

STATISTIČKI PRAKTIKUM

PISMENI ISPIT 24.2.2017.

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 učitajte s `library` (i po potrebi instalirajte s `install.packages`) paket `catdata`.
- Sve testove treba provesti na razini značajnosti 0.05.

ZADATAK 1: [5+5+5=15 bodova]

- Objasnite kako možemo testirati hipotezu o jednakosti proporcija u dva zavisna slučajna uzorka.
- Što je rang podatka x_i u nizu (x_1, \dots, x_n) u kojem su svi elementi međusobno različiti? Odredite rangove podataka u nizu (8.2, 5.4, 4.1, 2.2). Navedite barem jedan test koji se temelji na rangovima podataka.
- Navedite klasične pretpostavke regresijskog modela.

ZADATAK 2: [5+10+10+5+10+10+5+8+10+12=85 bodova]

U bazi podataka `birth` iz paketa `catdata` nalaze se podaci o 775 novorođenčadi (podatke učitati s `data(birth)` nakon instalacije i učitavanja paketa). Sve testove treba provesti na razini značajnosti 0.05.

- Varijabla `Weight` sadrži podatke o težini novorođenčadi u gramima. Procijenite vjerojatnost da je dijete teže od 4000 grama. Provjerite normalnu distribuiranost težine grafički i odgovarajućim testom.
- Varijabla `Height` sadrži podatke o visini (duljini) novorođenčadi u centimetrima. Odredite 95%-tni pouzdani interval za visinu. Je li očekivana visina veća od 50cm?
- Razlikuje li se očekivana težina novorođenčadi (varijabla `Weight`) ovisno o spolu (varijabla `Sex`, 1 - dječak, 2 - djevojčica)? Postoji li pozitivan pomak u distribuciji težine dječaka u odnosu na djevojčice?
- Imaju li djeca rođena carskim rezom (varijabla `Cesarean` jednaka 1) veći očekivani opseg glave (varijabla `Head`) od djece rođene prirodnim putem (varijabla `Cesarean` jednaka 0)? Pretpostavite da je opseg glave normalno distribuirano obilježje.
- Varijabla `Month` sadrži podatak o mjesecu u kojem je dijete rođeno. Je li svaki mjesec jednako vjerojatan?
- Varijabla `WeightBefore` sadrži podatak o težini majke prije trudnoće, a varijabla `WeightEnd` o težini majke nakon trudnoće. Možemo li tvrditi da je težina značajno veća nakon trudnoće? Usporedno prikažite kutijaste dijagrame te dvije varijable.
- Možemo li tvrditi da veća visina majke (varijabla `HeightMother`) ukazuje na veću visinu djeteta (varijabla `Height`)?
- U varijabli `Rest` naznačeno je ako je majci propisan strogi odmor nakon poroda. Je li ovo obilježje nezavisno o broju djece koje je majka rodila (varijabla `Previous`)?
- Procijenite parametre linearnog modela težine djeteta (`Weight`) u ovisnosti o visini djeteta (`Height`). Jesu li koeficijenti modela statistički značajno različiti od 0? Nacrtajte procijenjeni regresijski pravac zajedno s točkama podataka. Što možete reći o jakosti modela? Prema dobivenom modelu, koliko će biti teško dijete visine 50cm?
- Indeks tjelesne mase definiran je kao tjelesna masa u kilogramima podijeljena s kvadratom visine u metrima. Osoba se smatra pretilom ako je njezin indeks tjelesne mase veći od 30. Je li proporcija pretilih majki nakon poroda značajno veća nego prije poroda?