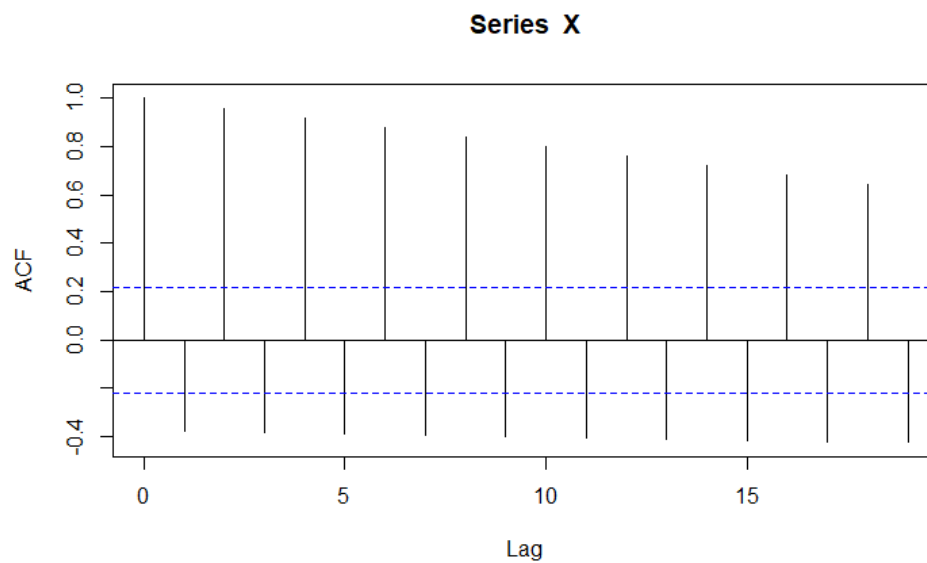
```
> plot.ts(x,col=2, main="simulirani aditivni model")
```



Slika 6.1.10. Grafički prikaz simuliranog aditivnog modela sa belim šumom

Dakle, jasni i stabilni sezonski obrasci daju nam do znanja da je u pitanju aditivni model. I za ovakav model se može pogledati kako izgleda grafički uzorački korelogram.

```
> acf(x)
```



Slika 6.1.11. Grafički uzorački korelogram simuliranog aditivnog modela

6.2. SIMULACIJA MULTIPLIKATIVNIH MODELA

Za simulaciju multiplikativnog modela mogu se uzeti slične vrednosti kao za simulaciju aditivnog modela i model prikazati grafički. Međutim, prilikom izbora komponenti mora se voditi računa da one imaju pozitivne vrednosti, kao i da prosek vrednosti sezonske komponente bude jednak 1. Ovde se za slučajnu komponentu uzima vremenska serija sa srednjom vrednošću 1 i standardnom devijacijom 0.2 koja je opisana u prethodnom primeru.

```
> e2 <- arima.sim(model = list(order = c(0, 0, 0)), n = 80, mean=1, sd=0.2)
```

Za trend i sezonsku komponentu uzimaju se sledeće vrednosti:

```
> s<-rep(c(0.25,0.75,1.25, 1.75),times=20)
> xt <- 1 : 80
> t<- 1 + 2 *xt
```

Kreira se multiplikativni model od ovih komponenti.

```
> M1<-t*s*e2
```

Umesto belog šuma može se koristiti i slučajni vektor koji je takođe kreiran u prethodnom primeru.

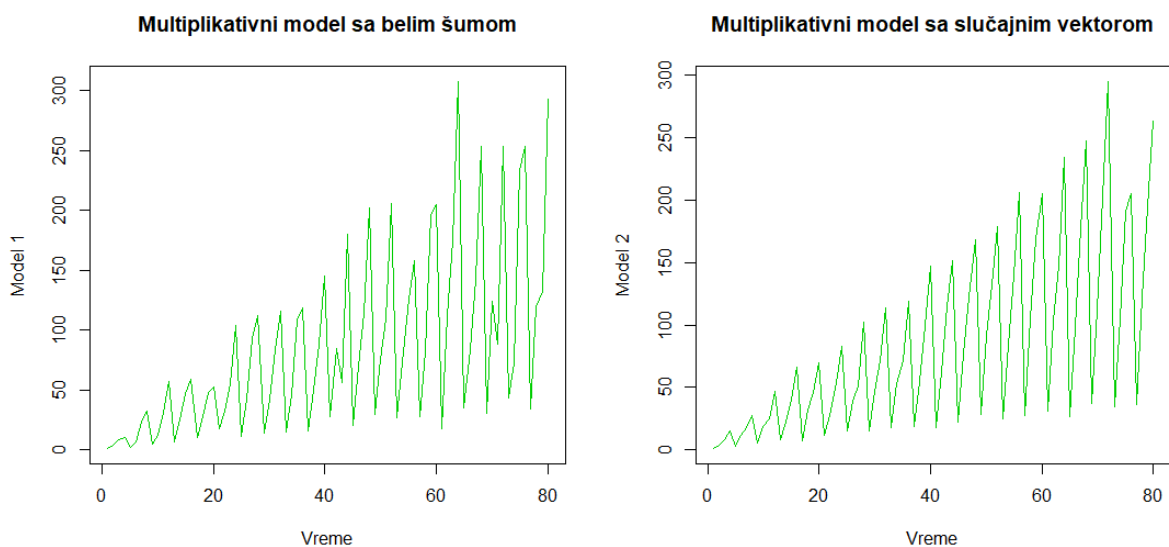
```
> r <- rnorm(80,mean=1,sd=0.08)
```

Kreira se novi multiplikativni model sa ovim slučajnim vektorom, trendom i sezonskom komponentom koji su isti kao i za prvi multiplikativni model M1.

```
> M2<-t*s*r
```

Oba modela se predstavljaju grafički na istoj slici radi lakšeg poređenja.

```
> par(mfrow=c(1,2))
> plot.ts(M1, col=3, main="Multiplikativni model sa belim šumom")
> plot.ts(M2, col=3, main="Multiplikativni model sa slučajnim vektorom")
```



Slika 6.2.1. Grafički prikaz multiplikativnog modela sa sa belim šumom(slika levo) i slučajnim vektorom(slika desno)

Na grafiku levo se jasno vidi trend i takođe se vidi da se sezonske varijacije povećavaju sa vremenom, iako one nisu jasno izražene.

Na grafiku desno se jasnije vide sezonske varijacije i vidi se da se one povećavaju sa vremenom. Takođe se vidi i postojanje trenda.

Sledi i prikaz sumarnih statistika, srednje vrednosti i standardne devijacije za oba modela.

```
> summary(M1)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.7437 26.9808 56.3631 83.0842 116.5646 307.8465

> sd(M1)
[1] 75.35601

> summary(M2)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.7965 24.8098 58.3720 83.4336 121.7785 295.4630

> sd(M2)
[1] 72.38967
```

Simulacija modela

U narednim koracima radi se simulacija multiplikativnog modela sa belim šumom. Prvo se kreira funkcija koja pravi multiplikativni model.

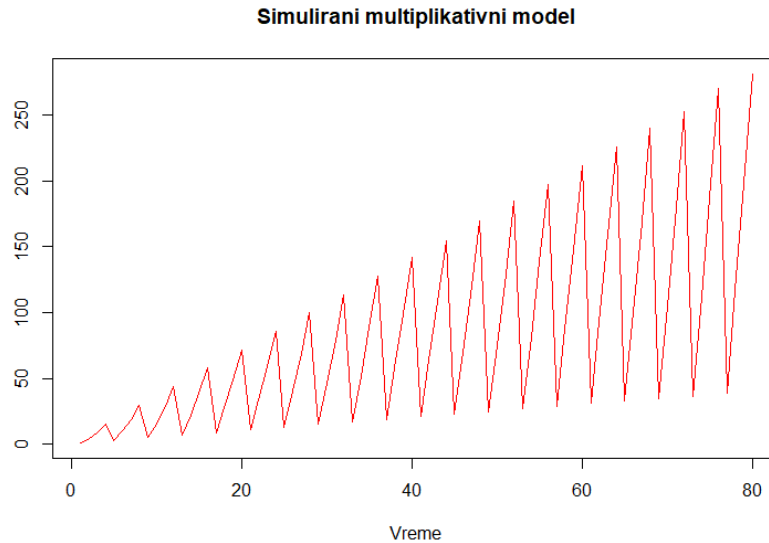
```
> multiplikativni_model1 <- function(){
+
+   e2 <- arima.sim(model = list(order = c(0, 0, 0)), n = 80, mean=1, sd=
0.2)
+   xt <- 1 : 80
+   t <- 1 + 2 *xt
+   s<-rep(c(0.25,0.75,1.25, 1.75),times=20)
+   M1<-t*s*e2
+
+ }
```

Zatim se funkcija poziva 1000 puta i vrednosti se smeštaju u matricu.

```
> v_simulacija <- replicate(1000, multiplikativni_model1())
```

Vrednosti modela upisuju se u kolone i s obzirom da se ponavlja 1000 puta, matrica ima 1000 kolona i 80 redova koliki je broj opservacija. Narednom petljom kreira se model sa prosečnom vrednošću kolona i smešta se u vektor X.

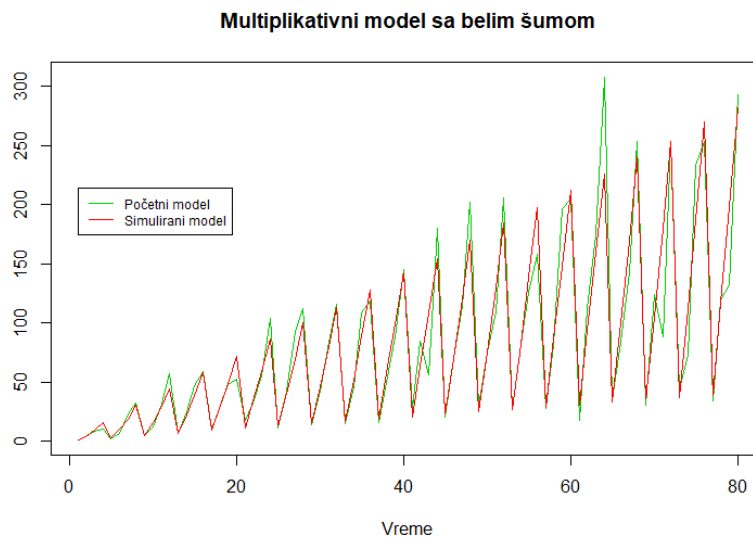
```
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
> par(mfrow=c(1,1))
> plot.ts(X,col=2, main="Simulirani multiplikativni model")
```



Slika 6.2.2. Grafički prikaz simuliranog multiplikativnog modela

Simulacijom se dobiju jasno izražene sezonske komponente koje se očigledno povećavaju tokom vremena, i ovaj grafik daje pravi prikaz multiplikativnog modela. Može se i prikazati na istom grafiku sa početnim modelom.

```
> plot.ts(M1, col=3, xlab="vreme", ylab=" ", main="Multiplikativni model sa
belim šumom")
> lines(x, col=2)
> legend(1, 95, legend=c("Početni model", "Simulirani model"), col=c(3, 2)
, lty=1:1, cex=0.75, yjust = -2 )
```



Slika 6.2.3. Grafički prikaz početnog i simuliranog multiplikativnog modela

Simulira se i model sa slučajnim vektorom r na isti način.

```
> multiplikativni_model2 <- function(){
+   r <- rnorm(80,mean=1,sd=0.08)
+   xt <- 1 : 80
+   t <- 1 + 2 *xt
+   s<-rep(c(0.25,0.75,1.25, 1.75),times=20)
+   M2<-t*s*r
+ }
```

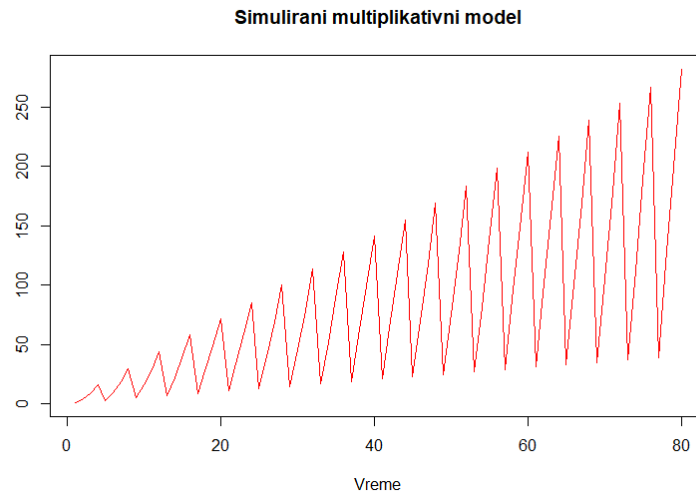
```

> v_simulacija <- replicate(1000, multiplikativni_model2())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)

> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

> plot.ts(X,col=2, xlab="Vreme", ylab=" ", main="Simulirani multiplikativni
model")

```



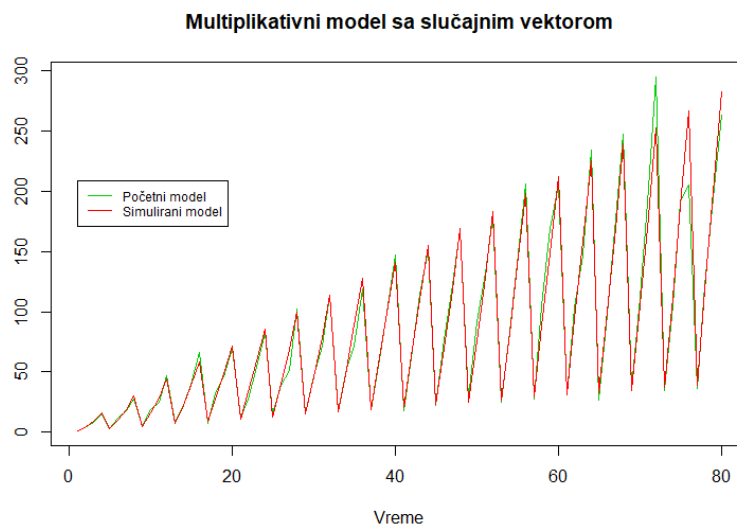
Slika 6.2.4. Grafički prikaz simuliranog multiplikativnog modela sa slučajnim vektorom

Ovom simulacijom dobija se sličan model kao i kad se koristi beli šum e_2 . Može se prikazati početni i simulirani model na jednom grafiku.

```

> plot.ts(M2, col=3, xlab="Vreme", ylab=" ", main="Multiplikativni model sa
slučajnim vektorom")
> lines(X, col=2)
> legend(1, 95, legend=c("Početni model", "Simulirani model"), col=c(3, 2),
lty=1:1,
+       cex=0.75, yjust = -2 )

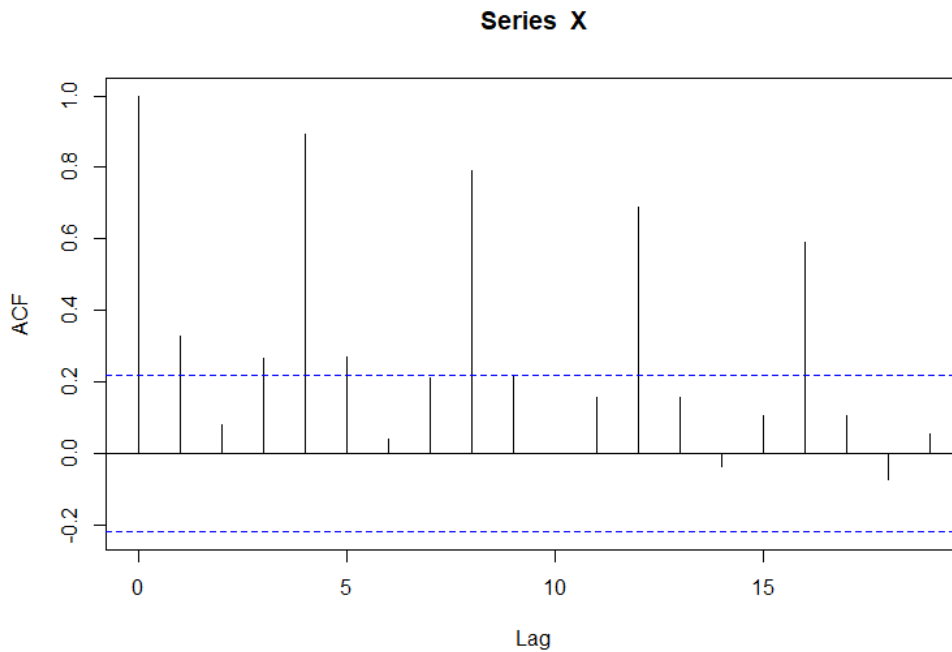
```



Slika 6.2.5. Grafički prikaz početnog i simuliranog multiplikativnog modela sa slučajnim vektorom

Može se pogledati i grafički uzorački korelogram dobijenog simuliranog modela **X**.

> `acf(X)`

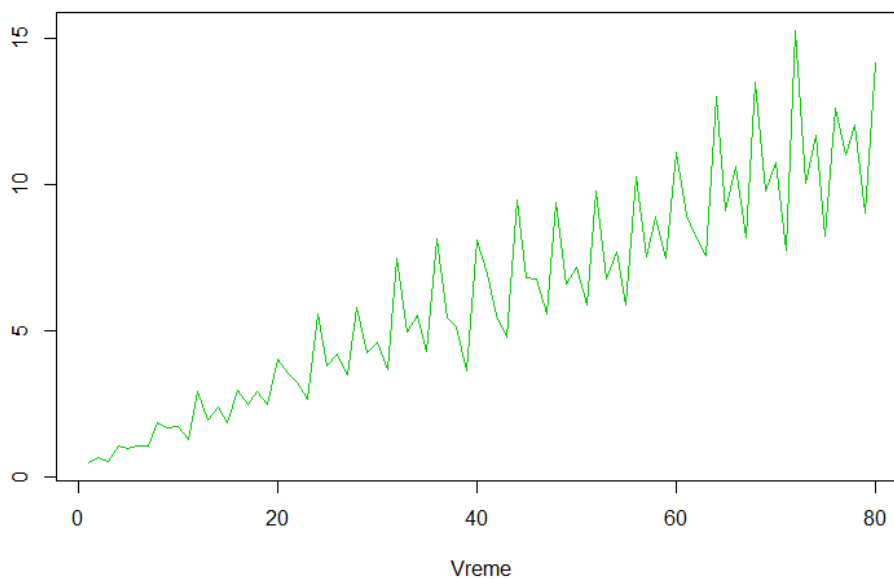


Slika 6.2.6. Grafički uzorački korelogram simuliranog multiplikativnog modela

Za sledeći primer uzima se malo drugačija sezonska komponenta i trend dobijen korišćenjem funkcije `seq()` kao u primeru 6.1. Prvo se uzima isti vektor slučajnih vrednosti `r` i model se predstavlja grafički.

```
> r <- rnorm(80, mean=1, sd=0.08)
> s <- rep(c(0.95, 1, 0.75, 1.3), times=20)
> t <- seq(0.5, 12, length.out=80)
> M3 <- t*s*r
> plot.ts(M3, col=3, xlab="Vreme", ylab=" ", main="Multiplikativni model sa
slučajnim vektorom")
```

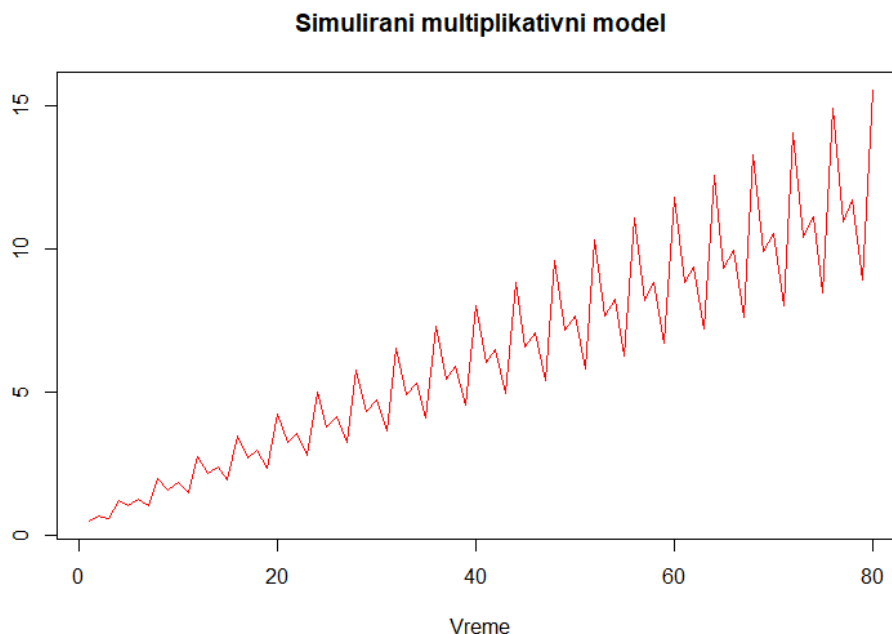
Multiplikativni model sa slučajnim vektorom



Slika 6.2.7. Grafički prikaz multiplikativnog modela sa slučajnim vektorom

Sa grafika se vidi postojanje trenda, vidi se da postoje neki sezonski obrasci koji se vremenom povećavaju, ali nisu baš najjasniji. Ali model opet može da se simulira 1000 puta i da se vidi prosečan model kako izgleda.

```
> multiplikativni_model3 <- function(){
+   r <- rnorm(80,mean=1,sd=0.08)
+   s<-rep(c(0.95,1,0.75,1.3),times=20)
+   t<-seq(0.5, 12, length.out=80)
+   M3<-t*s*r
+ }
> v_simulacija <- replicate(1000, multiplikativni_model3())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
> plot.ts(X,col=2, xlab="vreme", ylab=" ", main="Simulirani multiplikativni
i model")
```

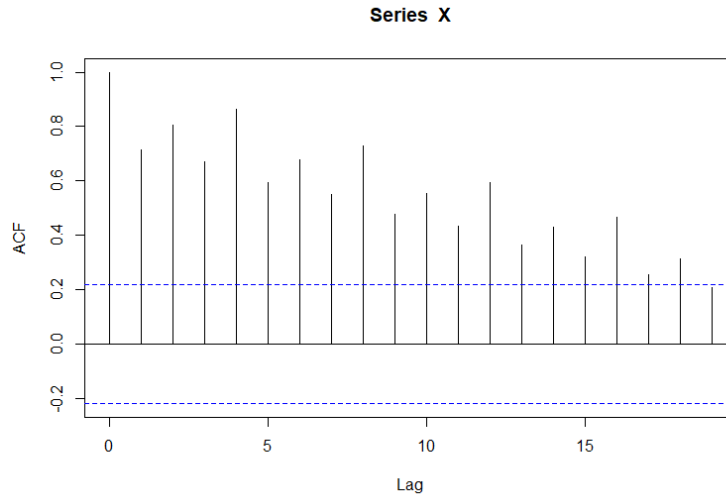


Slika 6.2.8. Grafički prikaz simuliranog multiplikativnog modela sa slučajnim vektorom

Sada se jasno vide sezonske varijacije i one se vidno povećavaju tokom vremena, pa se za ovakav model može jasno zaključiti da je multiplikativni. Sličan model se dobija i simulacijom modela korišćenjem belog šuma e_2 umesto vektora slučajnih vrednosti r .

Može se predstaviti i grafički uzorački korelogram ovog modela.

```
> acf(X)
```



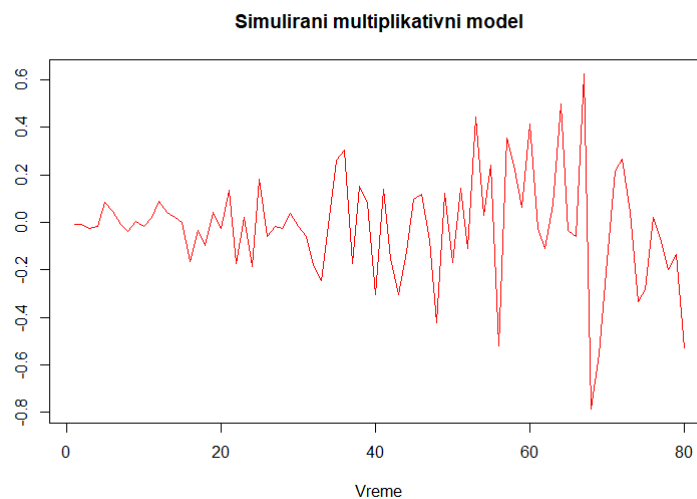
Slika 6.2.9. Grafički uzorački korelogram simuliranog modela

U sledećem primeru uzima se beli šum e , koji je korišćen u primeru 6.1. i pokušava da se simulira multiplikativni model. S obzirom da vektor e ima negativne vrednosti ovo nije dobar primer multiplikativnog modela.

```
> multiplikativni_model4 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   s<-rep(c(0.95,1,0.75,1.3),times=20)
+   t<-seq(0.5, 12, length.out=80)
+   M4<-t*s*e
+ }
> v_simulacija <- replicate(1000, multiplikativni_model4())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)

> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

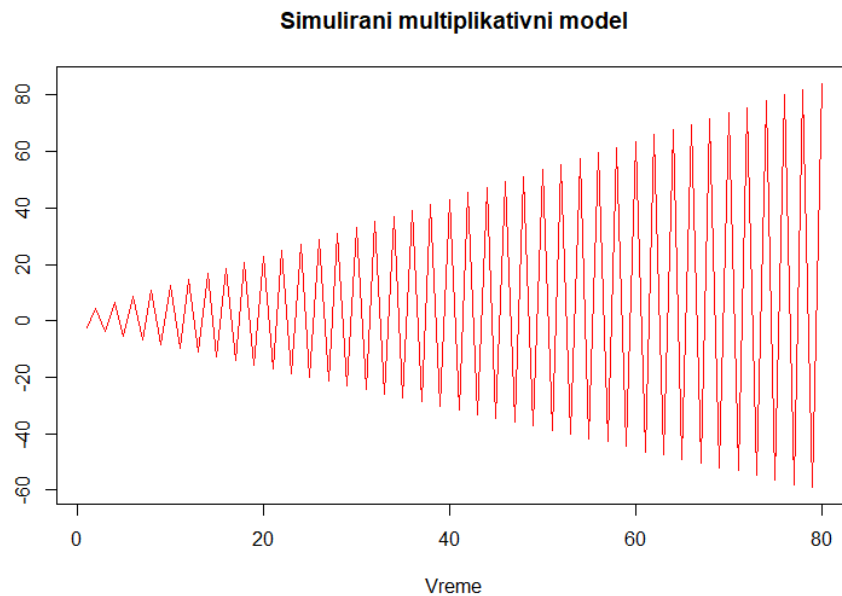
> plot.ts(X,col=2, xlab="Vreme", ylab=" ", main="Simulirani multiplikativni
model")
```



Slika 6.2.10. Grafički prikaz simuliranog multiplikativnog modela sa negativnim vrednostima

Može se simulirati model sa slučajnim vektorom r , trendom kao u prethodnom primeru i sezonskom komponentom koja ima vrednosti -5 i 7.

```
> multiplikativni_model5 <- function(){
+   r <- rnorm(80,mean=1,sd=0.08)
+   s<-rep(c(-5,7),times=40)
+   t<-seq(0.5, 12, length.out=80)
+   M5<-t*s*r
+ }
> v_simulacija <- replicate(1000, multiplikativni_model5())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
> plot.ts(X,col=2, xlab="Vreme", ylab=" ", main="Simulirani multiplikativni
model")
```



Slika 6.2.11. Grafički prikaz simuliranog multiplikativnog modela sa negativnim vrednostima

Ovo opet nije dobar primer multiplikativnog modela pošto model X ima negativne vrednosti.

6.3. SIMULACIJA MEŠOVITIH MODELA

Prilikom simulacije mešovitih modela koriste se slične vrednosti za beli šum, trend i sezonsku komponentu kao kod simulacije multiplikativnih modela. Simulira se prvo model koji množi trend i sezonsku komponentu, a dodaje slučajnu komponentu. Za slučajnu komponentu uzima se beli šum e . Njegova formula je $X_t = f_t \cdot S_t + e_t$.

```
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(0.25,0.75,1.25, 1.75),times=20)
> xt <- 1 : 80
> t<- 1 + 2 *xt
> N1<-t*s+e
```

U narednim koracima radi se simulacija ovog modela, na sličan način kao i u primerima 6.1. i 6.2.

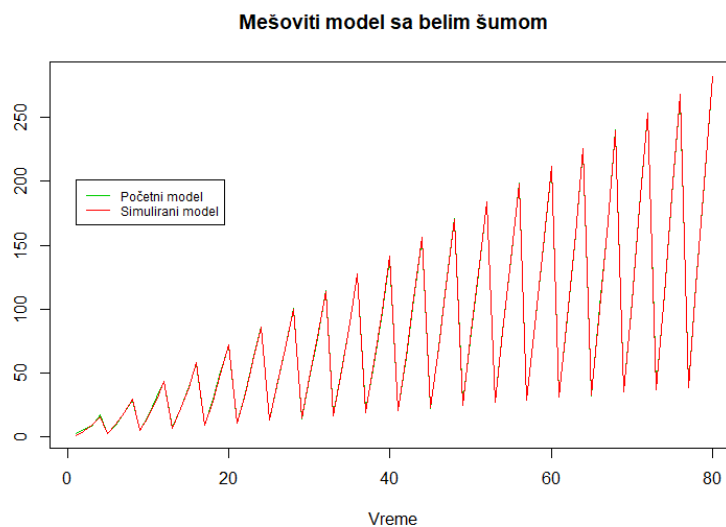
```
> mešoviti_model1 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   xt <- 1 : 80
+   t <- 1 + 2 *xt
+   s<-rep(c(0.25,0.75,1.25, 1.75),times=20)
+   N1<-t*s+e
+ }
> v_simulacija <- replicate(1000, mešoviti_model1())

> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)

> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
```

Početni model i simulirani model predstavljaju se na istom grafiku.

```
> plot.ts(N1, col=3, xlab="Vreme", ylab=" ", main="Mešoviti model sa belim
šumom")
> lines(x, col=2)
> legend(1, 95, legend=c("Početni model", "Simulirani model"),
+       col=c(3, 2), lty=1:1, cex=0.75,yjust = -2)
```

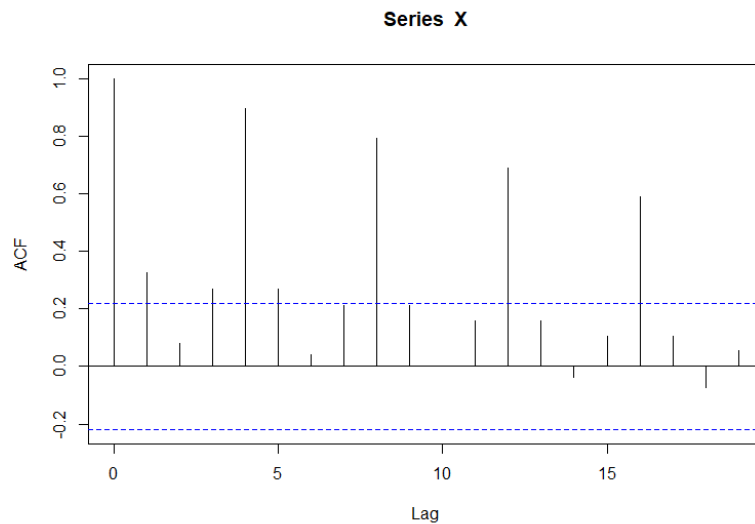


Slika 6.3.1. Grafički prikaz početnog i simuliranog mešovitog modela sa belim šumom

Simulirani model je skoro identičan kao početni. Vidi se da je ovaj model jako sličan multiplikativnom modelu. I ovde se sezonske varijacije postepeno povećavaju sa vremenom i jasno je izražen trend. Ovaj model ima samo pozitivne vrednosti, pa se sa grafika ne bi moglo zaključiti da li je mešoviti ili multiplikativni.

Može se predstaviti i grafički uzorački korelogram simuliranog modela.

> acf(X)

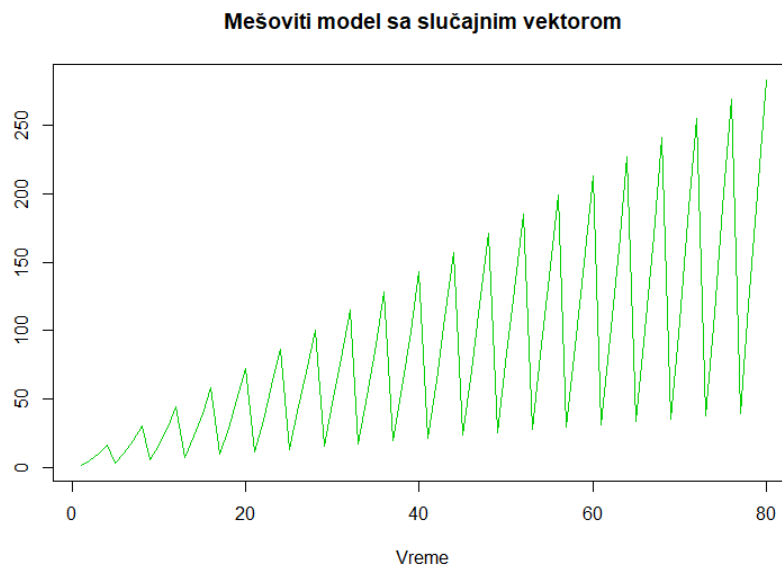


Slika 6.3.2. Grafički uzorački korelogram simuliranog mešovitog modela

I grafički uzorački korelogram simuliranog mešovitog modela je sličan kao i kod multiplikativnog.

Ako se koristi slučajni vektor r prilikom kreiranja mešovitog modela, takođe se dobije skoro identičan model multiplikativnom.

```
> r <- rnorm(80, mean=1, sd=0.08)
> s <- rep(c(0.25, 0.75, 1.25, 1.75), times=20)
> xt <- 1 : 80
> t <- 1 + 2 * xt
> N2 <- t * s + r
> plot.ts(N2, col=3, xlab="Vreme", ylab=" ", main="Mešoviti model sa slučajnim vektorom")
```

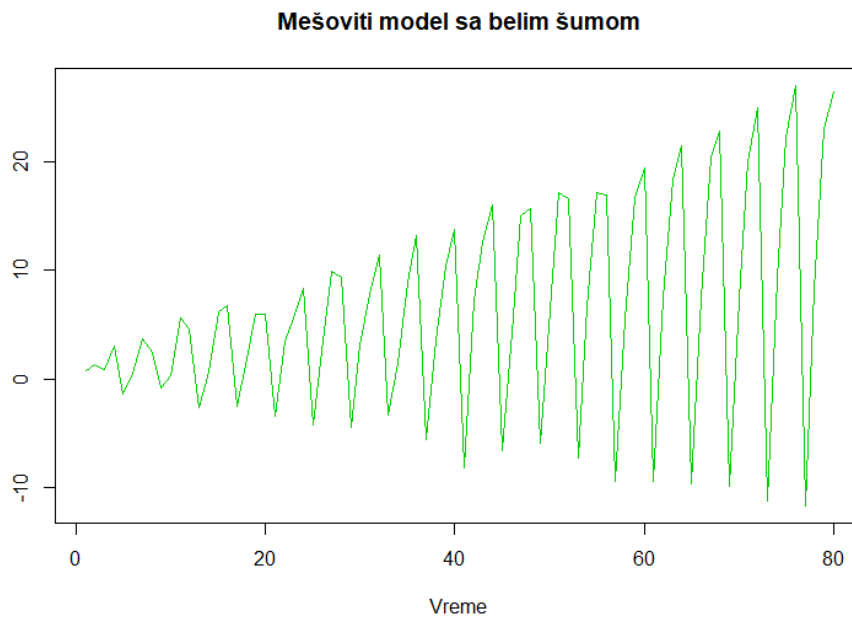


Slika 6.3.3. Grafički prikaz mešovitog modela sa slučajnim vektorom

Slični modeli se dobiju i korišćenjem Gausovog belog šuma g , kao i vremenske serije e_2 .

Mogu se i uzeti neke negativne vrednosti npr. za sezonsku komponentu i pogledati kakav se model dobija.

```
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(-1,0.75,2,2.25),times=20)
> t<-seq(0.5, 12, length.out=80)
> N3<-t*s+e
> plot.ts(N3, col=3,xlab="Vreme", ylab=" ", main="Mešoviti model sa belim šumom")
```



Slika 6.3.4. Grafički prikaz mešovitog modela sa belim šumom i negativnim vrednostima

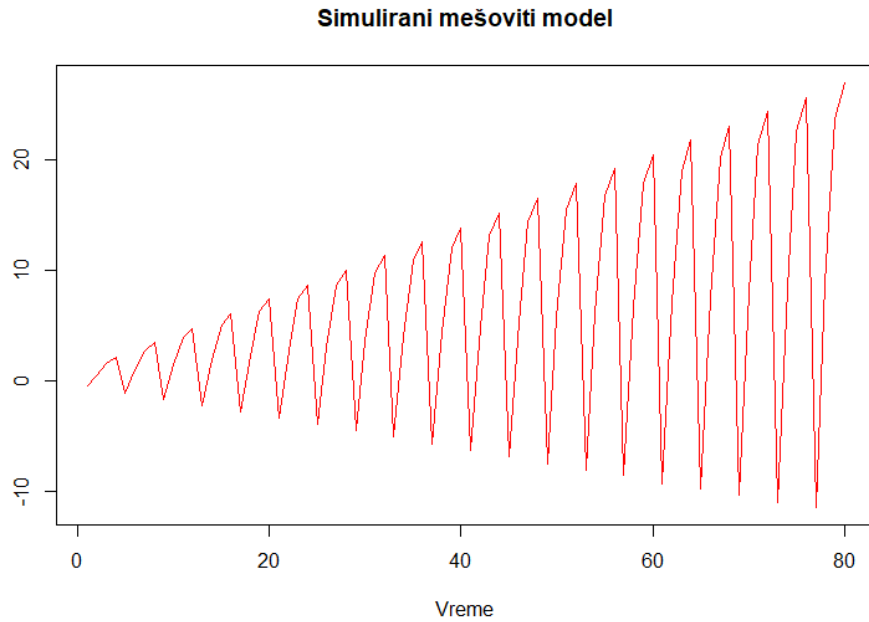
Sezonske varijacije se povećavaju vremenom kao kod multiplikativnog modela, s tim što ovde postoje i negativne vrednosti. Simulacijom se dobija sličan model kao u poslednjem primeru multiplikativnog modela.

```
> mešoviti_model2 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   s<-rep(c(-1,0.75,2,2.25),times=20)
+   t<-seq(0.5, 12, length.out=80)
+   N3<-t*s+e
+ }

> v_simulacija <- replicate(1000, mešoviti_model2())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)

> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

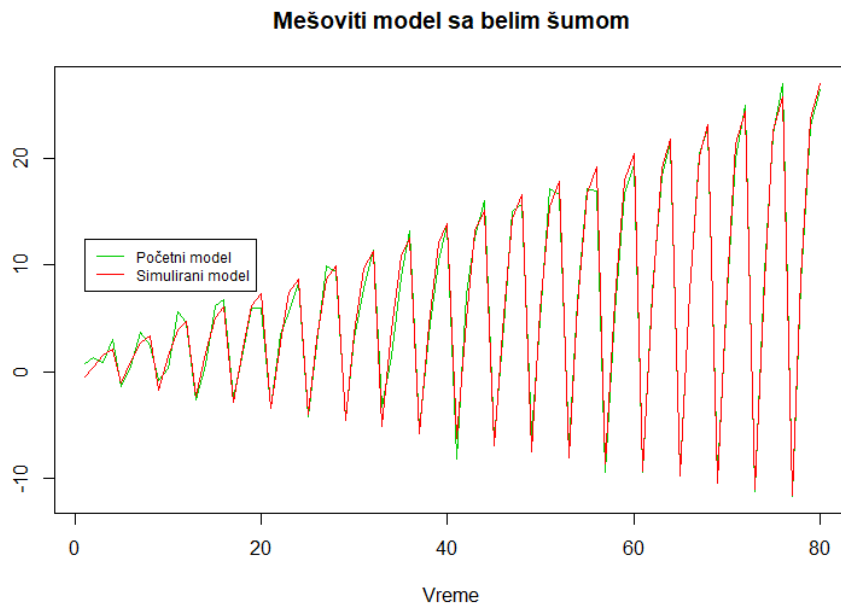
> plot.ts(X,col=2, main="Simulirani mešoviti model")
```



Slika 6.3.5. Grafički prikaz simuliranog mešovitog modela sa belim šumom i negativnim vrednostima

Početni i simulirani model se mogu predstaviti na istom grafiku.

```
> plot.ts(N3, col=3, xlab="vreme", ylab=" ", main="Mešoviti model sa belim
šumom")
> lines(x, col=2)
> legend(1, 10, legend=c("Početni model", "Simulirani model"),
+       col=c(3, 2), lty=1:1, cex=0.75, yjust = 0.5)
```

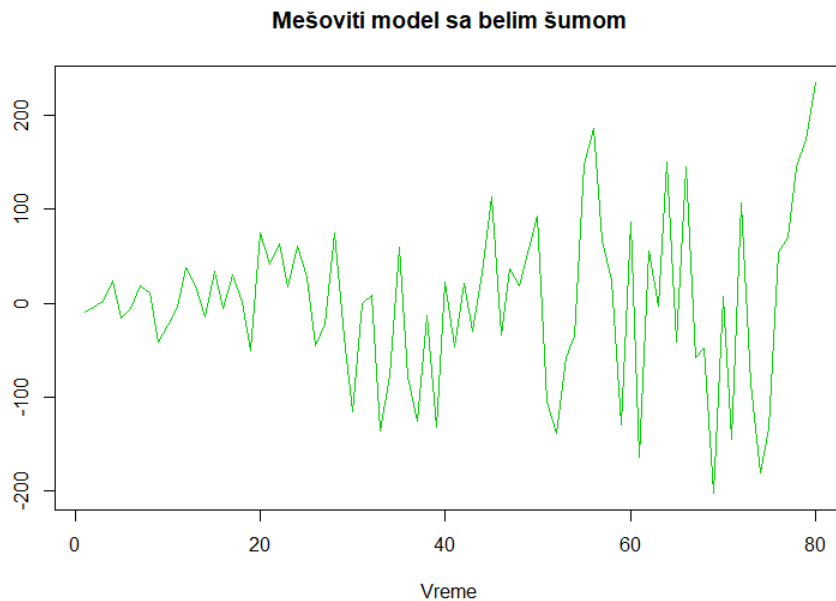


Slika 6.3.6. Grafički prikaz početnog i simuliranog mešovitog modela sa belim šumom i negativnim vrednostima

Dakle, ovakav model sa pozitivnim vrednostima je jako sličnog izgleda kao mešoviti model.

Može se prikazati i mešoviti model koji kombinuje drugačije komponente, npr. množi trend i slučajnu komponentu a dodaje sezonsku komponentu. Njegova formula je $X_t = f_t \cdot e_t + S_t$. Ovde je uzeta sezonska komponenta kao kod aditivnog modela.

```
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(-6,-2,2,6),times=20)
> xt <- 1 : 80
> t<- 1 + 2 *xt
> N4<-t*e+s
> plot.ts(N4, col=3, xlab="Vreme", ylab=" ", main="Mešoviti model sa belim
šumom")
```



Slika 6.3.7. Grafički prikaz mešovitog modela sa belim šumom

U narednim koracima radi se simulacija ovakvog modela.

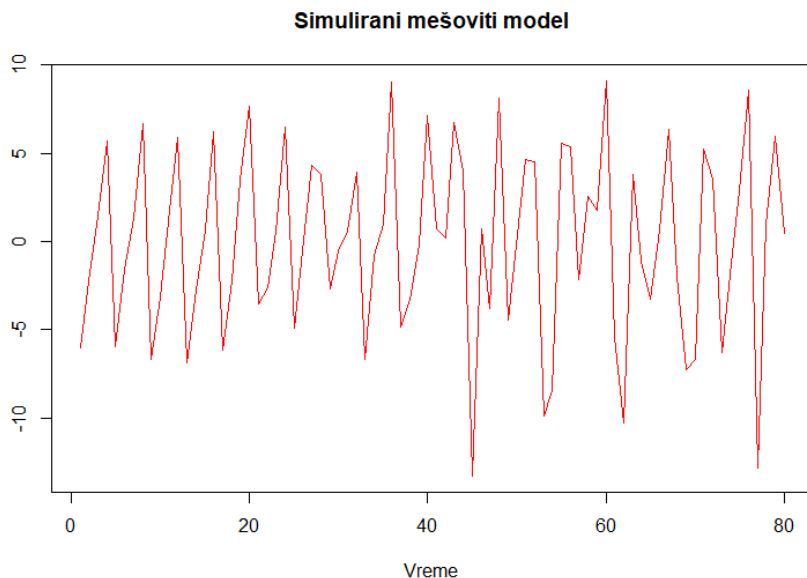
```
> mešoviti_model3 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   s<-rep(c(-6,-2,2,6),times=20)
+   xt <- 1 : 80
+   t<- 1 + 2 *xt
+   N4<-t*e+s
+ }

> v_simulacija <- replicate(1000, mešoviti_model3())

> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)

> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }

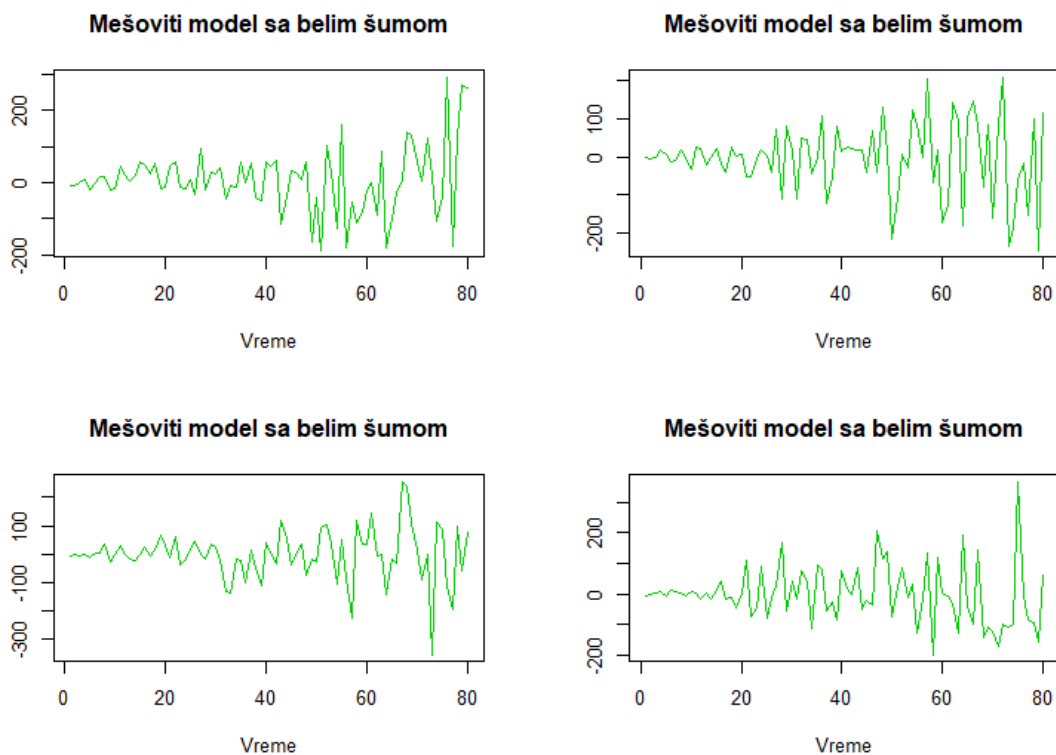
> plot.ts(X,col=2, xlab="Vreme", ylab=" ", main="Simulirani mešoviti model"
)
```



Slika 6.3.8. Grafički prikaz simuliranog mešovitog modela sa belim šumom

Prosečni model nije ni približnog izgleda kao početni. Ovakve mešovite modele bi bilo teško prepoznati. Mogu se pogledati i grafici prvih par simulacija.

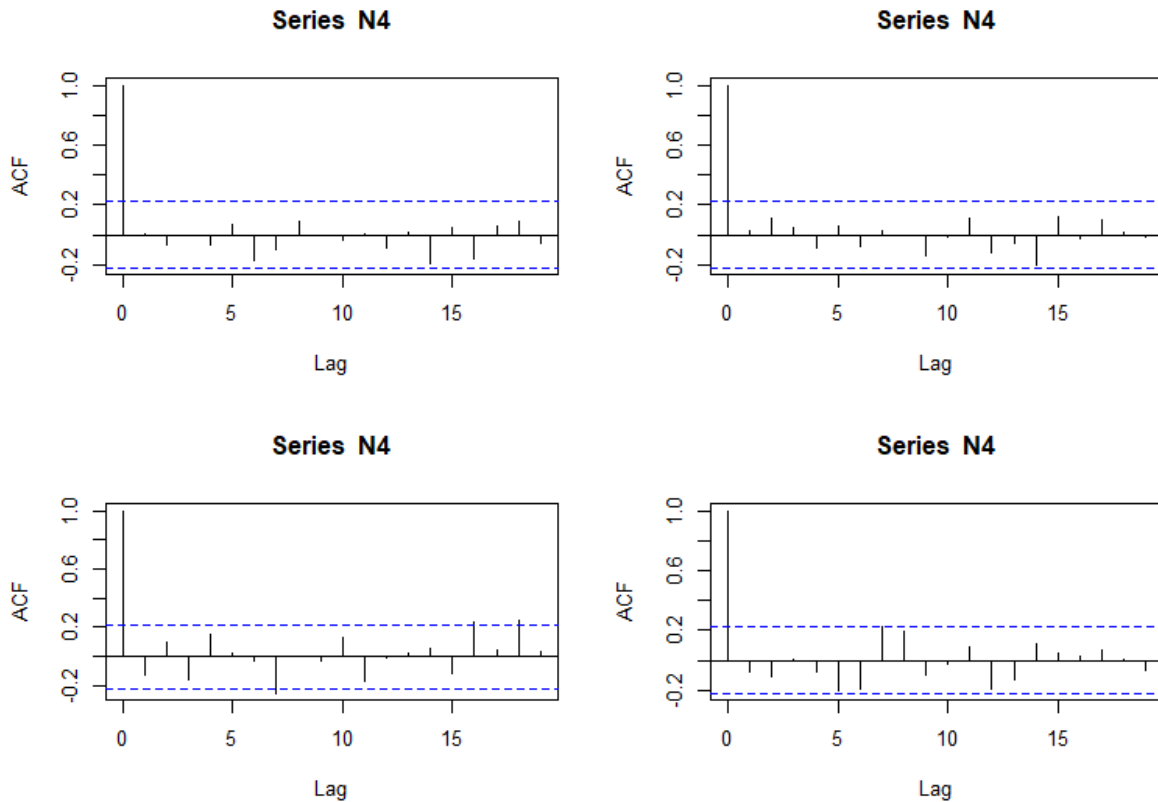
```
> par(mfrow=c(2,2))
> #ponavlja se 4 puta naredni kod
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(-6,-2,2,6),times=20)
> xt <- 1 : 80
> t<- 1 + 2 *xt
> N4<-t*e+s
> plot.ts(N4, col=3, xlab="Vreme", ylab=" ", main="Mešoviti model sa belim
šumom")
```



Slika 6.3.9. Grafički prikaz simuliranih mešovitih modela sa belim šumom

Takođe se može pogledati i grafički uzorački korelogram za takve modele.

```
> #ponavlja se 4 puta naredni kod
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(-6,-2,2,6),times=20)
> xt <- 1 : 80
> t<- 1 + 2 *xt
> N4<-t*e+s
> acf(N4)
```

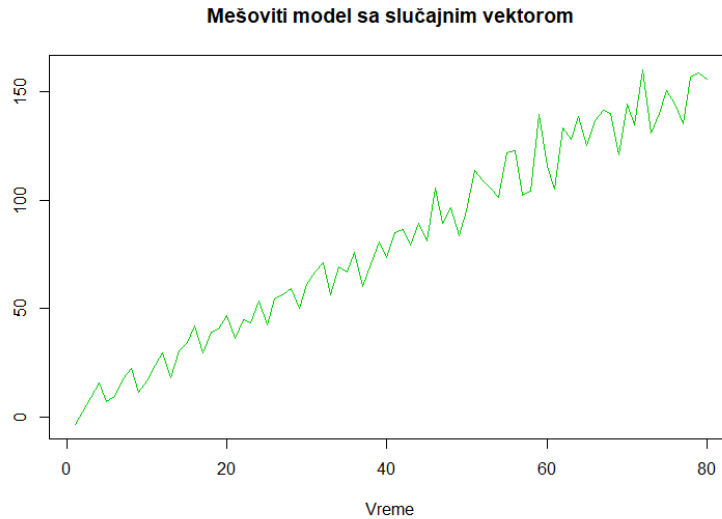


Slika 6.3.10. Grafički uzorački korelogram simuliranih mešovutih modela

Grafički uzorački korelogram simuliranih mešovutih modela je raznovrstan i na osnovu njega se ne bi mogao dobiti neki zaključak.

Umesto belog šuma može se kreirati model sa slučajnim vektorom r sa istom kombinacijom komponenti.

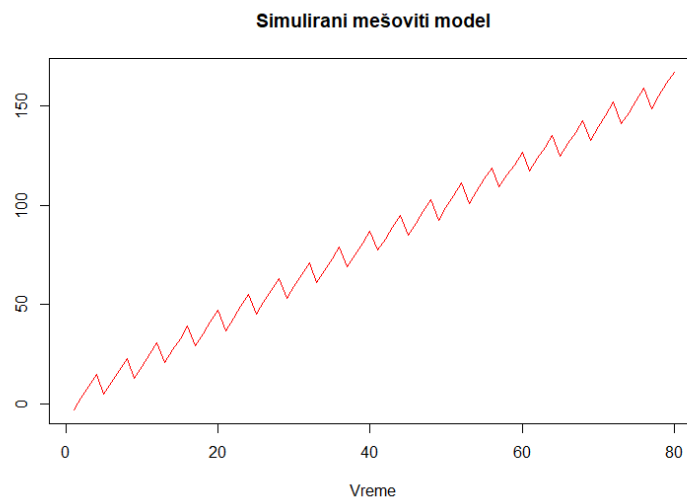
```
> par(mfrow=c(1,1))
> r <- rnorm(80,mean=1,sd=0.08)
> s<-rep(c(-6,-2,2,6),times=20)
> xt <- 1 : 80
> t<- 1 + 2 *xt
> N5<-t*r+s
> plot.ts(N5, col=3, xlab="vreme", ylab=" ", main="Mešoviti model sa slučajnim vektorom")
```

Slika 6.3.11. Grafički prikaz mešovitog modela sa slučajnim vektorom

Radi se i simulacija modela na već poznat način.

```
> mešoviti_model4 <- function(){
+   r <- rnorm(80,mean=1,sd=0.08)
+   s<-rep(c(-6,-2,2,6),times=20)
+   xt <- 1 : 80
+   t<- 1 + 2 *xt
+   N5<-t*r+s
+ }
> v_simulacija <- replicate(1000, mešoviti_model4())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
> plot.ts(X,col=2, xlab="vreme", ylab=" ", main="Simulirani mešoviti model"
)
```

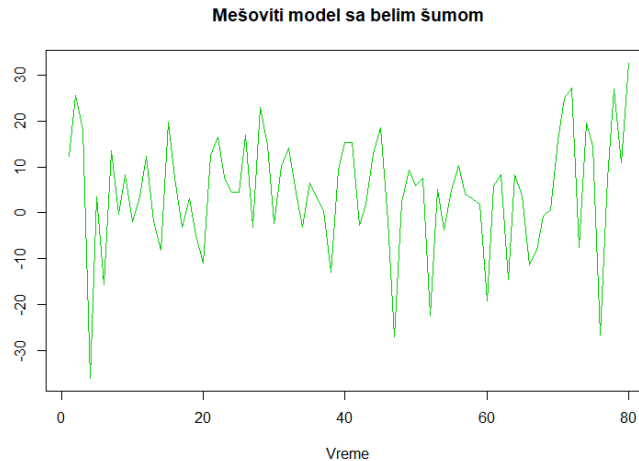


Slika 6.3.12. Grafički prikaz simuliranog mešovitog modela sa slučajnim vektorom

Ovakav model je skoro isti kao aditivni model, jer se sa grafika vidi stabilna sezonalnost. Dakle ovakav mešoviti model se može pomešati sa aditivnim.

I, na kraju, može se kreirati mešoviti model koji množi sezonsku i slučajnu komponentu i dodaje trend. Njegova formula je $X_t = S_t \cdot e_t + f_t$.

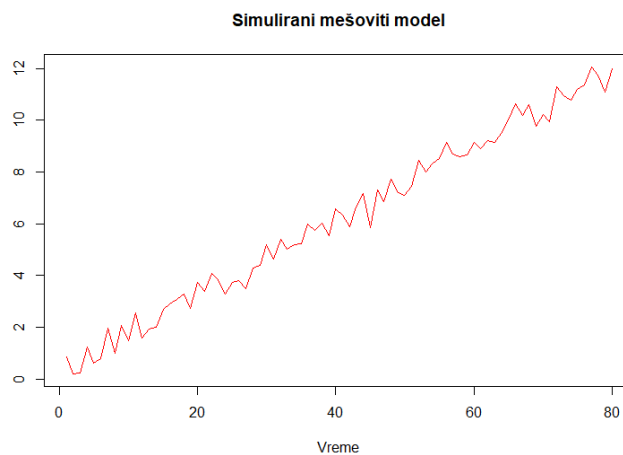
```
> e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
> s<-rep(c(-10,-12,12,14),times=20)
> t<-seq(0.5, 12, length.out=80)
> N6<-s*e+t
> plot.ts(N6, col=3, xlab="Vreme", ylab=" ", main="Mešoviti model sa belim
šumom")
```



Slika 6.3.13. Grafički prikaz mešovitog modela sa belim šumom

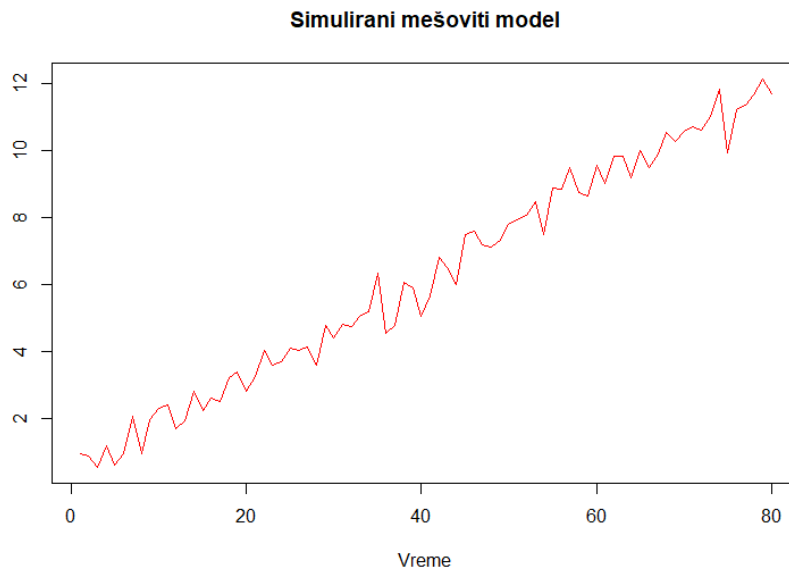
Radi se simulacija ovakvog modela.

```
> mešoviti_model5 <- function(){
+   e <- arima.sim(model = list(order = c(0, 0, 0)), n = 80)
+   s<-rep(c(-10,-12,12,14),times=20)
+   t<-seq(0.5, 12, length.out=80)
+   N6<-s*e+t
+ }
> v_simulacija <- replicate(1000, mešoviti_model5())
> X<-seq(0,0, length.out=80)
> Y<-seq(0,0, length.out=1000)
> for (j in 1:80) {
+   for (i in 1:1000) {
+     Y[i]<-v_simulacija[[j,i]]
+   }
+   X[j]<-sum(Y)/1000
+ }
> plot.ts(X,col=2, main="Simulirani mešoviti model")
```



Slika 6.3.14. Grafički prikaz simuliranog mešovitog modela sa belim šumom

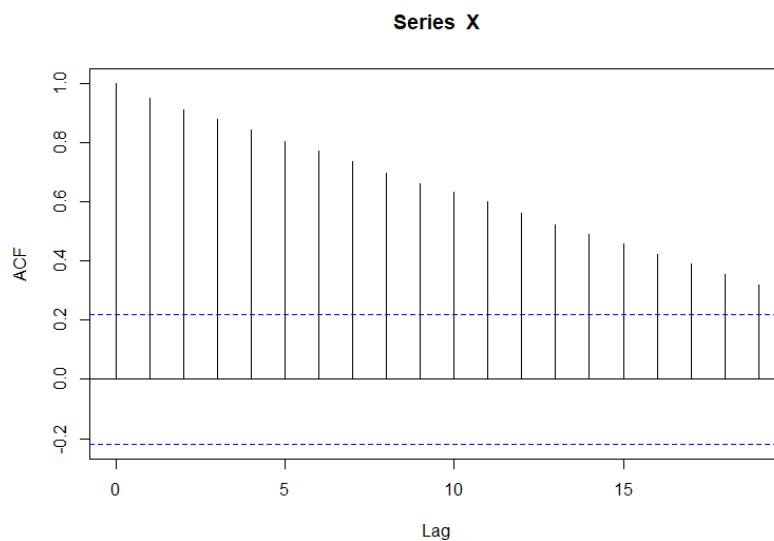
Sa grafika deluje da se poneki sezonski obrasci ponavljaju i to bi možda bio znak da je ovo neki mešoviti model i po tome bi se prepoznao. Simulacija se takođe može više puta ponoviti i dobiće se slični simulirani modeli:



Slika 6.3.15. Grafički prikaz simuliranog mešovitog modela sa belim šumom

Može se prikazati i grafički uzorački korelogram ovakvog modela.

> `acf(X)`



Slika 6.3.16. Grafički uzorački korelogram simuliranog mešovitog modela

Grafički uzorački korelogram simuliranog mešovitog modela je sličan kao kod simuliranog aditivnog modela u primeru 6.1.

7. ZAKLJUČAK

U ovom radu detaljno su opisani modeli vremenskih serija. Fokus je bio na multiplikativnim modelima, a takođe su opisani i aditivni i mešoviti modeli.

Kako bi se odabrao odgovarajući model vremenske serije, prvo je potrebno ispitati originalnu vremensku seriju. Nekada se odmah na osnovu grafika može videti koji model je u pitanju. Kada je sezonska varijacija relativno stabilna tokom vremena tada se bira aditivni model. Ako se sezonska varijacija npr. povećava tokom vremena onda je potrebno odabrati multiplikativni model. Međutim, često se ispostavi da ako je prvobitni zaključak da se može koristiti multiplikativni model, daljom analizom se dođe do toga da je ipak možda bolje primeniti neki mešoviti model.

Multiplikativni modeli su nepogodni ako su vrednosti nekih komponenti bliske nuli ili ako imaju negativan predznak, pa je u tom slučaju pogodniji mešoviti model.

Potrebno je i isprobati niz modela i izabrati onaj koji daje najstabilniju sezonsku komponentu, pogotovu na kraju serije.

Takođe, za datu vremensku seriju mogu se odrediti i autokorelacione i autokovarijacione funkcije, zato što one imaju specifične oblike za pojedinačne, različite modele, pa se prema tome isto može doći do zaključka o odabiru modela.

U poglavlju 5 kroz razne primere pokazano je kako se prema gore pomenutim kriterijumima bira koji model vremenske serije je najpogodniji. Korišćeni su razni skupovi podataka iz R baze, kao i jedna baza podataka preuzeta sa Interneta ([12]).

U poglavlju 6 prikazane su simulacije raznih modela. Ideja je bila krenuti sa druge strane, odnosno kreirati aditivne, multiplikativne i mešovite modele i prikazati kako izgledaju, kako bi takve vremenske serije bile lakše prepoznate. Zaključak je da se sa grafičkog prikaza ne može uvek jasno utvrditi koji model odabrati, ali može da bude dobra smernica za dalju analizu.

Multiplikativni modeli su značajni u raznim oblastima, te su razmatrani u mnogim naučno-istraživačkim radovima. U jednom od tih radova „*Daily Cash Forecasting with Multiplicative Models of Cash Flow Patterns*“ ([21]) multiplikativni modeli su primenjeni za dnevno predviđanje tokova novca, što predstavlja problem od velike praktične važnosti. Dobra predviđanja dnevnih tokova novca mogu poboljšati neto prinos za kratkoročna ulaganja, smanjiti troškove izdavanja vrednosnih papira i drugog kratkoročnog zaduživanja, smanjiti neaktivna stanja novca i generalno olakšati upravljanje novcem. Mnoge kompanije ulažu ozbiljne napore u predviđanju dnevnih tokova novca, pa je u ovom članku fokus na tome, kao i na razlozima zašto predviđanje nekad ne bude uspešno.

Glavni tokovi novca su definisani kao veliki tokovi koji nisu periodični i obično se ne mogu oceniti iz prošlosti, npr. porezi, dividende, isplate fondova, zakupi, isplate po završetku projekta, prodaja po ugovoru, prihod od raspolaganja imovinom... Kada bi se standardne ekstrapolativne, periodične ili auto-regresivne tehnike primenile na ovakve novčane tokove bez uklanjanja većih tokova, ne bi bile uspešne, jer glavni tokovi čine najveće tokove u danima kada se dogode i ne mogu se modelovati ni kao trend ni kao periodična funkcija. Na primer, plaćanja poreza na dohodak obično se dešavaju najviše šest puta godišnje, nemaju redovnu periodičnost i imaju iznose koji nemaju mnogo veze sa prošlim novčanim tokovima, uključujući čak i prošla plaćanja poreza. Čak i kad se svi veliki tokovi uklone iz novčanog toka, uglavnom je uvek potrebno podeliti preostali tok na komponente. Dnevni tokovi novca uvek imaju obrazac dan-u-nedelji i generalno takođe imaju obrazac dan-u-mesecu.

Mnogi događaji toka novca vezani su za mesečne ili nedeljne događaje, npr. mesečne ili nedeljne plate, uslovi kredita vezani za dan u mesecu, kao što je datum dospeća na kraju meseca i interne procedure obrade, kao što je priprema neke uplate određenim danima u nedelji itd.

Ovaj članak strukturira rešavanje efekata dana u nedelji i dana u mesecu formulišući nekoliko multiplikativnih modela. Takođe, poredi aditivne modele efekata dana u nedelji i dana u mesecu u trenutnoj literaturi, i zaključuje da su neki multiplikativni modeli ipak pogodniji.

U radu „*Additive and Multiplicative Models for Relative Survival Rates*“ ([22]) razmatra se i aditivni i multiplikativni model za relativnu stopu preživljavanja. Relativna stopa preživljavanja je odnos ukupne stope preživljavanja prema onoj „očekivanoj“ za demografski slične osobe u referentnoj populaciji. Obično se koristi za procenu uticaja određene bolesti na mortalitet, kada uzrok smrti nije pouzdano poznat. Efekat bolesti može biti multiplikativan ili aditivan. Pretpostavlja se da je efekat bolesti konstantan tokom čitavog perioda praćenja, ili da je konstantan u intervalima praćenja K . Prikazane su i razne statistike zasnovane na ocenama momenta efekta bolesti.

U radu „*Multiplicative Models for Interaction in Three-Way ANOVA, with Applications to Plant Breeding*“ ([23]) prikazana je primena multiplikativnih modela u svrhu bioloških tumačenja. U uzgajivanju biljaka, multiplikativni modeli za dvosmernu analizu interakcije disperzije postali su opšte sredstvo za opisivanje interakcije genotipa po okruženju. Nedostatak preovlađujuće dominacije dvosmernih multiplikativnih modela je to što su podaci sa složenijom strukturom životne sredine često primorani da se uklapaju u dvosmerni okvir. Kao delimično rešenje ovog problema, predstavljeni su trosmerni multiplikativni modeli koji se mogu koristiti kao dodatak poznatijim dvosmernim multiplikativnim modelima.

Dakle, multiplikativni modeli su bitni za mnoga praktična istraživanja u raznim oblastima, te su stoga jako važni u analizi vremenskih serija.

8. LITERATURA

- [1] Jovan Mališić, Vesna Jevremović: *Slučajni procesi i vremenske serije*, Matematički fakultet, Beograd 2008.
- [2] Zorica Mladenović, Aleksandra Nojković: *Primenjena analiza vremenskih serija*, Centar za izdavačku delatnost ekonomskog fakulteta u Beogradu, 2012.
- [3] Zlatko J. Kovačić: *Analiza vremenskih serija*, Ekonomski fakultet, Univerzitet u Beogradu, 1995.
- [4] <https://docplayer.gr/56030661-13-glava-analiza-vremenskih-serija.html>
- [5] Information paper: *An Introductory Course on Time Series Analysis* – Electronic Delivery, Australian Bureau of Statistics, 2005.
- [6] Ruey S. Tsay: *Analysis of Financial Time Series*, Third Edition, The University of Chicago Booth School of Business, 2010.
- [7] <https://newonlinecourses.science.psu.edu/stat510/lesson/5/5.1>
- [8] <http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/src/timeseries.html>
- [9] <https://www.udemy.com/course/time-series-analysis-and-forecasting-in-r/>
- [10] <https://www.rdocumentation.org>
- [11] <https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/JohnsonJohnson.html>
- [12] <https://www.kaggle.com/census/e-commerce-retail-sales-series-data-collection>
- [13] <https://nwfsc-timeseries.github.io/atsa-labs/sec-tslab-decomposition-of-time-series.html>
- [14] <https://financetrain.com/simulate-white-noise-wn-in-r/>
- [15] http://hikom.grf.bg.ac.rs/stari-sajt/web_stranice/KatZaHidr/Predmeti/StohHidrol/Elaborat2006.pdf
- [16] http://web.vu.lt/mif/a.buteikis/wp-content/uploads/2019/02/Lecture_03.pdf
- [17] https://rstudio-pubs-static.s3.amazonaws.com/387566_e7b840b30a934acc8acccc76a62f95d5.html
- [18] <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/box.test.html>
- [19] <https://rpubs.com/newajay/autoarima>
- [20] <https://rpubs.com/shichich/616110>
- [21] Bernell K. Stone, Tom W. Miller: *Daily Cash Forecasting with Multiplicative Models of Cash Flow Patterns*, Financial Management Vol. 16, Br. 4 (Winter, 1987), str. 45-54 (10 stranica)
- [22] J. D. Buckley: *Additive and Multiplicative Models for Relative Survival Rates*, Biometrics Vol. 40, Br. 1 (Mar., 1984), str. 51-62 (12 stranica)
- [23] Fred A. van Eeuwijk, Pieter M. Kroonenberg: *Multiplicative Models for Interaction in Three-Way ANOVA, with Applications to Plant Breeding*, Biometrics Vol. 54, Br. 4 (Dec., 1998), str. 1315-1333 (19 stranica)

BIOGRAFIJA

Sonja Jovanović rođena je 31.7.1992. godine u Užicu. Završila je Užičku gimnaziju, društveno- jezički smer i nosilac je Vukove diplome. Osnovne studije na Matematičkom fakultetu Univerziteta u Beogradu, smer Statistika, aktuarska i finansijska matematika završila je 2016. godine.

Radno iskustvo započela je 2016. godine u firmi Comtrade System Integration, u kojoj i sada radi. Njen tim se bavi analitikom podataka i poslovnom inteligencijom, a trenutno ima zvanje Business Consultant-a. Radila je na projektima implementacije Data warehouse-a (centralno skladište podataka) u raznim kompanijama privatnog i državnog sektora, kao i u industriji telekomunikacije. Kroz razne alate velikih vendora kao što su Oracle, IBM i Microsoft usavršila je znanje računarskog jezika SQL, a sastavni deo posla je i redovna komunikacija i sastanci sa klijentima. Tokom radne karijere prisustvovala je i učestvovala na mnogim konferencijama, poput Data Science Conference.