

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: [2+8+6+6+3=25 bodova]

U bazi podataka **Undergrad** (iz paketa **BSDA**) u stupcu **SAT** nalaze se rezultati američkog ekvivalenta državne mature 100 studenata jednog sveučilišta.

- (a) Nacrtajte box-plot. Ukazuje li box-plot na postojanje stršćih vrijednosti?
- (b) 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.
- (c) Nađite 99% interval pouzdanosti za očekivani rezultat na testu. Možemo li na nivou značajnosti 0.01 tvrditi da su raspoloživi rezultati prosječno bolji od 1000? Koja je najmanja razina značajnosti u odnosu na koju bi to mogli tvrditi?
- (d) Nađite jednostrani pouzdani interval za varijancu (95%). Testirajte je li varijanca različita od 28000 na razini značajnosti 0.05.
- (e) Testirajte dolazi li ovaj uzorak iz $\mathcal{N}(1045, 28000)$ distribucije (28000 je varijanca).

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

Za podatke iz prvog zadatka:

- (a) U stupcu **Gender** za svaki podatak naznačeno je radi li se o studentu (`=="Male"`) ili studentici (`=="Female"`). Provjerite jesu li rezultati SAT testa normalno distribuirani za svaku od ove dvije skupine na nivou značajnosti 0.01. Prikažite ih usporedno boxplotom i komentirajte.
- (b) Jesu li rezultati studenata bolji od rezultata studentica na nivou značajnosti 0.01?
- (c) Možemo li na nivou značajnosti 0.01 tvrditi da su rezultati u nekoj od skupina manje varijabilni nego u drugoj.

ZADATAK 3: [15 bodova]

U Zagrebu je anketirano 500 ljudi i njih 385 se izjasnilo da je za povećanje dozvoljene brzine u gradu za 10km/h . U Osijeku je od 400 ljudi njih 267 bilo za. Je li podrška povećanju dozvoljene brzine veća u Zagrebu, na nivou značajnosti 0.05?

ZADATAK 4: [15 bodova]

Iz pošiljke se na slučajan način odabere 100 pakiranja po šest limenki piva. Za limenke iz svakog pakiranja provjeri se jesu li ispravno napunjene. Dobiveni su sljedeći podaci:

broj neispravno napunjenih limenki	0	1	2	3	4	5	6
broj pakiranja	55	33	9	4	0	0	0

Tvornica tvrdi da stroj koji puni limenke pogriješi u 8% slučajeva. Možemo li tvrditi da tvornica laže na nivou značajnosti 0.05?

ZADATAK 5: [10 bodova]

U bazi **Inmate** (iz paketa **BSDA**) nalaze se podaci o evidentiranim slučajevima zlorabe droga s obzirom na vrstu droge i rasu uhićene osobe. Testirajte jesu li navedena obilježja nezavisna. Je li konzumacija kokaina uniformno raspoređena među rasama? (sve na razini značajnosti 0.05)

ZADATAK 6: [7+6+7=20 bodova]

Neka je (X_1, \dots, X_n) jednostavan slučajan uzorak iz $\mathcal{U}(0, 2\theta)$ distribucije.

- (a) Izvedite pouzdani interval za parametar θ na osnovu velikog uzorka.
- (b) Izvedite test statistiku i kritično područje na osnovu velikog uzorka za testiranje hipoteza

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta > \theta_0$$

- (c) Generirajte uzorak od 100 slučajnih brojeva iz $\mathcal{U}(0, 4)$. Koristeći kreirani test, testirajte je li za taj uzorak parametar jednak 2 ili veći na nivou značajnosti 0.05. Izračunajte p-vrijednost za test na ovom uzorku.