

STATISTIČKI PRAKTIKUM

PISMENI ISPIT 13.7.2016.

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: [5+5+5=15 bodova]

- Definirajte pogreške statističkog testa i interpretirajte razinu značajnosti statističkog testa.
- Kako interpretiramo vrijednosti -1 , 0 i 1 Pearsonovog, a kako Spearmanovog koeficijenta korelacije?
- Navedite klasične pretpostavke linearnog regresijskog modela.

ZADATAK 2: [10+10+10+10+10+10+5=65 bodova]

U R bazi podataka `Rainks` nalaze se prosječne izmjerene količine padalina u inčima ($1 \text{ inč} = 2.54 \text{ cm}$) za 35 uzastopnih godina za zapadni dio američke savezne države Kansas (varijabla `rain`) te za četiri susjedna područja (redom varijable `x1`, `x2`, `x3` i `x4`). Sve statističke testove trebate provesti na razini značajnosti $\alpha = 0.05$.

- Procijenite očekivanu prosječnu godišnju količinu padalina u zapadnom Kansasu 90%-tnim intervalom pouzdanosti te interpretirajte dobiveni rezultat. Budući da je Kansas česta meta tornada popraćenih jakim i obilnim kišama, savezna država Kansas odlučila je odobriti stalnu godišnju financijsku pomoć stanovnicima zapadnog dijela ukoliko se pokaže da je očekivana prosječna godišnja količina padalina veća od 3. Hoće li financijska pomoć biti odobrena ili će Dorothy i Toto biti prepušteni na milost i nemilost Čarobnjaka iz Oz-a?
- Odredite jednostrani 97%-tni pouzdani interval za varijancu prosječne godišnje količine padalina u zapadnom Kansasu te ga interpretirajte. Zbog promjene vlasti na izborima, došlo je i do promjene kriterija za dodjelu financijske pomoći - pomoć će biti odobrena ukoliko se pokaže da je varijanca prosječnih godišnjih padalina veća od 0.5. Hoće li nova vlast odobriti financijsku pomoć?
- Procijenite funkciju distribucije procječne godišnje količine padalina u području `x2`. Testirajte radi li se o normalno distribuiranom obilježju?
- Testirajte razlikuju li se očekivane prosječne godišnje količine padalina u zapadnom Kansasu i području `x2`? Je li prosječna godišnja količina padalina u području `x2` varijabilnija nego u zapadnom Kansasu?
- Procijenite koeficijente korelacije između prosječne godišnje količine padalina u zapadnom Kansasu i područjima `x1`, `x2`, `x3` i `x4`? Možemo li tvrditi da je prosječna godišnja količina padalina u zapadnom Kansasu korelirana s padalinama u nekom od područja `x1`, `x2`, `x3` i `x4`? Koja od varijabli `x1`, `x2`, `x3` i `x4` je najbolji prediktor prosječne godišnje količine padalina u zapadnom Kansasu?
- Procijenite parametre linearnog modela prosječne godišnje količine padalina u zapadnom Kansasu u ovisnosti o najboljem prediktoru detektiranom u prethodnom zadatku. Jesu li koeficijenti modela statistički značajno različiti od nule? Nacrtajte procijenjeni regresijski pravac zajedno s uređenim parovima podataka. Prema dobivenom modelu, kako porast prosječne godišnje količine padalina u području `x2` za 5 inča utječe na promjenu prosječne godišnje količine padalina u zapadnom Kansasu?
- Provjerite imamo li razloga sumnjati u nezavisnost grešaka modela iz prethodnog zadatka te obrazložite svoj zaključak.

ZADATAK 3: [10 bodova]

Pitamo se je li školski uspjeh učenika (kategoriziran u kategorije **ispodprosječan**, **prosječan** i **iznadprosječan**) u vezi sa sudjelovanjem učenika u izvannastavnim projektima (kategorija **Da** znači da učenik sudjeluje, a kategorija **Ne** da učenik ne sudjeluje u izvannastavnim projektima). Rezultati istraživanja na uzorku od 80 učenika trećeg razreda dani su u sljedećoj tablici:

	Da	Ne
Ispodprosječan uspjeh	17	9
Prosječan uspjeh	14	13
Iznadprosječan uspjeh	22	5

Možemo li na razini značajnosti $\alpha = 0.05$ tvrditi da su školski uspjeh učenika i njihovo sudjelovanje u izvannastavnim projektima zavisna obilježja?

ZADATAK 4: [10 bodova]

Anketom provedenom na uzorku od 50 ispitanika iz gradske populacije i 50 ispitanika iz seoske populacije želi se ispitati stav vezan uz jedno referendurno pitanje (jesu li "za" ili "protiv" tvrdnje sadržane u referendurnom pitanju). Rezultati ankete dani su u sljedećoj tablici:

	Za	Protiv
grad A	35	15
grad B	47	3

Razlikuju li se na razini značajnosti $\alpha = 0.05$ proporcije ispitanika koji su se izjasnili protiv tvrdnje sadržane u referendurnom pitanju (kategorija **Protiv**) s obzirom na to dolaze li ispitanici iz grada ili sa sela?