

STATISTIČKI PRAKTIKUM

DRUGI KOLOKVIJ

Napomena:

- Za sve zadatke popratni tekst piše se ili na papir ili u R skripti u obliku komentara. Konačni dokument rješenja zadataka u R-u generira se na "File → Knit Notebook → Word". Taj se dokument šalje na `statprak@mathos.hr` zajedno s R skriptom. U predmetu poruke treba napisati svoje ime i prezime.

ZADATAK 1: [2+2=4 bodova]

- Definirajte što je rang podatka iz uzorka i odredite rangove za podatke iz uzorka 5, 2, 7, 1.
- Definirajte procjene za uvjetne distribucije dvodimenzionalnog diskretnog slučajnog vektora na osnovu zajedničke tablice frekvencija.

ZADATAK 2: [6+7+4=17 bodova]

Instalirajte paket PASWR naredbom `install.packages("PASWR")` u konzoli (ne u skripti iz koje se generira izvješće). U tablici `titanic3` iz tog paketa nalaze se podaci o putnicima s Titanika, između ostalog:

- varijabla `pclass` označava klasu kojom je putnik putovao
 - varijabla `survived` je 1 ako je putnik preživio, a 0 ako nije
 - varijabla `sex` je "male" ako je putnik muškarac, a "female" ako je žena
 - varijabla `age` je dob putnika
 - varijabla `sibsp` je broj braće, sestara i supružnika s kojima je putnik putovao
 - varijabla `parch` je broj roditelja i djece s kojima je putnik putovao
- Izračunajte udio preživjelih posebno u prvoj, drugoj i trećoj klasi. Ovisi li preživljenje o klasi na razini značajnosti 0.05? Ako bi preživljenje bilo nezavisno o klasi, kolika bi bila očekivana frekvencija preživjelih u prvoj klasi?
 - Razlikuju li se na razini značajnosti 0.05 očekivane dobi preživjelih i stradalih? Usporedite podatke grafički. Možemo li tvrditi da se distribucije dobi razlikuju na razini značajnosti 0.05? (Za brojne putnike podatak o dobi je nepoznat, takve vrijednosti treba zanemariti što će većina funkcija prešutno i napraviti.)
 - Koristeći varijable `sibsp` i `parch` putnike možemo podijeliti na one koji su putovali sami i one koji su putovali s nekim. Na razini značajnosti 0.05, je li proporcija preživjelih veća kod onih koji su putovali s nekim?

ZADATAK 3: [4+4+3+6=17 bodova]

Instalirajte paket `Flury` naredbom `install.packages("Flury")` u konzoli (ne u skripti iz koje se generira izvješće). Nakon učitavanja paketa (`library("Flury")`), učitajte tablicu `m.twins` naredbom `data(m.twins)`. Tablica sadrži neke tjelesne mjere za uzorak blizanaca.

- Postoji li na razini značajnosti 0.05 razlika u očekivanoj visini (varijable `STA1` i `STA2`) između jednojajčanih blizanaca (varijabla `Type` jednaka "`monozygotic`")
- Usporedite koliko dobro linearna veza opisuje odnos između visine blizanaca u slučaju jednojajčanih i u slučaju dvojajčanih (varijabla `Type` jednaka "`dizygotic`") blizanaca.
- Možemo li tvrditi da su visine jednojajčanih blizanaca korelirane na razini značajnosti 0.05? Jesu li i zavisne?
- Procijenite parametre linearnog modela koji modelira visinu jednog blizanca na osnovu visine drugog posebno za jednojajčane i dvojajčane blizance. Usporedite kvalitetu dobivenih modela. Grafički prikazite dva skupa podataka različitim bojama te pripadne regresijske pravce.

ZADATAK 4: [6 bodova]

U svom istraživanju o genetskom nasljeđivanju, Mendel je postavio, između ostalog, hipotezu da križanjem graška sa žutim okruglim zrnom i graška sa zelenim naboranim zrnom, u drugoj generaciji dobije omjer biljaka 9:3:3:1 za redom žuti okrugli grašak, žuti naborani, zeleni okrugli, zeleni naborani.. Na osnovu posađenih stabljika, eksperimentom je dobio sljedeći broj biljki u drugoj generaciji

žuti okrugli	žuti naborani	zeleni okrugli	zeleni naborani
315	108	101	32

Ima li razloga sumnjati u Mendelovu hipotezu na razini značajnosti 0.05?

ZADATAK 5: [6 bodova]

Ispitanicima su postavljena dva pitanja: vjeruju li u raj i vjeruju li u pakao. Njih 95 vjeruje i u raj i u pakao, njih 29 ni u jedno ni u drugo, 19 ih vjeruje samo u raj, a 5 samo u pakao. Vjeruju li ljudi više u raj nego u pakao na razini značajnosti 0.05?