

**STATISTIČKI PRAKTIKUM**

## DRUGI KOLOKVIJ

**Napomena:**

- Rješenja teorijskih zadataka (Zadatak 1.) pišu se na papir koji se predaje predmetnom nastavniku.
- Kod zadataka koji se rješavaju u R-u sav 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 → Compile Notebook → Word". Taj se dokument šalje na `statistika@mathos.hr` - u predmetu poruke treba napisati svoje ime i prezime.
- Prije početka rješavanja kolokvija učitajte s `library` (i po potrebi instalirajte s `install.packages`) pakete `MASS`, `car` i `BSDA`.
- Sve testove treba provesti na razini značajnosti 0.05.

**ZADATAK 1:** [3+3+2=8 boda]

Odgovorite na sljedeća pitanja:

- Ako u uzorku veličine četiri koji dolazi iz jedne distribucije promatramo dva podatka, odredite distribuciju sume rangova ta dva podatka. Pretpostavite da su svi podaci različiti.
- Navedite i objasnite koje su moguće jednostrane hipoteze u MWW testu.
- Objasnite kako se test o usporedbi očekivanja za vezane uzorke svodi na odgovarajući test na jednom uzorku.

**ZADATAK 2:** [4+4+4+5+4+4+2=27 bodova]

Pima Indijanci imaju najveću rasprostranjenost dijabetesa od svih drugih naroda. Baza podataka `Pima.tr` iz paketa `MASS` sadrži neke podatke o ženama iz ove populacije. Varijabla `type` označava ima li osoba dijabetes (`Yes`) ili ne (`No`).

- Imaju li ispitanici s dijabetesom veći očekivani indeks tjelesne mase (BMI – varijabla `bmi`) u odnosu na one koji nemaju dijabetes?
- Možemo li tvrditi da je proporcija žena koje nisu bile trudne veća među oboljelima od dijabetesa? Varijabla `npreg` označava broj trudnoća.
- Možemo li tvrditi da su žene koje su bile trudne starije od onih koje nisu bile trudne (varijabla `age`)? Odaberite prikladan test s obzirom na veličine uzoraka.
- Prema jednom kriteriju, osoba je pretiła ako ima BMI veći od 30. Pretilost se može dijagnosticirati i ako je debljina nabora kože iznad tricepsa (varijabla `skin`) veća od 20. Razlikuju li se proporcije pretiilih po ova dva kriterija?
- Možemo li tvrditi da veća debljina nabora kože iznad tricepsa ukazuje na veći BMI?
- Procijenite parametre odgovarajućeg linearnog model kojim bi mogli procijeniti BMI na osnovu debljine kože iznad tricepsa. Što možete reći o jakosti modela? Procijenite BMI za osobu čija je koža iznad tricepsa debela točno 17 mm.
- Provjerite ima li razloga sumnjati u homogenost varijanci grešaka modela dobivenog u (f). Obrazložite.

ZADATAK 3: [5 bodova]

Novčić bacamo dok ne padne pismo, ali najviše četiri puta. Ponavljanjem ovog pokusa 100 puta utvrđeno je da je 49 puta novčić bačen jednom, 13 puta je bačen dva puta, 11 puta je bačen tri puta i 27 puta je bačen četiri puta. Možemo li tvrditi da je novčić nepravilan?

ZADATAK 4: [5 bodova]

U R bazi `Burt` iz paketa `car` nalaze se podaci o rezultatima IQ testa za 27 parova blizanaca koji su po rođenju odvojeni. Jedan od blizanaca odrastao je s biološkim roditeljima (varijabla `IQbio`), a drugi u udomiteljskoj obitelji (varijable `IQfoster`). Provjerite ima li razloga sumnjati u normalnu distribuiranost razlika rezultata IQ testova. Koristeći taj rezultat, odgovorite ima li dokaza da okolina utječe na očekivanu inteligenciju mjerenu IQ testovima. (Napomena: postoje opravdane sumnje da je Sir Cyril Burt izmislio neke podatke u svojoj studiji.)

ZADATAK 5: [5 bodova]

U tablici `Seatbelt` is paketa `BSDA` zapisani su podaci o ozljedama u prometnim nesrećama (redom: nikakve, minimalne, lakše i ozbiljne) i podatak o tome je li osoba koristila pojas. Ovisi li težina ozljeda o korištenju pojasa?