

$$P(\text{Bin}(n, \pi) = k) = \binom{n}{k} \pi^k (1-\pi)^{n-k}$$

$$P(\text{Bin}(12, 0.5) \geq 11) = P(\dots = 11) + P(\dots = 12) = \binom{12}{11} 0.5^{12} + \binom{12}{12} 0.5^{12}$$

Povaja: Deli malo brez D. → tako bi resili na izpitku

Univerza v Ljubljani

Uporabna statistika, Uvod v statistiko

Inštitut za biostatistiko in medicinsko informatiko

4. vaja

Naloga 1 - eksaktni binomski test → Ali je verjetnost (dogodeka) = cifra?

Želimo preveriti, ali je kovanec pošten. Naredili smo poizkus, kjer smo 12-krat vrgli kovanec in dobili, da je grb padel 11-krat. Označimo z X število grbov, ki jih dobimo, če 12-krat vržemo kovanec. $X = 11$

- Zapišite ničelno domnevo za vaš primer. Ali je ničelna domneva enostavna ali sestavljena? **Trenarst (oz. ni povzročeni)**

$$H_0: \pi = 0.5$$

\uparrow
"T₀" teorija

$$\pi = P(\text{grb})$$

statistična pogema

- Denimo, da je vaša alternativna domneva $H_A : \pi > 0.5$. Ali je ta domneva enostavna ali sestavljena? Ali je domneva enostranska ali dvostranska?

→ oddeli stranovanje

- Ali lahko domnevi zamenjamo glede na naše raziskovalno vprašanje oz. delovno hipotezo?

		H_0 velja	H_0 ne velja	
		H_0 zavrnemo napaka 1. vrste	pravilen step	\downarrow ... postavljen naprej (tipično 0.05)
H_0 obdržimo	H_0 velja	pravilen step	napaka 2 vrste ni omejena	$\Rightarrow H_0$ nisodi ne spregmemmo! H_0 obdržimo / H_0 NE zavrnemo
	H_0 ne velja	napaka 1. vrste ni omejena		

↓
Ne, v H_0 vedno enacst.

- Kaj je vaša testna statistika? Koliko znaša za vaš vzorec?

↳ sluc. spram.
↳ pove veraj o H_0
↳ pozorni moramo porazdelilov ob vlegavnosti H_0

$$X = 11$$

Tu: testna stat.: X

$$X \stackrel{H_0}{\sim} \text{Bin}(n, \pi_0) = \text{Bin}(12, 0.5)$$

- Pri kakšnih vrednostih X boste zavrnili ničelno domnevo v prid alternativni? Zapisi obliko zavrnitvenega oz. kritičnega območja.

↓ pri velikih

$$\hookrightarrow [F, 12]$$

- Za trenutek denimo, da je $X = 6$. Kolikšna je verjetnost, da se na vzorcu zgodi ta dogodek, če ničelna domneva drži? Koliko bi pa dobili, če bi padel grb v 50% pri 200 metih kovanca? Kakšne dogodke moramo torej preučevati?

$$\hookrightarrow P(X=6) = P(\text{Bin}(12, 0.5)=6) = \text{dbinom}(6, 12, 0.5) = 0.226$$

$$\hookrightarrow P(\text{Bin}(200, 0.5)=100) = 0.056$$

↑ prizadana vrednost

→ sklep ne more temeljiti na verjetnosti dogodeka, tj. zato gledamo repe porazdelitev

- Denimo, da je območje zavrnitve sestavljeno iz vrednosti $\{11, 12\}$. Kakšna je $X = \text{cifra}$ velikost testa v tem primeru?

def. $\hookrightarrow P(\text{napaka 1. vrste}) = P_{H_0}(\text{Ho zavrnemo}) = P_{\pi=0.5}(X \geq 11) = P(\text{Bin}(12, 0.5) \geq 11) =$

$$= P(\text{Bin}(12, 0.5) > 10) = 0.003 \rightarrow \text{velikost testa}$$

vR-u:

- Določite območje zavrnitve pri stopnji tveganja oz. značilnosti $\alpha = 0.05$.

plbinom(10, 12, 0.5, lower.tail = F)

↳ največje možno težino, da bo velikost testa $= P(\text{napaka 1. vrste}) < \alpha$

• krit. območje {11, 12} \Rightarrow velikost testa $= 0.003 < \alpha$

• Ali lahko {10, 11, 12}?

$$\text{velikost testa} = P_{H_0}(\text{Ho zavrnemo}) = P_{\pi=0.5}(X \geq 10) = P(\text{Bin}(12, 0.5) \geq 10) = 0.019 < \alpha$$

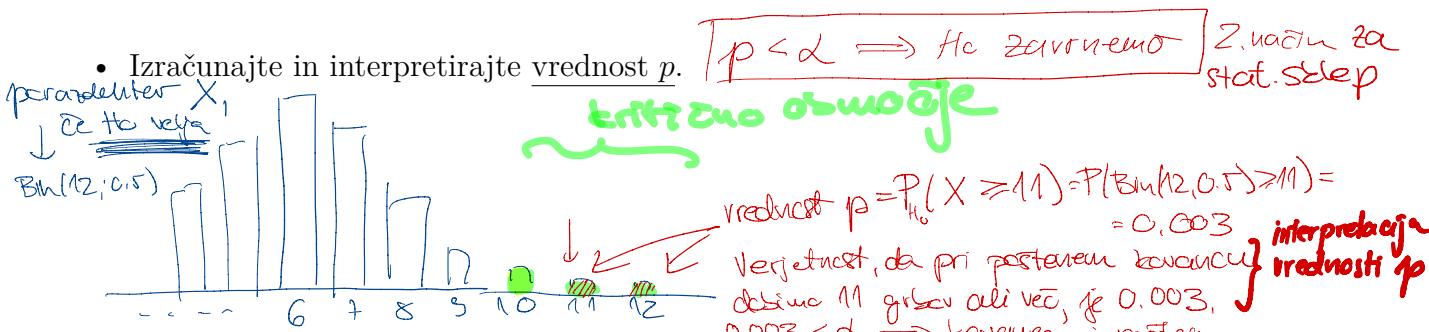
• {5, 10, 11, 12}?

$$\text{velikost testa} = P(\text{Bin}(12, 0.5) \geq 9) = 0.073 > \alpha$$

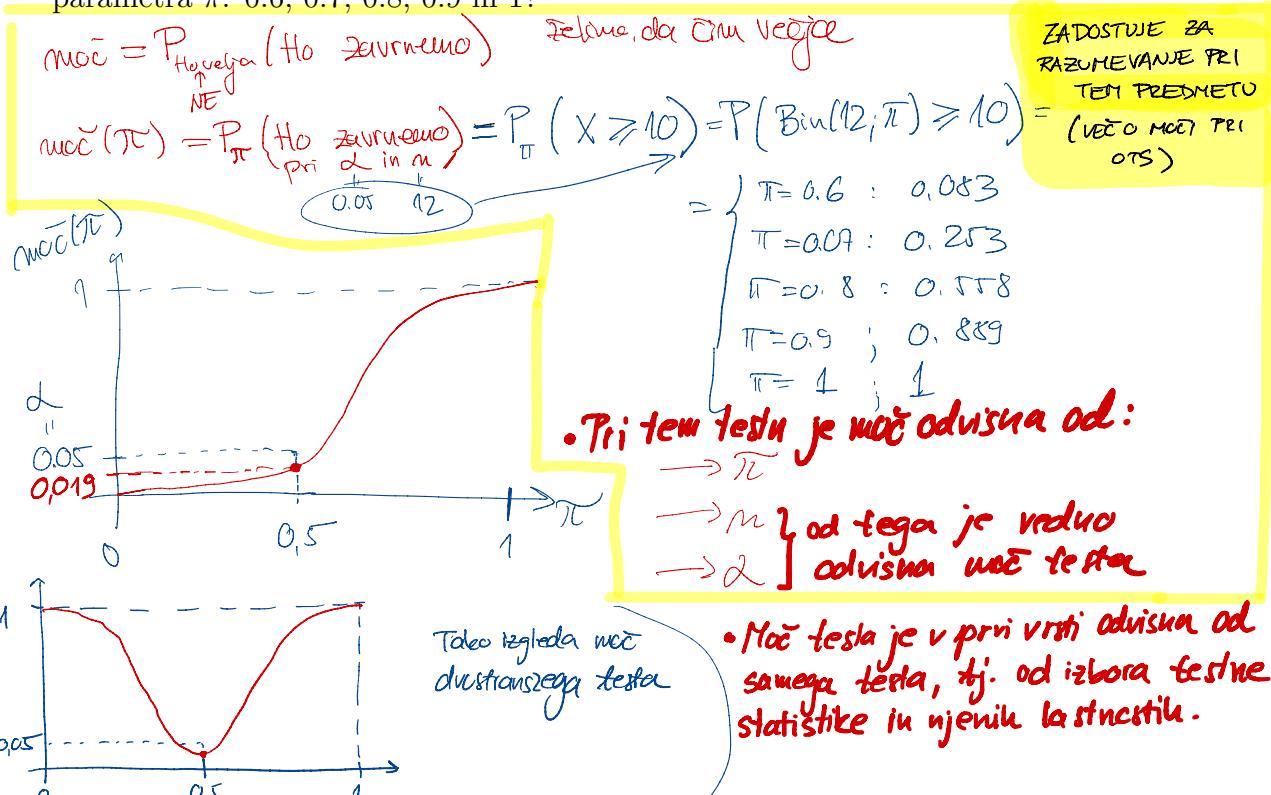
→ kritično območje (oz. območje zavrnitve) : {10, 11, 12}

$$X=11$$

- Ali lahko na podlagi dobljenih podatkov zavrnete ničelno domnevo pri stopnji tveganja $\alpha = 0.05$? To je statistični sklep. H_0 zavrnemo 1. način za stat. sklep
 - $D\rightarrow X \in \text{krit. območje} \Rightarrow H_A \text{ velja}$
- Zapišite vsebinski sklep. \rightarrow z besedami pokaže H_A
Verjetnost gibanja je večja od 0,5.



- Kolikšna je v našem primeru moč testa, če predpostavimo, da je prava vrednost parametra π : 0.6, 0.7, 0.8, 0.9 in 1?



- Kolikšna je verjetnost napake druge vrste, če predpostavimo, da je prava vrednost parametra π : 0.6, 0.7, 0.8, 0.9 in 1?

$$\begin{aligned} P(\text{neprič. 2. vrste}) &= P_{H_0 \text{ velja}}(H_0 \text{ obdržimo}) = 1 - P_{H_0 \text{ velja}}(H_0 \text{ zavrnemo}) \\ &= 1 - \text{moč} = \dots \text{izračunajte sami} \end{aligned}$$

\Rightarrow verjetnost napake 2. vrste je lahko velika, odvisno od π in H_A , ki pa ga ne poznamo

$\Rightarrow H_0$ nikoli ne sprejmemo

Uporabljamo izraz „ H_0 ne zavrnemo“, dovolimo si lahko uporabo izraza „ H_0 obdržimo“

(*) Znati je potrebno:

izračunati kritično vrednost in vrednost p :

→ pri poljutnem H_0 v primeru enostranskega testa $H_0: \pi = \pi_0$
→ pri $\pi_0 = 0.5$ v primeru dvostranskega testa $H_0: \pi = 0.5$

če v Hc primerjava π z $\pi_0 \neq 0.5$, Biu (n, π_0) ni simetrična
 $\pi_0 = 0.5$ → Ni tako enostavno določiti kritično območje, lahko pa uporabimo R in binom. test.

- Naj bo sedaj vaša alternativna domneva $H_A: \pi \neq 0.5$. Ali je ta domneva enostavna ali sestavljena? Ali je domneva enostranska ali dvostranska?

- Pri kakšnih vrednostih X boste v tem primeru zavrnili ničelno domnevo v prid alternativni? Zapišite obliko zavnitvenega območja.

pri zelo majhnih in velikih velikosti X → $[0, c_1] \cup [c_2, 1]$

- Določite območje zavnitve pri stopnji tveganja $\alpha = 0.05$.

prej: $\{10, 11, 12\}$

$$\rightarrow \{0, 1, 2, 10, 11, 12\} \Rightarrow \text{velikost testa} = 2 \cdot 0.019 = 0.038 < \alpha$$

sim. pri 0.5

reno je od enostranskega testa

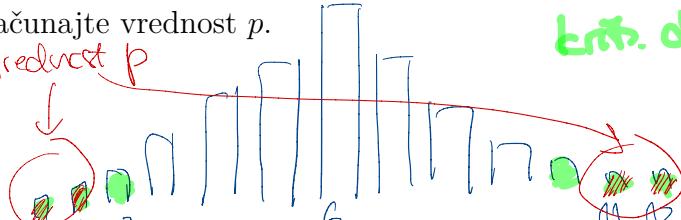
$$\{0, 1, 2, 3, 5, 10, 11, 12\} \Rightarrow \text{velikost testa} > \alpha$$

simetrično glede na $6 (u=12)$

$$\{0, 1, 2, 10, 11, 12\}$$

- Izračunajte vrednost p .

vrednost p



krit. območje

$$p = 2 \cdot P(X \geq 11) = \\ = 2 \cdot 0.003 = 0.006$$

- Ali na podlagi dobljenih podatkov zavrnete ničelno domnevo pri stopnji tveganja $\alpha = 0.05$?

Da, Hc zavrnemo, ker

$11 \in \text{krit. območje}$

- Zapišite vsebinski sklep.

H_A velja $\Rightarrow \pi \neq 0.5$ ↓ ↓ Smer odstopanja iz vzorca

Vrednost grba je večja od 0.5.

Eksatni binomski test v R:

```
binom.test(11, 12, alternative = "greater", p = 0.5) ##enostranski test  
binom.test(11, 12, alternative = "two.sided", p = 0.5) ##dvostranski test
```

Opozimo (spodnje velja splošno za testiranje domnev):

- Vsebinski sklep eno- in dvostranskega testa je enak.
- Pri tem je pri enostranskem vrednost p 2x manjša → imamo 2x močnejši "dokaz proti H_0 " pri enačih podatkih
 - To lahko dosežemo pri vsaki domnvi, t.j. pogledamo, v katero smer je odstopanje na vzorcu, in mato testiramo to smer kot alternativno - to se sveda ne sme - hipoteze postavimo preden vidimo podatke oz. sploh začnemo raziskavo.
 - Da preprečimo "goljufanje", se vedno uporabi dvostranski test oz. moramo imeti za enostranskega zelo močne razloge.

Naloga 2 - test za en vzorec

Preveriti želite vašo raziskovalno domnevo, da študenti medicine v Sloveniji v povprečju za uporabo interneta ne namenijo toliko časa kot njihovi kolegi v Avstriji. Izvedeli ste, da študenti medicine v Avstriji v povprečju internet uporabljajo 23 ur na teden. Za slučajen vzorec 96 študentov medicine (Ankete1011.txt) ste izračunali povprečje in standardni odklon za število ur, ki jih študenti tedensko namenijo uporabi interneta, in dobili: $\bar{x} = 16.77$ in $s = 14.12$.

- Zapišite ničelno in alternativno domnevo.

$$H_0: \mu = 23$$

\uparrow populacija je povprečje

$$H_a: \mu \neq 23$$

miso vzorec - imam le 1 vzorec (SLO)

$H_0:$ Slovenski študenti medicine v povprečju uporabljajo internet
 \uparrow 23 ur/teden (enako kot AUT kolegi).

POPULACIJA

- Izračunajte testno statistiko.

$$T = \frac{\bar{X} - 23}{\frac{s}{\sqrt{n}}} = \frac{16.77 - 23}{\frac{14.12}{\sqrt{96}}} = -4.32$$

Standardizacija \bar{X}

$$T \stackrel{H_0}{\sim} t_{n-1}$$

\uparrow teorija

$$T = \frac{16.77 - 23}{\frac{14.12}{\sqrt{96}}} = -4.32$$

\uparrow t

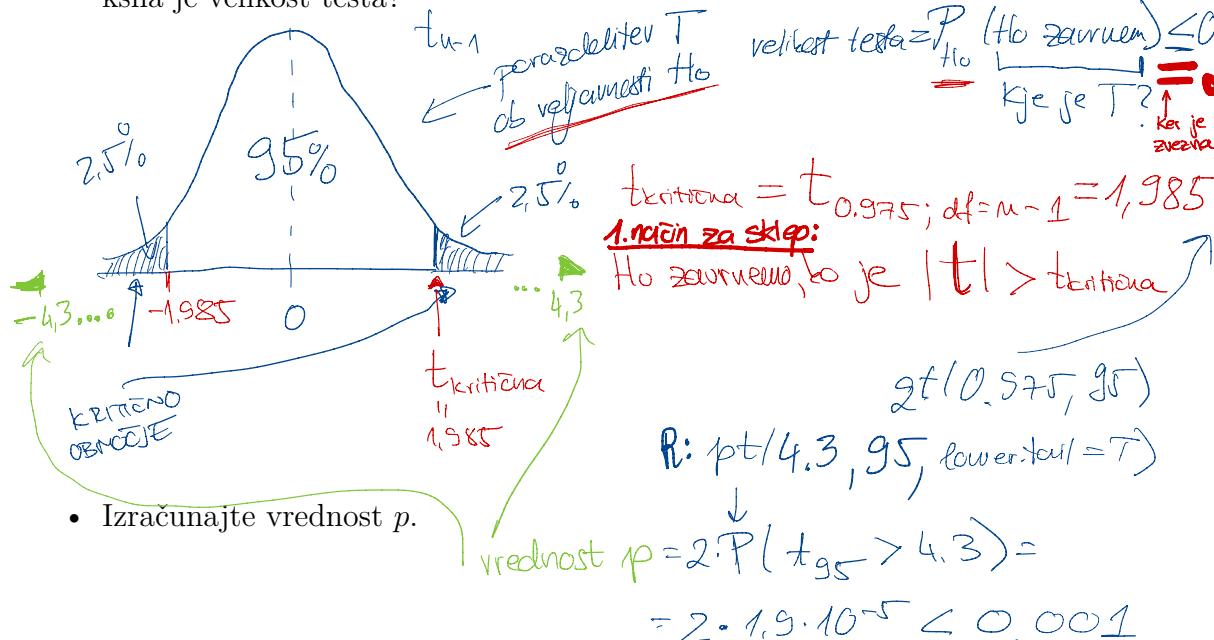
- Zapišite predpostavke testa.

$\rightarrow X$ = uporaba interneta je porazdeljen normalno
 \downarrow dolžina velikosti n , da je $\bar{X} \sim$ normal

\uparrow pridobiveno porazdeljeno

\uparrow Vodilo je predosem, da porazdelitev X ne sme biti tako asimetrična, da bi \bar{X} bistveno odstopalo od mediane.

- Določite kritično območje in kritično vrednost pri stopnji tveganja $\alpha = 0.05$. Kolikšna je velikost testa?



- Izračunajte vrednost p .

- Ali lahko zavrnete ničelno domnevo v prid alternativni? 2. nacin za sklep:

Da, H_0 zavrnemo (t v krit. območju, $p < \alpha$)

- Izračunajte 95% IZ za povprečno uporabo interneta. Kako sta povezana IZ in izračunana vrednost p ?

$$95\% \text{ IZ: } \bar{x} \pm \text{tant.} \cdot \frac{\text{SE}}{\sqrt{n}} = 16,77 \pm 1,985 \cdot \frac{16,12}{\sqrt{96}}$$

$\frac{16,12}{\sqrt{96}}$

to 95%
ista kot pri testu

$\rightarrow [13,9; 19,6]$
E 95%

~~3. vrednost za sklep~~

$23 \notin \text{IZ} \Rightarrow H_0 \text{ zavemo} \quad \mu$

- Zapišite vsebinski sklep. (H_0 velja, $SLO < AUT$)

populacija
Studenci medicine uporabljajo v povprečju internet wayi
kot boljgi iz AUT in sicer vsaj 3,4 h manj na tednu,
nepa več kot za 9 h. \rightarrow IZ

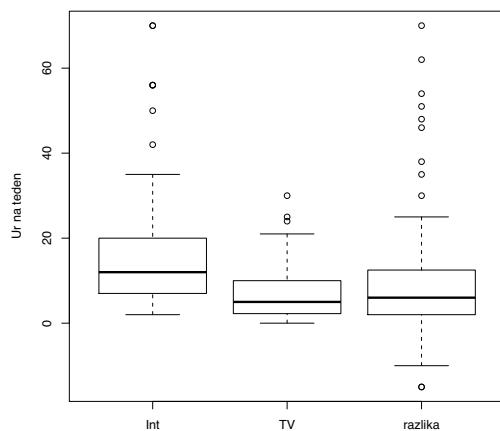
Test t v R:

```
dd <- read.table("Ankete1011.txt", header = T, dec = ",", sep = "\t", fill = T)
id <- which(dd$Studij != "Veterina") #izloci studente veterine
dd.med <- dd[id, ] #podatki samo za studente medicine
dim(dd.med) #n in st. spremenljivk za medicince
summary(dd.med$Internet) ##med drugim izracuna povprecje za vzorec
sd(dd.med$Internet) ##izracuna s

t.test(dd.med$Internet, mu=23, alternative = "two.sided")
```

Naloga 3 - test t za parne meritve oz. dva odvisna vzorca

Želite preveriti vašo domnevo, da študenti medicine v povprečju več časa namenijo uporabi interneta kot gledanju televizije. Za vsakega od 96 slučajno izbranih študentov ste izračunali razliko, razlika = internet – TV. Porazdelitev razlike in nekaj opisnih statistik je prikazano spodaj.



Tipični primeri:

- Prej VS potecu
- Zdrava VS poškodovana noge

Parni test t je natanko

test + za

1 vzorec

ne razlike

	internet	TV	razlika
povprečje	16.7708	6.8510	9.9198
std. odklon	14.1172	6.2331	14.9666
n	96	96	96

- Zapišite ničelno in alternativno domnevo.

$$H_0: \mu_{TV} = \mu_{INT} \text{ oz. } \mu_{INT} - \mu_{TV} = 0 \quad H_A: \mu_{TV} \neq \mu_{INT}$$

↓ Slo. studentje v medicini uporabljajo enako časa TV kot Internet.
pop.

- Izračunajte testno statistiko.

$$T = \frac{\bar{x}_{\text{razlika}} - 0}{\frac{s_r}{\sqrt{n}}} = \frac{9,92}{14,97 / \sqrt{96}} = 6,49 \left(\begin{array}{l} \text{je ne moremo} \\ \text{izraziti z } s_{TV}, s_{INT} \end{array} \right)$$

- Zapišite predpostavke testa.

→ razlika ~ normal
→ velik n } kjer ga testi t za 1 vzorec,
→ dovolj velik n } vendar za sprem. $X = INT - TV$

- Določite kritično vrednost pri stopnji tveganja $\alpha = 0.05$.

$$T \stackrel{H_0}{\sim} t_{n-1}$$

$$t_{\text{krit.}} = t_{0.975, 95} = 1.985$$

isti princip sklike kot prej!

$$6,49 > 1,985 \Rightarrow H_0 \text{ zavrnemo}$$

- Izračunajte vrednost p.

$$p = 2 \cdot P(T > 6,49) = 3.8 \cdot 10^{-9} < \alpha \Rightarrow H_0 \text{ zavrnemo}$$

- Ali lahko zavrnete ničelno domnevo v prid alternativni?

Da.

- Izračunajte 95% IZ za povprečno razliko med uporabo interneta in TV. Kako sta povezana IZ in izračunana vrednost p ?

$[6,89; 12,95]$ $\text{IZ} \Rightarrow \text{to zavrnemo}$
 $\hookrightarrow 95\%$

- Zapišite vsebinski sklep. $(\text{Avejca}; \text{Int} > \text{TV})$

pop- Slo. Studentji medverne v poprejji vec dosa uporabljajo Int
 Int TV, in sicer je razlika med 6,96 in 13h na
 teden.

12

Test t za dva odvisna vzorca v R:

```
dd.med$razlika <- dd.med$Internet - dd.med$Televizija #izracuna sprem. razlika
boxplot(dd.med$Internet, dd.med$Televizija, dd.med$razlika,
        names = c("Int", "TV", "razlika"), ylab = "Ur na teden")
```

```
t.test(dd.med$Internet, dd.med$Televizija, paired = T)
t.test(dd.med$razlika)
```

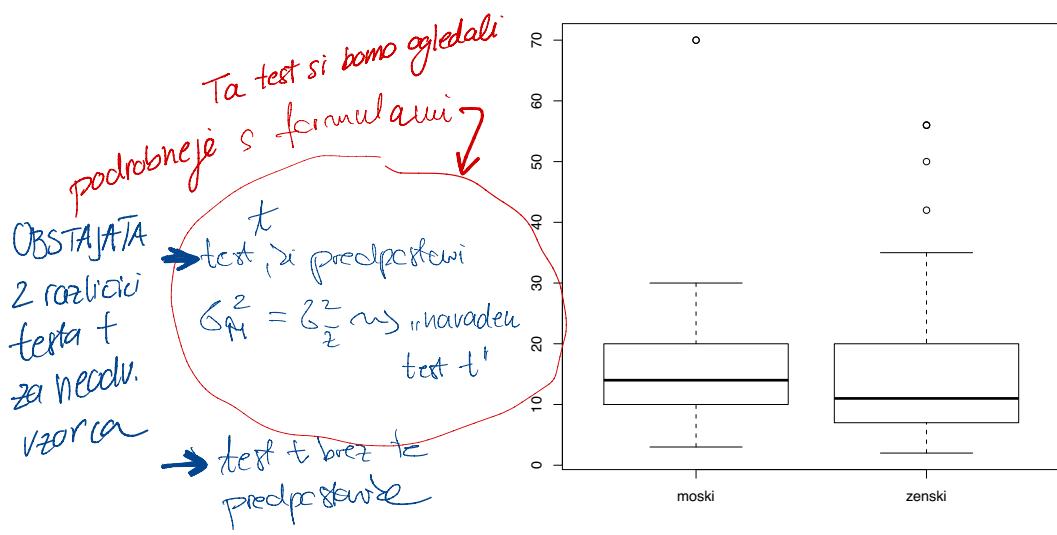
! mu=0 je prednastavljena reduct

primerjava povprečij

2 neodvisnih skupin

Naloga 4 - test t za dva neodvisna vzorca (predpostavljene enake variance)

Želite preveriti raziskovalno domnevo, da pri študentih medicine moški in ženske različno ur namenijo uporabi interneta. Za slučajni vzorec 96 študentov ste pridobili podatek o tedenskem številu ur namenjenih uporabi interneta. Podatke ste grafično prikazali v okviru z ročaji, opisne statistike pa so navedene v spodnji tabeli.



	moški	ženski
povprečje	18.9200	16.0141
std. odklon	16.9802	13.0148
n	25	71

- Zapišite ničelno in alternativno domnevo.

$$H_0: \mu_M = \mu_Z \text{ ali } \mu_M - \mu_Z = 0 \quad H_A: \mu_M \neq \mu_Z$$

Povprečna uporaba int. pri M je enaka kot pri Z v populaciji študentov medicine.

- Izračunajte testno statistiko.

$$T = \frac{\bar{X}_M - \bar{X}_Z}{S_p} = \frac{\bar{X}_M - \bar{X}_Z - 0}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{1}{S_p}$$

$$S_p^2 = \frac{(n_1-1) \cdot s_1^2 + (n_2-1) \cdot s_2^2}{n_1 + n_2 - 2}$$

- Zapišite predpostavke testa.

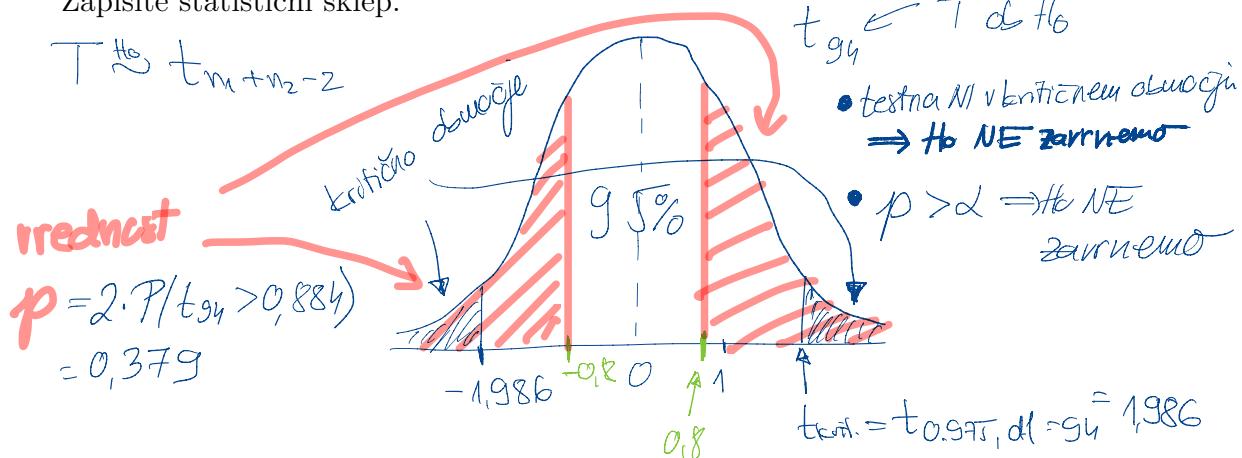
→ internet je porazdeljen normalno
tako pri Z kot pri M
(\rightarrow oz. dolgi veliki skupini)

$$S_p^2 = 199,7532$$

$$t = 0,884$$

→ pri temu testu test: $\& S_p^2 = s^2$
 $\& p > \alpha$

- Določite kritično vrednost pri stopnji tveganja $\alpha = 0.05$. Izračunajte vrednost p . Zapišite statistični sklep.



- Izračunajte 95% IZ za povprečno razliko med uporabo interneta za moške in ženske. Kako sta povezana IZ in izračunana vrednost p ?

• 95% IZ: $\bar{X}_M - \bar{X}_Z \pm t_{0.975, df=m+n_2-2} \cdot \hat{SE}_r = 2.91 \pm 6.53$

$$\Rightarrow [-3.62, 9.44]$$

$$\hat{SE}_r = \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}$$

- Zapisite vsebinski sklep. $H_0: \text{ne moremo spregledi}$ $\rightarrow H_0 \text{ NE ZAVRNEMO}$

Ne moremo spregledi, da H_0 velja. \rightarrow Ne moremo slediti, da H_0 velja.

(Akot tudi ne moremo slediti, da H_0 velja, vendar to ni isto, ker nas zanima, soj H_0 ne moremo veljeti – v populaciji ne morebiti razlike preprost.)

Ne moremo slediti, da H_0 ni \bar{Z} v pop. Stadentov metodice razlike časa uporabljanja interneta; razlika med preprosti je med -3.6 in 9.44 / sledi.

POZOR

NAPACNO: Pop. preprosti sta enaki.

Test t za dva neodvisna vzorca (s predpostavljenimi enakimi variancami in brez te predpostavke) v R:

boxplot(dd.med\$Internet ~ dd.med\$Spol)

t.test(dd.med\$Internet ~ dd.med\$Spol, paired = FALSE, var.equal = TRUE)
t.test(dd.med\$Internet ~ dd.med\$Spol, paired = FALSE, var.equal = FALSE)

la hko damo v mes, ali ~

**Se priporoča,
(*) t.j. Welchov test**

- (*)** • Če je predpostavka o enakih variankah izpolnjena, potem bosta vrednosti p teh dveh testov podobni.
• Če ta predpostavka ni izpolnjena, potem moramo uporabiti Welchov test.

Ko funkcijo $\text{asym. porazdelitev}$, kje je X zelo razlozen od N . "predstavlja strukturjo iz podatkov, ne statistik"

Naloga 5 - Neparametrična testa

Isto raziskovalno vprašanje in vzorec kakor pri prejšnji nalogi. Ugotovili, ste da je spremenljivka internet v populaciji porazdeljena asimetrično v desno, zato želite uporabiti neparametrični test. Podatke ste ustrezno analizirali in dobili spodnji izpis:

testi t so parcijski (parcijski = popravki)
Wilcoxon rank sum test with continuity correction

data: dd.med\$Internet by dd.med\$Spol
 $W = 1016.5$, p-value = 0.281

alternative hypothesis: true location shift is not equal to 0

- Kaj so predpostavke tega testa?

mi jih oz. so predpostavka o neodvisnosti enot

- Zapišite ničelno in alternativno domnevo.

H_0 : Porazdelitev interneta je enaša pri N in Z v Pop. St. medicine.

H_A : različna

- Koliko je znašala vrednost p ?

$$p = 0.281 > \alpha$$

- Ali lahko zavrnete ničelno domnevo v prid alternativni?

✓ H_0 ne zavrnemo

- Zapišite vsebinski sklep.

Ne moremo trditi, da je internet drugače porazdeljen
pri N kot pri Z v Pop. St. medicine.

to mi test primerjave med tem

V R:

Mann-Whitneyev test (drugo poimenovanje: Wilcoxonov test vsote rangov)
je neparametrična varianta testu za neodvisna vzorca:

wilcox.test(dd.med\$Internet ~ dd.med\$Spol, paired = FALSE)

tu damo lahko vmes tudi ~

Wilcoxonov test predznačenih rangov

je neparametrična varianta testu za odvisna vzorca:

wilcox.test(dd.med\$Internet, dd.med\$Televizija, paired = T)

Pro et Contra neparametričnih testov:

④ Ničesar predpostavk.

⑤ Ker ne primerjajo posamezne, vendar pa tudi za celoten posameznik.

Ko moremo uporabiti neparametrični test, zato pogosto naredimo nekoliko kompromis:

* vrednost p pripada neparametričnemu testu

* posamezno dodatnik t^2 za posamezne

* dodamo neke vrste opozicije ce. opombe

(*) Enote uredimo od najmanjših do največjih glede na prenovevalo spremenljivko (obe skupini skupaj).

Poi vsaj si zapomnimo, kje je rang v tej razvrstitvi.

Neparametrični testi obravnavajo tako določene ravne in pozabijo na samo vrednost spremenljivke.

Razmislite do naslednje – povezano tudi s h. nalogo pri DN3

Resimo pri 5. vaji:

Naloga 6 - Statistična značilnost in strokovna pomembnost

Preučevali ste dejavnike, ki so povezani s količino treninga (merjeno v urah na teden). Zbrali ste podatke za 50 športnih plezalcev in 50 alpinistov, ki so registrirani pri Planinski zvezi Slovenije. Strokovno pomembna razlika v količini treninga je 2 uri na teden.

Za povprečno razliko v količini treninga med alpinisti in športnimi plezalci, tj. trening alpinista – trening plezalca, izračunamo 95 % IZ. V spodnji tabeli so zapisane različne vrednosti intervala zaupanja, ki bi jih lahko dobili. Pri vsaki varianti zapišite, ali je rezultat statistično značilen in strokovno pomemben.

najni odgovor – odgovor stranjaka, ne pa statistika

95 % IZ	statistična značilnost	strokovna pomembnost	komentar
[3;6]	da	da	😊
[0,5;1,5]	da	ne	To se zgodi, ko je m velik (upr. pri študijah na poletki). Povezanost ni pomembna, četudi je statistično značilna.
[1;5]	da	merila na populaciji razlika je strokovno pomembna na vzorcu	Nadaljujemo preučevanje tega raziskovalnega vprašanja, morebitno novo raziskavo z večjim m. (Enak razmislek bi lahko morebiti tudi, da razlika ne bi bila strokovno pomembna niti na vzorcu.)
[-1;5]	ne	merila na populaciji razlika je strokovno pomembna na vzorcu	/ /
[-1;1,5]	ne	ne	Povezanost ni pomembna.