

STATISTIČKI PRAKTIKUM

PISMENI ISPIT 29.6.2016.

Napomena:

Prije početka rješavanja pismenog ispita 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 e-mail adresu

`statistika@mathos.hr`

i u predmetu poruke napisati svoje ime i prezime.

ZADATAK 1: [3+4+3=10 bodova]Neka je (X_1, \dots, X_n) jednostavan slučajni uzorak iz binomne distribucije s parametrima $n \in \mathbb{N}$ i $p \in (0, 1)$.

- Interpretirajte centralni granični teorem u kontekstu j.s.u. (X_1, \dots, X_n) . (3 boda)
- Koristeći rezultat pod a), napišite i interpretirajte asimptotski pouzdani interval pouzdanosti γ za parametar p . (4 boda)
- Simulirajte uzorak veličine 100 iz binomne distribucije s parametrima $n = 10$ i $p = 0.2$ i izračunajte realizaciju 90%-tnog pouzdanog intervala za parametar p . (3 boda)

ZADATAK 2: [5+5+10+5+10+5+10+10+10=70 bodova]Baza podataka `Wastewat` iz paketa `BSDA` sadrži podatke o količini otpadnih voda u galonima koje su pod visokim tlakom i uz pomoć specijaliziranih bušilica injektirane u podnožje planine Rocky Mountain u SAD-u (varijabla `gallons`) te podatke o broju potresa na tom području u određenom vremenskom periodu nakon svakog injektiranja (varijabla `number`). Sve statističke testove trebate provesti na razini značajnosti $\alpha = 0.05$.

- Kakvim biste slučajnim varijablama modelirali količinu otpadnih voda (varijabla `gallons`) i broj potresa (varijabla `number`). Procijenite funkciju distribucije obiju slučajnih varijabli i nacrtajte njihove grafove.
- Za količinu otpadnih voda nacrtajte histogram podataka, QQ-graf za normalnu distribuciju, te graf procijenjene funkcije gustoće usporedno s grafom gustoće normalne distribucije s procijenjenim parametrima. Možemo li na osnovu svake od tih slika naslutiti da je količina otpadnih voda normalno distribuirana slučajna varijabla? Provjerite tu slutnju odgovarajućim statističkim testom.
- Procijenite očekivanu količinu otpadnih voda 97%-tnim pouzdanim intervalom te interpretirajte dobiveni rezultat.
- Možemo li tvrditi da je broj potresa koreliran s količinom otpadnih voda injektiranih u podnožje planine?
- Procijenite parametre linearnog modela koji opisuje ovisnost broja potresa o količini otpadnih voda. Jesu li koeficijenti modela statistički značajno različiti od 0? Nacrtajte procijenjeni regresijski pravac zajedno s uređenim parovima podataka. Prema dobivenom modelu, kako porast količine injektiranih otpadnih voda za 5 galona utječe na promjenu broja potresa?
- Prikladnim grafičkim prikazima provjerite imamo li razloga sumnjati u homogenost varijanci grešaka linearnog modela promatranog u zadatku (e).

Nadalje, promotrimo posebno mjerenja u kojima je, s obzirom na izmjerene vrijednosti, u podnožje planine injektirana ispodprosječna količina otpadnih voda te mjerenja u kojima je u podnožje planine injektirana iznadprosječna količina otpadnih voda.

- Razlikuje li se očekivani broj potresa u slučaju iznadprosječne i ispodprosječne količine otpadnih voda injektiranih u podnožje planine?

- (h) Je li broj potresa varijabilniji u slučaju iznadprosječne količine otpadnih voda injektiranih u podnožje planine?
- (i) Razlikuju li se distribucije broja potresa u slučaju iznadprosječne i ispodprosječne količine otpadnih voda injektiranih u podnožje planine?

ZADATAK 3: [10 bodova]

British Medical Journal objavio je članak u kojemu se proučava povezanost hrkanja i srčanih oboljenja. Ispitanici su kategorizirani s obzirom na "učestalost hrkanja" (ne hrče, povremeno hrče, hrče gotovo svaku noć, hrče svaku noć) te s obzirom na to imaju li neku bolest srca ili obiteljsku povijest bolesti srca (kategorija 1) ili nemaju (kategorija 0). Rezultati istraživanja prikazani su u sljedećoj tablici:

| | Ne hrče | Povremeno hrče | Hrče gotovo svaku noć | Hrče svaku noć |
|---|---------|----------------|-----------------------|----------------|
| 1 | 24 | 35 | 21 | 30 |
| 0 | 1355 | 603 | 192 | 224 |

Možemo li na nivou značajnosti 0.05 tvrditi da su hrkanje i oboljenje srca zavisna obilježja?

ZADATAK 4: [10 bodova]

U sklopu jednog istraživanja proučavana je uspješnost učenika u rješavanju matematičkih zadataka zadanih riječima. Utvrđeno je da je od 53 djevojčice njih 17 zadovoljilo kriterije i mogu se kategorizirati kao uspješne u rješavanju zadataka riječima, dok je od 48 dječaka njih 15 svrstano među uspješne. Možemo li na razini značajnosti $\alpha = 0.05$ tvrditi da se djevojčice bolje snalaze u rješavanju zadataka riječima?