

STATISTIČKI PRAKTIKUM 1

1. kolokvij - 23. listopada 2019.

Zadatak 1. (12 bodova) U datoteci *statpr-1920-kol1-zad1.dat* nalaze se podaci o učenicima jedne srednje škole. U stupcu *razred* piše koji razred učenik trenutno pohađa, a u stupcima *ime*, *visina* (u m) i *tezina* (u kg) dani su osobni podaci o tim učenicima.

- (a) Učitajte podatke u okvir podataka. U tablicu dodajte stupac *bmi* u kojem će za svakog učenika pisati indeks njegove tjelesne mase koji se računa po formuli: $bmi = \text{tezina} / \text{visina}^2$.
- (b) Ispišite imena onih učenika čiji je indeks tjelesne mase veći od 25.
- (c) Razdvojite podatke o indeksu *bmi* u četiri vektora prema razredu u koji je učenik trenutno upisan i za svaki vektor izračunajte raspon uzorka, aritmetičku sredinu i varijancu. Analizirajte te podatke: sugeriraju li da svi uzorci dolaze iz iste distribucije?
- (d) Sortirajte te uzorke od najmanjeg prema najvećem elementu i nacrtajte ih na istom grafu (s obzirom na odgovarajuće indekse na *x*-osi). Točke neka budu povezane linijama, a svaki razred neka bude prikazan drugom bojom. Graf uredite (dodajte naslov grafa i nazovite *y*-os "bmi" te pripazite na raspon *x* i *y*-osi tako da svi podaci budu vidljivi) i spremite pod nazivom *usporedba1d.pdf*.
- (e) Crtajući histogram usporedite svaki od uzoraka za prvi i četvrti razred s $\chi^2(25)$ razdiobom. Slike spremite pod nazivima *usporedba1e-prvi.pdf* i *usporedba1e-cetvrti.pdf*. Što zaključujete?

Zadatak 2. (18 bodova) U jednoj državi promatrali su broj ljudi s plavom bojom očiju. U tu svrhu uzeli su više uzoraka od po 10 ljudi i za svaki zapisali koliko među njima ima onih s plavom bojom očiju. Podaci su dani u *statpr-1920-kol1-zad2.txt*.

- (a) Promatrajmo sad obilježje X = broj osoba s plavom bojom očiju u uzorku od 10 osoba u toj državi. Učitajte podatke u okvir podataka. Koliko su uzoraka od po 10 ljudi ukupno promatrali?
- (b) Napišite funkciju koja računa frekvencije svih mogućih vrijednosti koje ova varijabla može poprimiti (brojeva od 0 do 10) i pomoću nje odredite frekvencije za vaše podatke.
- (c) Razmislite kojoj bi distribuciji ovo promatrano obilježje moglo pripadati. Na temelju toga, metodom maksimalne vjerodostojnosti odredite procjenu parametra p = vjerojatnost pojave plavih očiju u uzorku od 10 ljudi tako da: nacrtate graf funkcije vjerodostojnosti te nađete multočku njene derivacije (na intervalu na kojem ste vidjeli u prethodnom grafu da se nalazi). Graf funkcije vjerodostojnosti spremite u *graf2c.pdf*.
- (d) Uspredite dobivene apsolutne frekvencije iz uzorka sa onima koje biste očekivali (koristeći rezultat dobiven u (c)) na stupčastom dijagramu. Spremite dobivenu sliku u *usporedba2d.pdf*.
- (e) Sprovedite test pripadnosti pretpostavljenoj distribuciji iz (c) i dajte zaključak pomoću *p*-vrijednosti.



- (f) Ako se zna da je udio ljudi s plavim očima u populaciji te države 20%, usporedite na stupčastom dijagramu pripadajuće relativne frekvencije (za vaš uzorak i tu državu općenito). Spremite dobivenu sliku u *usporedba2f.pdf*

Napomena

Svaki zadatak sačuvajte u svojoj skripti. Rješenje **SVAKOG** podzadatka također napišite (u komentaru) u skripti (uz naredbe).

Pošaljite obje skripte i sve grafove u komprimiranom obliku na statpr1920@gmail.com pod naslovom: prezime_ime (bez kvačica).