

3. DOMAĆA ZADAĆA

STATISTIČKI PRAKTIKUM 2

Zadatak 5. Zarada u slastičarnici

U datoteci `slasticarnica.csv` nalaze se podaci o dnevnoj zaradi jedne slastičarnice (u tisućama kuna) u razdoblju od 2015. do kraja 2017. godine. Analizirajte podatke na 3 načina.

1. Prilagodite linearni model danim podacima.
 - (a) Grafički i inferencijalno provjerite pretpostavke tog linearnog modela. Po potrebi izmijenite model (napravite potrebne transformacije) kako bi bolje odgovarao danim podacima.
 - *Napomena:* Kao test koreliranosti ostataka možete primijeniti Durbin-Watsonov test iz paketa `lmtest`, a za homogenost ostataka Jarque-Berov test.
2. Prilagodite GLS linearnu regresiju danim podacima (za koreliranost ostataka uzmite u obzir $AR(1)$ strukturu, kao na vježbama).
3. Prilagodite podacima pogodan ARIMA model (možete koristiti gotovu naredbu `stl`).

U sva tri slučaja procijenite mjesečnu prodaju za siječanj 2018. godine zajedno s 95% pouzdanim intervalom za opaženu vrijednost i usporedite sa stvarnim podacima o prodaji za taj mjesec koji su dani u datoteci `slasticarnica_siječanj.csv`. Usporedite sve procjene.

(15 bodova)

Zadatak 6. Zaraza koronavirusom

U datoteci `korona_hrv2022.xlsx` se nalaze podaci o dnevnom broju novih slučajeva zaraze koronavirusom u Hrvatskoj u razdoblju od 1. siječnja do 31. listopada 2022. godine.

- (a) Grafički prikažite podatke. Što možete reći o homogenosti podataka na temelju ovog grafa?
- (b) Napravite potrebne transformacije kako bi podaci odgovarali našem modelu za vremenske nizove te analizirajte podatke (uz odgovarajuće grafičke prikaze).
- (c) Procijenite broj novoizaraženih osoba u idućih 17 dana mjeseca studenog i odredite 95% pouzdani interval za procjenu. Usporedite dobivenu procjenu sa stvarnim podacima o novoizaraženim osobama koji se nalaze u datoteci `korona_hrv2022_studeni.xlsx`.

(10 bodova)

Zadatak 7. Podaci o proizvodnji električne opreme

U datoteci `electric.tx` nalaze se podaci o mjesečnoj proizvodnji električne opreme za 17 europskih zemalja u razdoblju od siječnja 1996. do veljače 2011. godine.

Podatke o proizvodnji prikažite grafički i zatim procijenite sezonalnost nekom od neimplementiranih metoda (tj. bez korištenja procedure `stl`). Grafički usporedite tu procjenu s procjenom sezonalnosti iz procedure `stl`.

Ostatak niza analizirajte na dva načina:

- (a) Predložite pogodan model za trend i dodajte ga na graf podataka. Analizirajte ostatke. Radi li se o bijelom šumu? Ukoliko ne, predložite pogodan model za ostatke.
- (b) Analizirajte podatke (bez procjene trenda) i pronađite pogodan ARIMA model za njih.

Ispitajte podobnost dobivenog modela u oba slučaja.

U oba slučaja procijenite mjesečnu proizvodnju do kraja 2011. godine zajedno s 95% pouzdanim intervalom i usporedite sa stvarnim podacima o proizvodnji za to razdoblje koji su dani u datoteci `electric_new.txt`. Usporedite obje procjene.

(15 bodova)