

# Magistrski študij Uporabna statistika

## Uvod v statistiko, študijsko leto 2020/2021

17. 11. 2020

1. Za vsako izmed spodnjih trditev zapišite, ali je PRAVILNA ali NAPAČNA. Vse svoje odločitve kratko utemeljite - pri napačnih trditvah napišite, kaj je narobe in kaj bi bilo prav; kratko utemeljite tudi pri pravilnih.
  - a. Bolečino pri topem udarcu v koleno merimo tako, da jo ovrednotimo z vrednostjo med 1 in 100, kjer 1 predstavlja "brez bolečin". Ta spremenljivka je številska.
  - b. Mediano je smiselno izračunati za vse opisne spremenljivke.
  - c. Ustrezen par mere središčnosti in razpršenosti asimetrično porazdeljene številske spremenljivke sta mediana in standardni odklon.
  - d. Porazdelitev številske spremenljivke na vzorcu lahko grafično predstavimo s histogramom ali z okvirjem z ročaji.
  - e. Vzorčni standardni odklon je slučajna spremenljivka.
  - f. naj bo  $\mu$  populacijsko povprečje. Pri vzorcu velikosti  $n$  je pričakovana vrednost vzorčnega povprečja enaka  $\mu/\sqrt{n}$ .
  - g. Spremenljivka  $X$  naj bo na populaciji A porazdeljena enakomerno na intervalu  $[-100, 100]$ , na populaciji B pa enakomerno na intervalu  $[-200, 200]$ . Standardna napaka  $\bar{X}$  za vzorec velikosti  $n$  iz populacije A je večja kot za vzorec velikosti  $n$  iz populacije B.
  - h. Zanima nas razlika v teži med rokometaši in nogometaši. Ničelna domneva testa Mann-Whitney je, da je v populaciji mediana teže rokometašev enaka mediani teže nogometašev.

2. Na izpitu dobi študent 12 trditev. Za vsako se mora odločiti, ali je pravilna ali napačna. Zanima nas, ali so pravilne trditve enako verjetne kot napačne. Na zadnjem izpitu je bilo 10 trditev napačnih.

Preverite domnevo z dvostranskim testom pri stopnji tveganja  $\alpha = 0.01$ .

- a. Kateri test boste uporabili? Odgovor kratko utemeljite.
- b. Natančno z besedami zapišite alternativno domnevo.
- c. Kaj je testna statistika in koliko znaša?
- d. Določite območje zavrnitve oz. kritično območje.
- e. Kolikšna je velikost testa? Interpretirajte jo na tem konkretnem primeru, tj. uporabite kontekst oz. besedilo naloge.
- f. Kolikšna je vrednost  $p$ ? Interpretirajte jo na tem konkretnem primeru, tj. uporabite kontekst oz. besedilo naloge.
- g. Ali ničelno domnevo zavrnemo? Odgovor utemeljite.
- h. Na podlagi izračunanega zapišite vsebinski sklep testa.
- i. Izračunajte moč testa, če je verjetnost, da je trditev napačna, enaka 0.8. Interpretirajte jo na tem konkretnem primeru, tj. uporabite kontekst oz. besedilo naloge.

3. Pri indeksu telesne mase (ITM) je (glede na določene smernice) meja za prekomerno telesno težo 25. Zanima nas, ali imajo slovenski petošolci prekomerno telesno težo. Zbrali smo vzorec 100 slovenskih petošolcev. Povprečen ITM na vzorcu je enak 27, standardni odklon pa 3. Domnevo preverite z ustreznim dvostranskim testom pri stopnji tveganja  $\alpha = 0.05$ .

- a. Kateri test boste uporabili? Odgovor kratko utemeljite.
- b. Natančno z besedami zapišite ničelno domnevo.
- c. Izračunajte testno statistiko.
- d. Ali ničelno domnevo zavrnemo? Odgovor utemeljite.
- e. Vrednosti  $p$  ne morete natančno izračunati, vendar lahko iz naslednjih možnosti izberete pravilno. Kateri odgovor je pravilen? Vrednost  $p$  je
  - i) 0.066
  - ii) 0.044
  - iii) 0.666
  - iv)  $< 0.001$

Odgovor kratko utemeljite.

- f. Na podlagi izračunanega zapišite vsebinski sklep.
- g. Naše rezultate želimo primerjati z avstrijsko raziskavo, v kateri so preučevali 50 avstrijskih petošolcev. V raziskavi poročajo, da je povprečen ITM enak 25.5, standardni odklon pa 3. Zanima nas, ali imajo slovenski in avstrijski petošolci različen ITM. Domnevo bomo preverili z izračunom primerne 95% intervala zaupanja - sledite spodnjim korakom. Strokovno pomembna razlika v ITM je 1.
  - i) Izračunajte primerene 95% interval zaupanja.
  - ii) Interpretirajte izračunani interval zaupanja.
  - iii) Zapišite ničelno domnevo, ki jo preverjamo.
  - iv) Ali ničelno domnevo zavrnemo?. Utemeljite brez dodatnih izračunov.
  - v) Na podlagi izračunanega zapišite vsebinski sklep, ne pozabite komentirati strokovne pomembnosti.

4. Raziskovalec je s testom  $\chi^2$  preučeval, ali obstaja povezanost med starostjo ( $< 65$  let in  $\geq 65$  let) in indeksom telesne mase (ITM: do 20, 20 do 25, 25 do 30, 30 do 40, 40 ali več). Pričakovane frekvence so prikazane v spodnji tabeli. V raziskavi je preučeval 113 oseb.

	$< 65$	$\geq 65$
do 20	14,31	6,69
20 do 25	20,44	9,56
25 do 30	14,31	6,69
30 do 40	21,12	9,88
40 ali več	6,81	3,19

- a. Izmed spodnjih trditev je natanko ena pravilna. Katera trditev je pravilna? (Izjemoma utemeljitev ni potrebna.)
  - i) Na vzorcu je imelo 21 ljudi ITM manjši od 20, od tega jih je bilo 14 mlajših od 65 let.
  - ii) Na vzorcu je imelo 21 ljudi ITM manjši od 20, od tega jih je bilo 14,31 mlajših od 65 let.
  - iii) Na vzorcu je imelo 21 ljudi ITM manjši od 20. Imamo premalo podatkov, da bi določili, koliko izmed teh je bilo mlajših od 65 let.
  - iv) Imamo premalo podatkov, da bi določili, koliko ljudi je imelo na vzorcu ITM manjši od 20.
  - v) Vse zgornje trditve so napačne.
- b. Ali sta ob veljavnosti ničelne domneve naslednji verjetnosti enaki?
 

$P_1$ : Verjetnost, da ima v populaciji oseba ITM pod 20 in je mlajša od 65 let.  
(Angl.: *The probability that a person from the population has body mass index (BMI) less than 20 and is younger than 65.*)

$P_2$ : Verjetnost, da ima v populaciji oseba ITM od 20 do 25 in je mlajša od 65 let.  
(Angl.: *The probability that a person from the population has body mass index (BMI) from 20 to 25 and is younger than 65.*)

Svoj Odgovor utemeljite tako, da izračunate verjetnosti.
- c. Ali sta ob veljavnosti ničelne domneve naslednji verjetnosti enaki?
 

$P_3$ : Verjetnost, da je v populaciji nekdo z ITM pod 20 mlajši od 65 let.  
(Angl.: *The probability that a person from the population with body mass index (BMI) less than 20 and is younger than 65.*)

$P_4$ : Verjetnost, da je v populaciji nekdo y ITM od 20 do 25 in je mlajši od 65 let.  
(Angl.: *The probability that a person from the population with body mass index (BMI) from 20 to 25 and is younger than 65.*)

Svoj Odgovor utemeljite tako, da izračunate verjetnosti.
- d. Kaj je pri tej raziskavi bolj močen dokaz proti ničelni domnevi - če je na vzorcu 6 ljudi, ki so mlajši od 65 let in imajo ITM vsaj 40, ali če je takih na vzorcu 8? Odgovor kratko utemeljite.  
(Angl.: *What is a stronger proof against the null hypothesis - if there are 6 people on our sample that are younger than 65 and have BMI at least 40, or if there are 8 of them? Justify your answer shortly.*)
- e. Vrednost testne statistike na vzorcu je bila enaka 3.7. Ali ničelno domnevo zavrnamo pri stopnji tveganja  $\alpha = 0.05$ ? Odgovor utemeljite.
- f. Zapišite vsebinski sklep testa.
- g. Denimo, da spremenljivke ITM ne bi kategorizirali in bi nas zanimalo, ali se ITM pri osebah mlajših od 65 let in osebah, ki so starejše od 65 let, razlikuje. Kateri test bi uporabili? Odgovor kratko utemeljite.
- h. Denimo, da spremenljivk ITM in starost ne bi kategorizirali in bi nas zanimalo, ali sta ITM in starost povezana. Kateri test bi uporabili? Odgovor kratko utemeljite.