

TREĆI KOLOKVIJ IZ VJEROJATNOSTI I STATISTIKE: A grupa

ZADATAK 1:

djelatnici.sta

Baza podataka djelatnici.sta sadrži podatke o uzorku djelatnika jedne tvornice. Za svakog djelatnika iz uzorka zabilježene su vrijednosti sljedećih obilježja:

- varijabla **Spol** sadrži spol svakog djelatnika iz uzorka (M - muško, Z - žensko),
- varijabla **Odjel** sadrži naziv odjela u kojem je djelatnik zaposlen (TR - transport, P - pakiranje, IS - isporuka),
- varijabla **Obrazovanje** sadrži stručnu spremu djelatnika (SSS - srednja stručna sprema, VŠSS - viša stručna sprema, VSS - visoka stručna sprema),
- varijabla **Dob** sadrži starost djelatnika u godinama,
- varijabla **Visina** sadrži visinu djelatnika u centimetrima,
- varijabla **Rukovostvo** sadrži broj godina rada koje je djelatnik proveo na nekoj od rukovodećih pozicija u toj tvornici,
- varijabla **Placa** sadrži iznos mjesečne neto plaće djelatnika u kunama.

Na temelju podataka dostupnih u ovoj bazi odgovorite na sljedeća pitanja:

- a) Kojeg su tipa varijable **Obrazovanje** i **Visina**? (6 bodova)
- b) Odredite tablicu frekvencija i relativnih frekvencija za podatke sadržane u varijabli **Obrazovanje**. (6 bodova)
- c) Procijenite vjerojatnost da je djelatnik stariji od 35 godina (varijabla **Dob**). (8 bodova)
- d) Nacrtajte kružni dijagram (pitu) relativnih frekvencija za podatke koji su sadržani u varijabli **Odjel**. (8 bodova)
- e) Odredite tablicu frekvencija i relativnih frekvencija za podatke sadržane u varijabli **Odjel** posebno za kategoriju ispitanika ženskog spola. (8 bodova)
- f) Nacrtajte zajednički histogram frekvencija i relativnih frekvencija tipa **Overlaid** svih podataka sadržanih u varijabli **Odjel** kategoriziran prema spolu klijenta (varijabla **Spol**). (8 bodova)
- g) Za podatke sadržane u varijabli **Dob** odredite vrijednosti aritmetičke sredine, moda (da li je jedinstven), varijance i standardne devijacije. Ukratko protumačite značenje svake od navedenih numeričkih karakteristika. (8 bodova)
- h) Skicirajte i protumačite kutijasti dijagram na bazi medijana za podatke sadržane u varijabli **Visina**. (8 bodova)
- i) Intervalom pouzdanosti 95% procijenite očekivanje varijable **Visina**. (8 bodova)
- j) Provođenjem prikladnog statističkog testa provjerite možemo li na razini značajnosti $\alpha = 0.05$ tvrditi da je očekivana dob djelatnika tvornice (varijabla **Dob**) statistički značajno veća od $\mu_0 = 30$ godina? (12 bodova)
- k) Možete li na razini značajnosti $\alpha = 0.05$ tvrditi da je varijabla **Dob** normalno distribuirana? (8 bodova)

ZADATAK 2: Vlasnika poznate slastičarnice koja prodaje najbolje krempite u gradu zanima postoji li dio dana u kojemu se kod građana budi veća želja za konzumacijom ovog kolača. Počevši od 10:00 sati odabrao je 5 vremenskih intervala duljine 2 sata i bilježio broj ljudi koji su kupili krempitu. Na razini značajnosti $\alpha = 0.05$ provjerite konzumiraju li građani krempite više u nekom od ponuđenih vremenskih intervala ili ih konzumiraju jednoliko tijekom cijelog mjerenog perioda.

Vremenski interval	10:00 - 12:00	12:00 - 14:00	14:00 - 16:00	16:00 - 18:00	18:00 - 20:00
Broj kupaca	16	24	30	20	10

Koji ste test koristili? (12 bodova)

TREĆI KOLOKVIJ IZ VJEROJATNOSTI I STATISTIKE: B grupa

ZADATAK 1:**crabs.sta**

Baza podataka **crabs.sta** sadrži podatke o jednom biološkom istraživanju u kojem su bilježene reprezentativne karakteristike ženki raka *Carpilius convexus* koje uključuju broj antenica (satelita), stanje kralježnice, boju, težinu, itd. U nastavku je opisano značenje svih varijabli.

- varijabla I sadrži podatke o broju antenica (satelita) kod pojedinog raka: 1 - rak ima barem jednu antenicu; 0 - rak nema antenicu;
- varijabla B sadrži boju jedinke iz uzorka: SS - srednje svijetla; S - svijetla; ST - srednje tamna; T - tamna,
- varijabla KR sadrži podatke o stanju kralježnice: 2D - obje kralježnice u dobrom stanju; 1D - jedna kralježnica je u dobrom stanju dok je druga u lošem; OD - obje kralježnice su u lošem stanju,
- varijabla D sadrži duljina antenica (u cm),
- varijabla NS sadrži točan broj antenica kod jedinke iz uzorka,
- varijabla M predstavlja masu jedinke iz uzorka (u kg).

Na temelju podataka dostupnih u ovoj bazi odgovorite na sljedeća pitanja:

- a) Kojeg su tipa varijable B i D? [6 bodova]
- b) Odredite tablicu frekvencija i relativnih frekvencija varijable KR. [6 bodova]
- c) Procijenite vjerojatnost da slučajno odabrana jedinka nije tamne boje (varijabla B). [8 bodova]
- d) Nacrtajte histogram frekvencija i histogram relativnih frekvencija za podatke koji su sadržani u varijabli B. [8 bodova]
- e) Odredite tablicu frekvencija i relativnih frekvencija za podatke sadržane u varijabli KR posebno za kategoriju jedinki koje imaju antenice. [8 bodova]
- f) Nacrtajte zajednički histogram frekvencija i relativnih frekvencija tipa **Separate** svih podataka sadržanih u varijabli B kategoriziranih prema tome imaju li odgovarajuće jedinke iz uzorka antenice ili ne (varijabla I). [8 bodova]
- g) Za podatke sadržane u varijabli M odredite vrijednosti aritmetičke sredine, moda (da li je jedinstven), varijance i standardne devijacije. Ukratko protumačite značenje svake od navedenih numeričkih karakteristika. [8 bodova]
- h) Skicirajte i protumačite kutijasti dijagram na bazi medijana za podatke sadržane u varijabli NS. [8 bodova]
- i) Intervalom pouzdanosti 95% procijenite vjerojatnost da je masa slučajno odabrane jedinke (varijabla M) manja od 3 kg. [12 bodova]
- j) Kategorizirajte¹ numeričku varijablu M tako da procijenite vjerojatnost da je masa jedinke veća od 3 a manja ili jednaka od 5 kg. [8 bodova]
- k) Možete li na razini značajnosti $\alpha = 0.05$ tvrditi da je varijabla NS normalno distribuirana? [8 bodova]

ZADATAK 2: Voditelj pjevačkog zbora nastoji poštovati zahtjev o jednakoj zastupljenosti prvog, drugog i trećeg glasa u svom zboru. Trenutačno zbor broji 90 pjevača, čije su frekvencije po glasovima dane u sljedećoj tablici:

Glas	Prvi	Drugi	Treći
Broj pjevača	39	28	23

Razlikuje li se ova distribucija na nivou značajnosti $\alpha = 0.05$ statistički značajno od zahtijevane distribucije? Koji ste test koristili? [12 bodova]

¹skicirajte odgovarajući histogram ili tablicu frekvencija i relativnih frekvencija

TREĆI KOLOKVIJ IZ VJEROJATNOSTI I STATISTIKE: C grupa

ZADATAK 1:

djelatnici.sta

Baza podataka djelatnici.sta sadrži podatke o uzorku djelatnika jedne tvornice. Za svakog djelatnika iz uzorka zabilježene su vrijednosti sljedećih obilježja:

- varijabla **Spol** sadrži spol svakog djelatnika iz uzorka (M - muško, Z - žensko),
- varijabla **Odjel** sadrži naziv odjela u kojem je djelatnik zaposlen (TR - transport, P- pakiranje, IS - isporuka),
- varijabla **Obrazovanje** sadrži stručnu spremu djelatnika (SSS - srednja stručna sprema, VŠSS - viša stručna sprema, VSS - visoka stručna sprema),
- varijabla **Dob** sadrži starost djelatnika u godinama,
- varijabla **Visina** sadrži visinu djelatnika u centimetrima,
- varijabla **Rukovodstvo** sadrži broj godina rada koje je djelatnik proveo na nekoj od rukovodećih pozicija u toj tvornici,
- varijabla **Placa** sadrži iznos mjesečne neto plaće djelatnika u kunama.

Na temelju podataka dostupnih u ovoj bazi odgovorite na sljedeća pitanja:

- a) Kojeg su tipa varijable **Dob** i **Odjel**? (6 bodova)
- b) Odredite tablicu frekvencija i relativnih frekvencija za podatke sadržane u varijabli **Odjel**. (6 bodova)
- c) Procijenite vjerojatnost da je djelatnik niži od 160 cm (varijabla **Visina**). (8 bodova)
- d) Nacrtajte histogram frekvencija i histogram relativnih frekvencija za podatke koji su sadržani u varijabli **Spol**. (8 bodova)
- e) Odredite tablicu frekvencija i relativnih frekvencija za podatke sadržane u varijabli **Obrazovanje** posebno za kategoriju ispitanika muškog spola. (8 bodova)
- f) Nacrtajte kružni dijagram (pitu) frekvencija svih podataka sadržanih u varijabli **Obrazovanje** kategoriziran prema spolu klijenta². (8 bodova)
- g) Za podatke sadržane u varijabli **Visina** odredite vrijednosti aritmetičke sredine, moda (da li je jedinstven), varijance i standardne devijacije. Ukratko protumačite značenje svake od navedenih numeričkih karakteristika. (8 bodova)
- h) Skicirajte i protumačite kutijasti dijagram na bazi medijana za podatke sadržane u varijabli **Dob**. (8 bodova)
- i) Intervalom pouzdanosti 95% procijenite očekivanje varijable **Rukovodstvo**. (8 bodova)
- j) Provođenjem prikladnog statističkog testa provjerite možemo li na razini značajnosti $\alpha = 0.05$ tvrditi da je vjerojatnost da je djelatnik tvornice ženskog spola (varijabla **Spol**) statistički značajno veća od $p_0 = 0.4$? (12b.)
- k) Možete li na razini značajnosti $\alpha = 0.05$ tvrditi da je varijabla **Visina** normalno distribuirana? (8 bodova)

ZADATAK 2: Voditelj ZOO-vrta nastoji poštovati zahtjev o jednakoj zastupljenosti određenih skupina životinja.

Podaci o broju životinja koje pripadaju pojedinoj skupini u njegovom ZOO-vrtu dani su u sljedećoj tablici:

divlje mačke	ptice	majmuni	glodavci	morske životinje	ostale životinje
24	36	22	32	60	30

Razlikuje li se ova distribucija na nivou značajnosti $\alpha = 0.05$ statistički značajno od zahtijevane distribucije? Koji ste test koristili? (12 bodova)

²ako ne možete sve očitati, uključite opciju: Appearance: Use document style: System default 000: OK

TREĆI KOLOKVIJ IZ VJEROJATNOSTI I STATISTIKE: D grupa

ZADATAK 1:

crabs.sta

Baza podataka **crabs.sta** sadrži podatke o jednom biološkom istraživanju u kojem su bilježene reprezentativne karakteristike ženki raka *Carpilius convexus* koje uključuju broj antenica (satelita), stanje kralježnice, boju, težinu, itd. U nastavku je opisano značenje svih varijabli.

- varijabla I sadrži podatke o broju antenica (satelita) kod pojedinog raka: 1 - rak ima barem jednu antenicu; 0 - rak nema antenica;
- varijabla B sadrži boju jedinke iz uzorka: SS - srednje svijetla; S - svijetla; ST - srednje tamna; T - tamna,
- varijabla KR sadrži podatke o stanju kralježnice: 2D - obje kralježnice u dobrom stanju; 1D - jedna kralježnica je u dobrom stanju dok je druga u lošem; OD - obje kralježnice su u lošem stanju,
- varijabla D sadrži duljina antenica (u cm),
- varijabla NS sadrži točan broj antenica kod jedinke iz uzorka,
- varijabla M predstavlja masu jedinke iz uzorka (u kg).

Na temelju podataka dostupnih u ovoj bazi odgovorite na sljedeća pitanja:

- a) Kojeg su tipa varijable KR i M? [6 bodova]
- b) Odredite tablicu frekvencija i relativnih frekvencija varijable B. [6 bodova]
- c) Procjenite vjerojatnost da je barem jedna kralježnica kod ženke raka *Carpilius convexus* u dobrom stanju (varijabla KR). [8 bodova]
- d) Nacrtajte histogram frekvencija i histogram relativnih frek. za podatke koji su sadržani u varijabli KR. [8 b.]
- e) Odredite tablicu frekvencija i relativnih frekvencija za podatke sadržane u varijabli B posebno za kategoriju jedinki koje nemaju antenice. [8 bodova]
- f) Nacrtajte zajednički histogram frekvencija i relativnih frekvencija tipa **Overlaid** svih podataka sadržanih u varijabli KR kategoriziranih prema tome imaju li odgovarajuće jedinke iz uzorka antenice ili ne. [8 bodova]
- g) Za podatke sadržane u varijabli NS odredite vrijednosti aritmetičke sredine, moda (da li je jedinstven), varijance i standardne devijacije. Ukratko protumačite značenje svake od navedenih numeričkih karakteristika. [8 b.]
- h) Skicirajte i protumačite kutijasti dijagram na bazi medijana za podatke sadržane u varijabli M. [8 bodova]
- i) Provođenjem prikladnog statističkog testa provjerite možemo li na razini značajnosti $\alpha = 0.05$ tvrditi da je očekivana masa slučajno odabrane jedinke (varijabla M) statistički značajno manja od $\mu_0 = 2.5$ kg. [12 b.]
- j) Kategorizirajte³ numeričku varijablu D tako da procjenite vjerojatnost da je duljina antenice veća od 26 a manja ili jednaka od 28 cm. [8 bodova]
- k) Možete li na razini značajnosti $\alpha = 0.05$ tvrditi da je varijabla NS normalno distribuirana? [8 bodova]

ZADATAK 2: Slastičar želi istražiti imaju li posjetitelji njegove slastičarnice neke posebne sklonosti prema kolačima. Bilježeći koje je kolače naručilo 100 posjetitelja dobio je sljedeću tablicu frekvencija:

tiramisu	rum kocke	kolač s višnjama	baklava	rafaelo
14	29	19	28	10

Slastičara zanima da li na nivou značajnosti $\alpha = 0.01$ posjetitelji preferiraju neke kolače ili je sklonost posjetitelja prema svim kolačima jednaka. Koji ste test koristili? [12 bodova]

³skicirajte odgovarajući histogram ili tablicu frekvencija i relativnih frekvencija