

STATISTIKA

DRUGI KOLOKVIJ

Napomena:

Prije početka rješavanja kolokvija instalirajte i učitajte paket BSDA

```
install.packages("BSDA")
```

```
library(BSDA)
```

Rješenja trebate kopirati u Word dokument. Komentare možete pisati u isti dokument ili na papir. Po završetku pisanja, trebate poslati dokument na

`statistika@mathos.hr`

i u predmetu poruke napisati svoje ime i prezime.

ZADATAK 1: [4+6+6+6+8=30 bodova]

U bazi podataka `Rat` (iz paketa `BSDA`) u stupcu `survival.time` nalaze se podaci o životnom vijeku štakora koji su izloženi velikim dozama radijacije.

- Nacrtajte box-plot. Ukazuje li box-plot na postojanje stršućih vrijednosti?
- Nacrtajte histogram, QQ-plot za normalnu distribuciju, te graf uzoračke funkcije gustoće usporedno s grafom gustoće normalne distribucije čije parametre će te procijeniti. Komentirajte možemo li na osnovu svake od tih slika naslutiti da je obilježje normalno distribuirano. Potvrdite to odgovarajućim testom.
- Nadite 99% interval pouzdanosti za očekivani životni vijek štakora. Možemo li na nivou značajnosti 0.01 tvrditi da će štakori živjeti dulje od 80 dana.
- Nadite jednostrani pouzdani interval za varijancu (95%). Testirajte je li varijanca različita od 1280 na razini značajnosti 0.05.
- Provjerite je li očekivanje različito od 113 na nivou značajnosti 0.05. Sad testirajte dolazi li ovaj uzorak iz $\mathcal{N}(113, 1280)$ distribucije (1280 je varijanca).

ZADATAK 2: [5+5+5=15 bodova]

U bazi podataka `Rehab` (iz paketa `BSDA`) nalaze se ocjene koje su dva psihijatra davali istim pacijentima koji su prošli kroz rehabilitaciju.

- Provjerite jesu li ocjene oba psihijatra normalno distribuirane.
- Nacrtajte usporedno boxplot dvije skupine podataka, različitom bojom.
- Razlikuju li se prosječne ocjene dvaju sudaca na nivou značajnosti 0.05?

ZADATAK 3: [10 bodova]

Potomstvo dobiveno križanjem dvije vrste biljaka se pojavljuje u tri genotipa označena s A, B i C. Teoretski model o genskom nasljeđivanju sugerira da bi se tipovi genotipa A, B i C trebali pojavljivati u omjeru 1 : 2 : 1. Izvedena je eksperimentalna provjera tog genetskog modela na 90 biljaka i dobiveni su sljedeći rezultati:

genotip	A	B	C
frekvencija	18	44	28

Jesu li ovi podaci u skladu s teoretskim modelom na nivou značajnosti 0.05?

ZADATAK 4: [10 bodova]

Jane Austen je preminula prije nego je završila svoj posljednji roman. Imitator je pokušao dovršiti njezin roman. Sljedećom tablicom prikazane su frekvencije nekoliko čestih riječi u engleskom jeziku u slučajno odabranom dijelu teksta koji je napisala Austen i imitator.

	a	an	this	that	with	without
Austen	434	62	86	236	161	38
Imitator	83	29	15	22	43	4

Testirajte jesu li učestalosti odabranih riječi jednake kod Austen i kod imitatora, tj. je li imitator uspio dobro iskopirati stil Jane Austen. ($\alpha = 0.05$)

ZADATAK 5: [15 bodova]

U jednoj društvenoj igri, kocka je bačena 144 puta i palo je ukupno 32 šestice. Je li kocka simetrična (na nivou značajnosti 0.05)?

ZADATAK 6: [10+10=20 bodova]

Neka je (X_1, \dots, X_n) jednostavan slučajni uzorak iz $\mathcal{E}(\lambda)$ distribucije

(a) Izvedite test statistiku i kritično područje na osnovu velikog uzorka za testiranje hipoteza

$$H_0 : \lambda = \lambda_0$$

$$H_1 : \lambda \neq \lambda_0$$

(sjetite se i da je $E[\mathcal{E}(\lambda)] = \frac{1}{\lambda}$ i $\text{Var}(\mathcal{E}(\lambda)) = \frac{1}{\lambda^2}$)

(b) Generirajte uzorak od 100 slučajnih brojeva iz $\mathcal{E}(5)$. Koristeći kreirani test, testirajte je li za taj uzorak parametar jednak 5 ili ne na nivou značajnosti 0.05.