Testiranje statističkih hipoteza

22.12.2023.

Rosa Karlić

Testiranje statističkih hipoteza

- **Statistička hipoteza –** pretpostavka o parametru populacije (može i ne mora biti istinita)
- **Nulta hipoteza** (H_0) – hipoteza koju testiramo, uglavnom hipoteza koja govori da su naša opažanja rezultat slučajnosti
- **Alternativna hipoteza**(H_1 ili H_a) – hipoteza koja govori da na ispitanike u uzorku utječe neki ne-slučajni (non-random) uzrok

- Ispitujemo nasumičan uzorak iz populacije
- Ako su podaci iz uzorka u skladu s nultom hipotezom (na određenoj razini pouzdanosti) ne odbacujemo ju, u suprotnom – odbacujemo nultu hipotezu

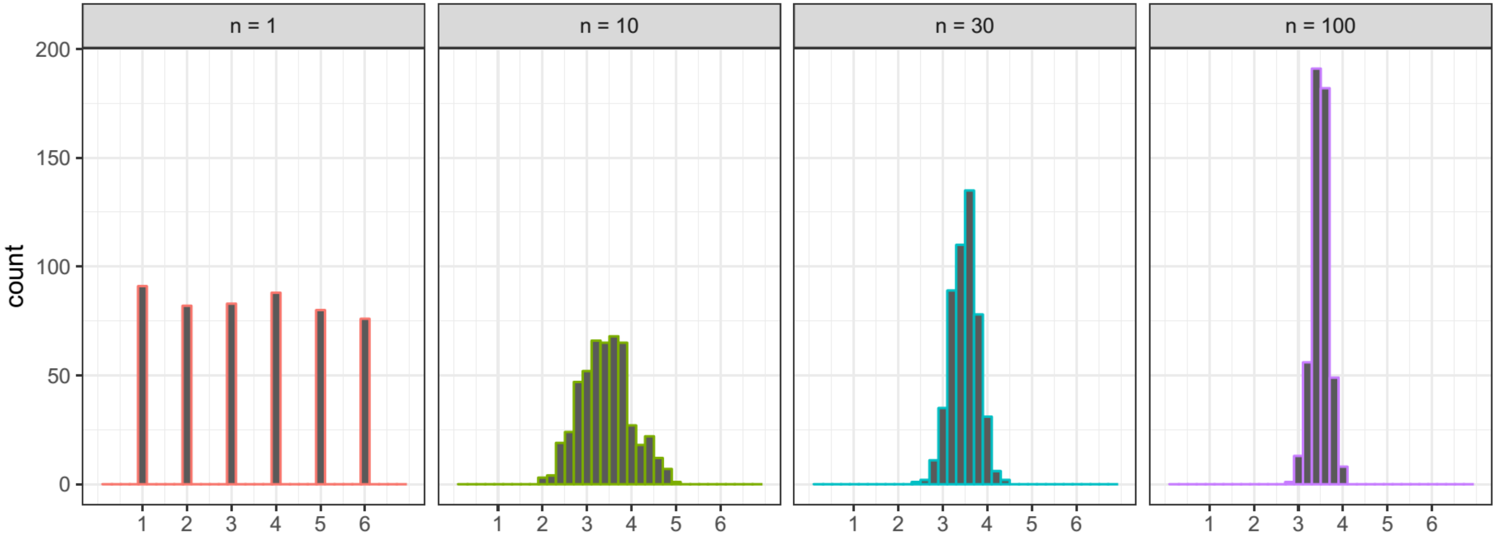
Logička podloga statističkih testova

$$\text{test statistika} = \frac{\text{sistematska varijacija}}{\text{nesistematska varijacija}} = \frac{\text{učinak}}{\text{greška}}$$

- Poznata nam je distribucija test statistike
- To nam omogućuje da izračunamo vjerojatnost da dobijemo određenu vrijednost test statistike ukoliko je nulta hipoteza točna (P-vrijednost)

Središnji granični teorem

- Distribucija procjene statistike (zbroja ili srednje vrijednosti) i.i.d. slučajnih varijabli biti će normalna ili gotovo normalna ukoliko je veličina uzorka dovoljno velika



Standardna devijacija statistike

$$\sigma_{\overline{x}} = \sigma / \sqrt{n}$$

Standardna greška statistike

$$SE_{\overline{x}} = s / \sqrt{n}$$

Studentov t-test (jedan uzorak)

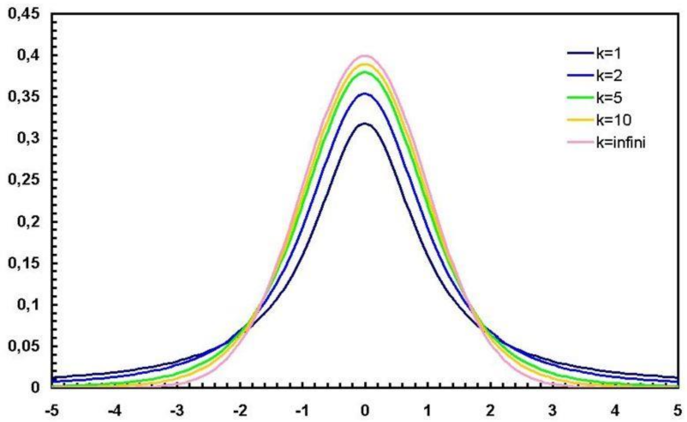
- t-statistika

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

→ učinak

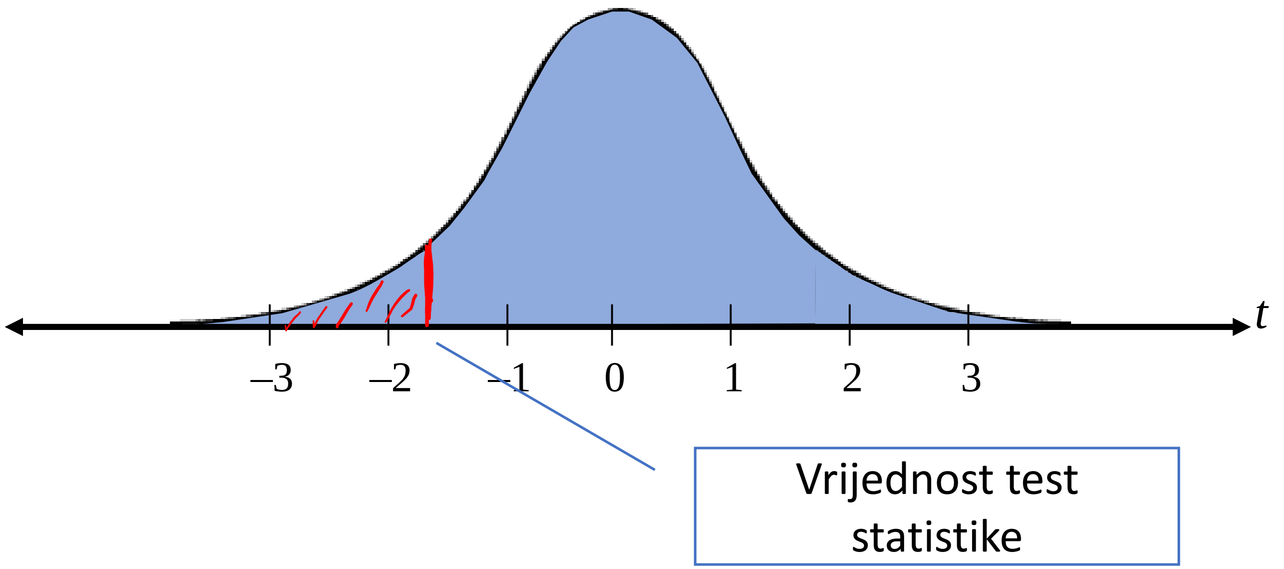
→ greška

$$SE_{\bar{x}} = s / \sqrt{n}$$



$$f_{\nu}(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

- Ukoliko je nulta hipoteza točna t-statistika će slijediti t-distribuciju sa n-1 stupnjeva slobode

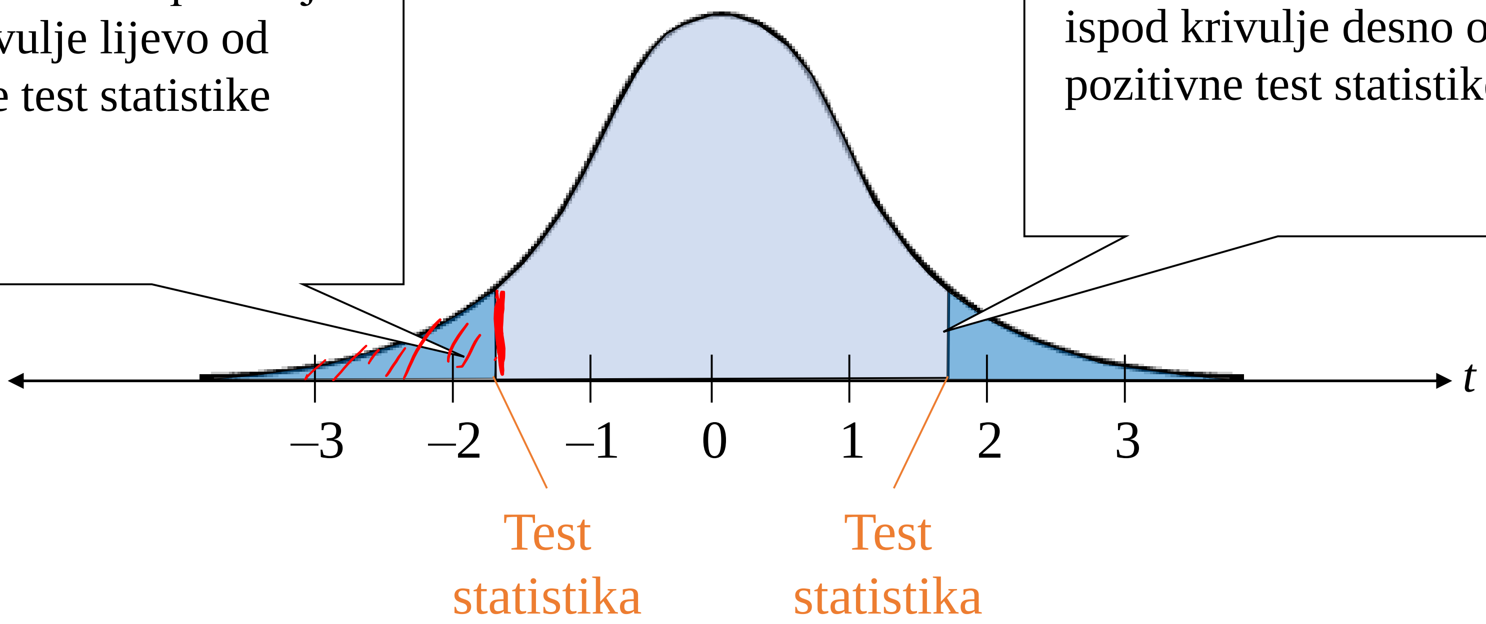


Studentov t-test (jedan uzorak)

$$H_0: \mu = k$$
$$H_a: \mu \neq k$$

P-vrijednost =2 x područje ispod krivulje lijevo od negativne test statistike

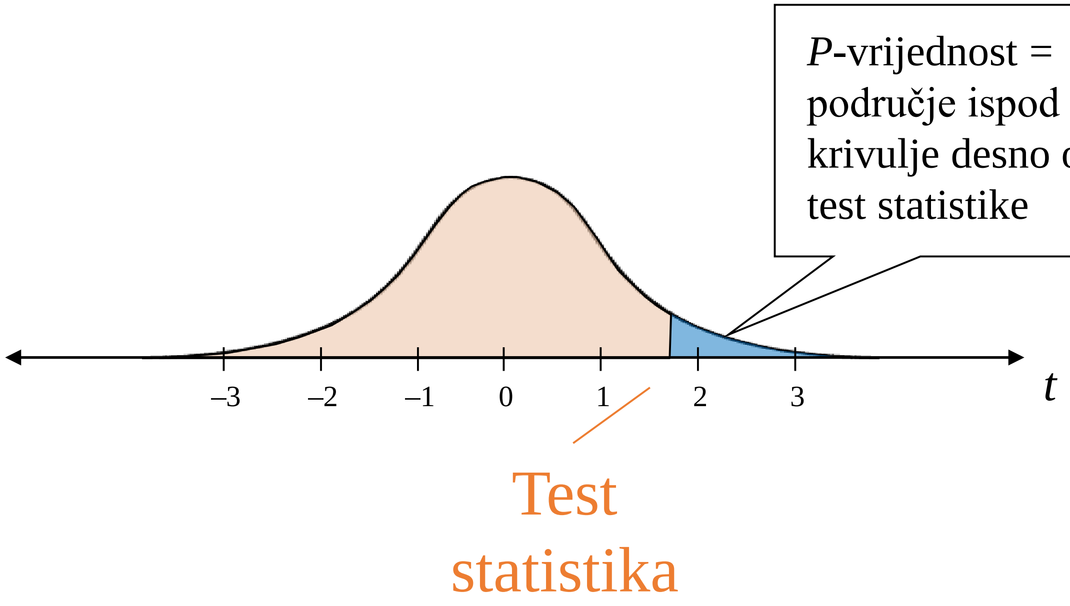
P-vrijednost =2 x područje ispod krivulje desno od pozitivne test statistike



Studentov t-test (jedan uzorak)

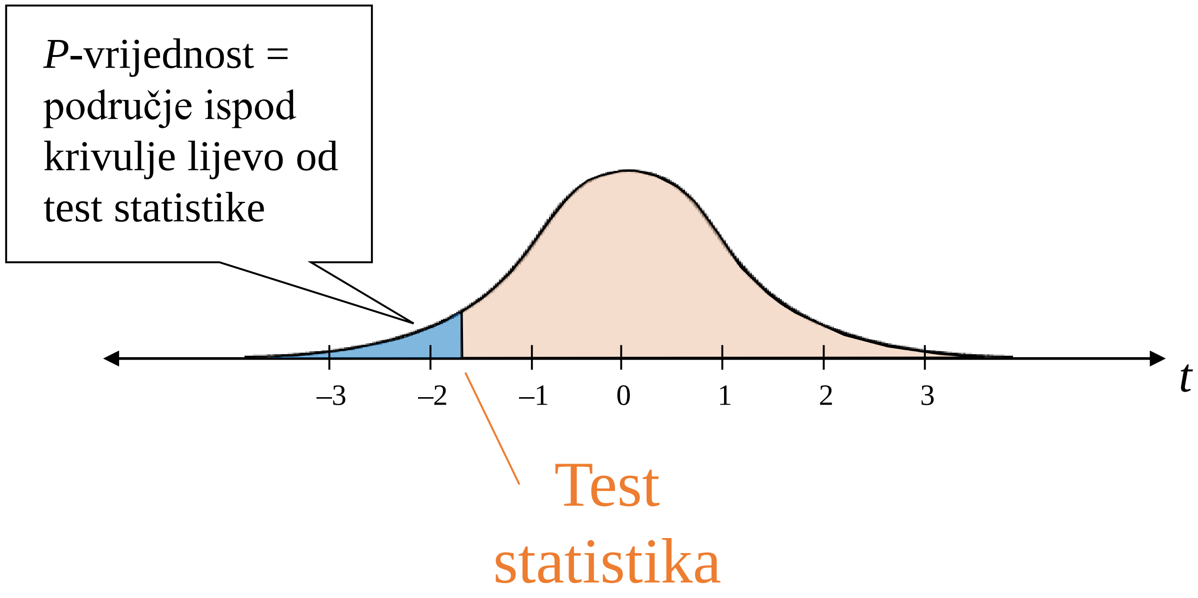
$H_0: \mu \leq k$

$H_a: \mu > k$



$H_0: \mu \geq k$

$H_a: \mu < k$



Studentov t-test (jedan uzorak)

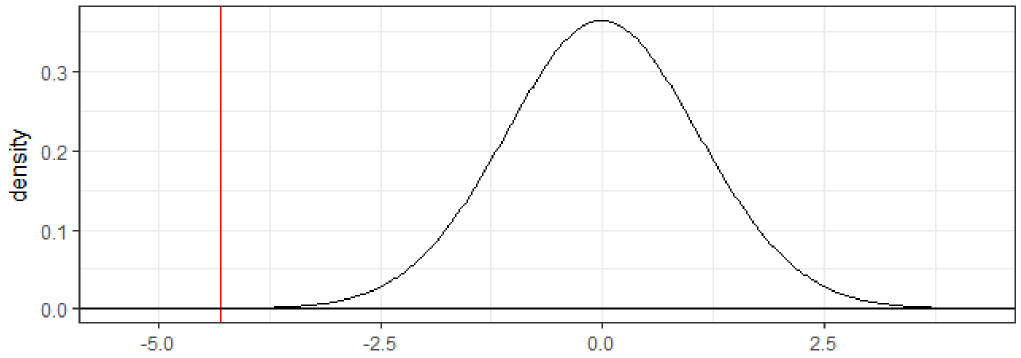
ID	STAROST	SPOL	TRETMAN	GEN1	GEN2	GEN3	STATUS
87	53	F	trt2	17.41	28.23	4.17	zdrav
119	53	M	ctrl	19.84	52.56	47.24	zdrav
67	52	M	trt2	19.19	27.49	12.07	zdrav
62	54	M	trt2	22.77	28.45	9.62	bolestan
131	55	F	ctrl	24.17	49.91	49.55	bolestan
50	54	F	trt1	17.15	15.32	10.67	zdrav
106	54	M	ctrl	17.92	44.95	51.39	zdrav
127	58	F	ctrl	20.06	53.19	49.71	bolestan
30	54	M	trt1	19.97	16.18	13.78	zdrav
72	54	F	trt2	25.44	27.58	11.81	zdrav

$$H_0: \mu = 25$$

$$H_a: \mu \neq 25$$

$$t_{n-1} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

```
        One Sample t-test

data:  data[, "GEN1"]
t = -4.3036, df = 9, p-value = 0.00198
alternative hypothesis: true mean is not equal to 25
95 percent confidence interval:
 18.64107 23.02285
sample estimates:
mean of x
 20.83196
```

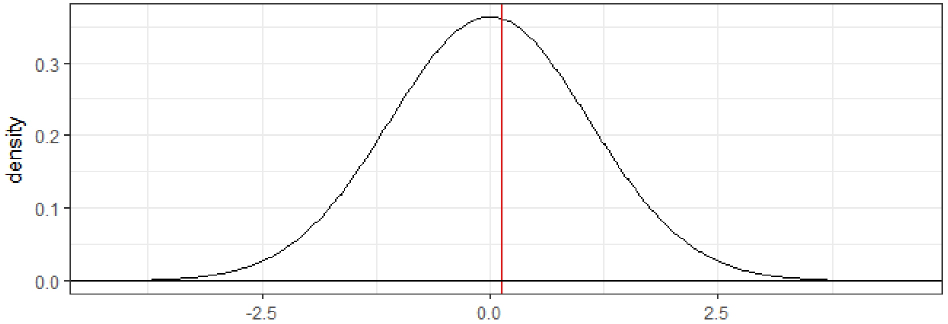


Studentov t-test (dva uzorka)

ID	STAROST	SPOL	TRETMAN	GEN1	GEN2	GEN3	STATUS
87	53	F	trt2	17.41	28.23	4.17	zdrav
119	53	M	ctrl	19.84	52.56	47.24	zdrav
67	52	M	trt2	19.19	27.49	12.07	zdrav
62	54	M	trt2	22.77	28.45	9.62	bolestan
131	55	F	ctrl	24.17	49.91	49.55	bolestan
50	54	F	trt1	17.15	15.32	10.67	zdrav
106	54	M	ctrl	17.92	44.95	51.39	zdrav
127	58	F	ctrl	20.06	53.19	49.71	bolestan
30	54	M	trt1	19.97	16.18	13.78	zdrav
72	54	F	trt2	25.44	27.58	11.81	zdrav

$$t_{n_1+n_2-2} = \frac{\bar{X}_1 - \bar{X}_2}{SE_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_P^{\,2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

```
        Two Sample t-test

data:  data[data$TRETMAN == "trt2", "GEN1"] and data[data$TRETMAN == "ctrl", "GEN1"]
t = 0.13371, df = 98, p-value = 0.8939
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.7221867  0.8265337
sample estimates:
mean of x mean of y 
 20.38154  20.32937 
```



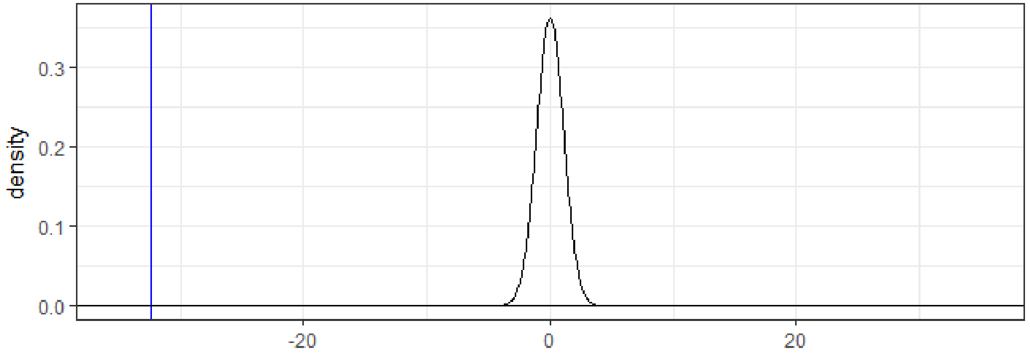
Studentov t-test (dva uzorka)

ID	STAROST	SPOL	TRETMAN	GEN1	GEN2	GEN3	STATUS
87	53	F	trt2	17.41	28.23	4.17	zdrav
119	53	M	ctrl	19.84	52.56	47.24	zdrav
67	52	M	trt2	19.19	27.49	12.07	zdrav
62	54	M	trt2	22.77	28.45	9.62	bolestan
131	55	F	ctrl	24.17	49.91	49.55	bolestan
50	54	F	trt1	17.15	15.32	10.67	zdrav
106	54	M	ctrl	17.92	44.95	51.39	zdrav
127	58	F	ctrl	20.06	53.19	49.71	bolestan
30	54	M	trt1	19.97	16.18	13.78	zdrav
72	54	F	trt2	25.44	27.58	11.81	zdrav

$$t_{n_1+n_2-2} = \frac{\bar{X}_1 - \bar{X}_2}{SE_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_P{}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$s_p^2 = \frac{\sum_{i=1}^{k}(n_i-1)s_i^2}{\sum_{i=1}^{k}(n_i-1)} = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2+\cdots+(n_k-1)s_k^2}{n_1+n_2+\cdots+n_k-k}$$

```
        Two Sample t-test

data:  data[data$TRETMAN == "trt2", "GEN2"] and data[data$TRETMAN == "ctrl", "GEN2
t = -32.433, df = 98, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -20.7178 -18.3287
sample estimates:
mean of x mean of y
 30.10788  49.63114
```

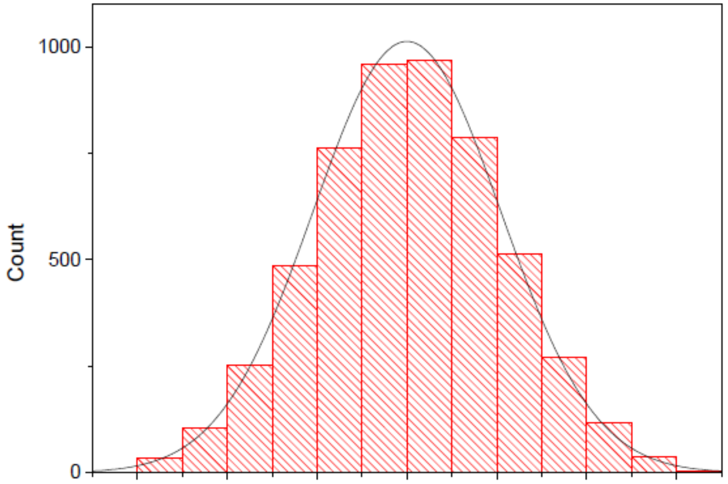


Pretpostavke za t-test

- Varijabla koju testiramo je normalno distribuirana
- Uzorci imaju jednake varijance
- Uzorci su međusobno neovisni

Testiranje normalnosti:

- Histogram, druge grafičke metode
- Statistički testovi (Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling)



Što ako naši podaci krše neku od pretpostavki?

- **Welch test** – koristi se u slučaju nejednakih varijanci
- Modificirana t-statistika
- Obično se koristi za uzorke jako različitih veličina

- **Studentov t-test za povezane uzorke** – koristi se u slučaju da su podaci u uzorcima međusobno ovisni

Studentov t-test za povezane uzorke

- Kada uspoređujemo dva povezana uzorka često nas ne zanimaju apsolutne vrijednosti već razlika između izmjerenih vrijednosti za svaki par observacija u povezanih uzorcima

Uzorak 1	Uzorak 2	Razlika
X_{11}	X_{21}	$X_{11} - X_{21}$
X_{12}	X_{22}	$X_{12} - X_{22}$
.	.	.
.	.	.
.	.	.
X_{1n}	X_{2n}	$X_{1n} - X_{2n}$

- Uzorak razlika u mjerenjima uzet iz populacije razlika
- Karakterizira ga srednja vrijednost i standardna devijacija
- Srednja vrijednost razlika u mjerenjima $\bar{D}$

Studentov t-test za povezane uzorke

- Nulta hipoteza: Nema razlike među dva međusobno ovisna mjerenja u našem uzorku
- Zbog varijabilnosti očekujemo da će neke razlike biti pozitivne, neke negativne, ali srednja vrijednost tih razlika biti će jednaka nuli

Studentov t-test za povezane uzorke

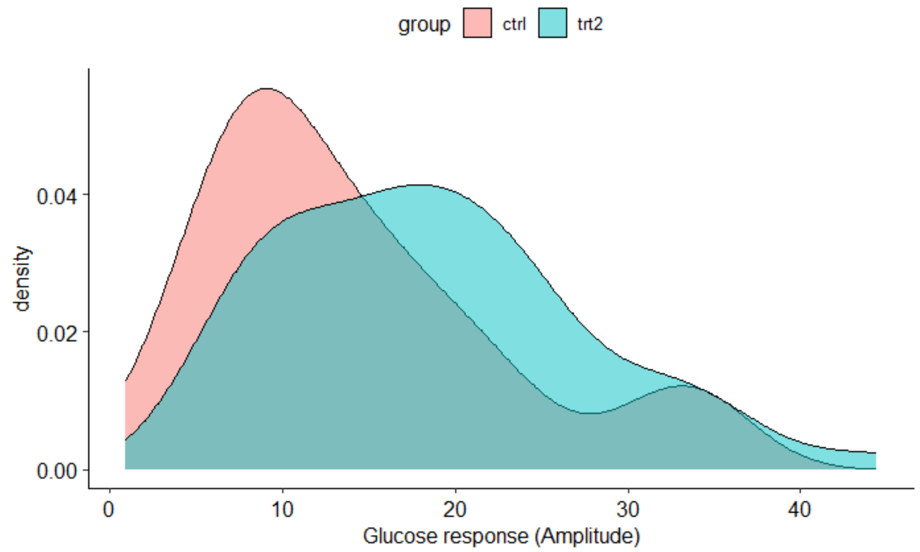
- Koristimo t-test kako bismo utvrdili je li srednja vrijednost razlika značajno različita od nule

$$t_{n-1} = \frac{\bar{D} - \mu_{\bar{D}}}{SE_{\bar{D}}}$$

- Naš uzorak je zapravo stupac s razlikama u mjerenjima na kojem ćemo napraviti t-test na jednom uzorku s nultom hipotezom da je $\mu = 0$

Što ako podaci nisu normalno distribuirani?

ID	AGE	SEX	TRETMAN	GEN1	GEN2	GEN3	STATUS	GLUCOSE RESPONSE (Amplitude)
87	53	F	trt2	17.41	28.23	4.17	healthy	6.95
119	53	M	ctrl	19.84	52.56	47.24	healthy	6.41
67	52	M	trt2	19.19	27.49	12.07	healthy	12.85
62	54	M	trt2	22.77	28.45	9.62	sick	7.88
131	55	F	ctrl	24.17	49.91	49.55	sick	6.05
50	54	F	trt1	17.15	15.32	10.67	healthy	9.76
106	54	M	ctrl	17.92	44.95	51.39	healthy	5.16
127	58	F	ctrl	20.06	53.19	49.71	sick	13.51
30	54	M	trt1	19.97	16.18	13.78	healthy	11.41
72	54	F	trt2	25.44	27.58	11.81	healthy	30.85



Neparametarski testovi

- Testovi koji nemaju pretpostavki o distribuciji proučavanih varijabli
- Testiraju se rangovi podataka, a ne same vrijednosti podataka
- Mogu se koristiti za male uzorke (N<30) i za varijable koje ne slijede normalnu distribuciju (!)
- Manja statistička snaga testa

- **Wilcoxon-Mann-Whitney rank-sum test (Mann-Whitney U)** – usporedba dva neovisna uzorka
- **Wilcoxon signed rank test** – usporedba dva ovisna uzorka

Wilcoxon-Mann-Whitney Rank Sum Test

- Test statistika računa se na rangovima, a ne sirovim podacima

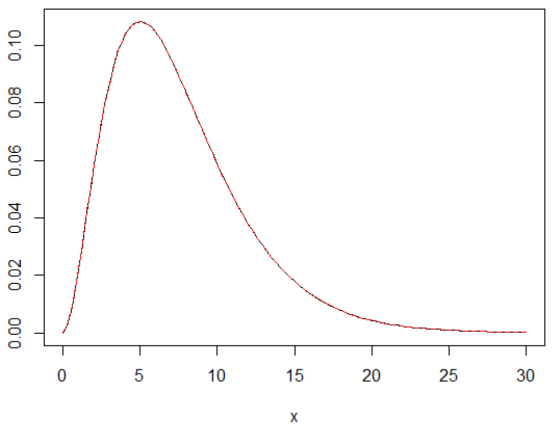
- Hipoteza
  - H_0 – dvije populacije imaju identičnu distribuciju
  - H_1 – "lokacije populacija se razlikuju"

- Pretpostavke:
  - Podaci dolaze iz nasumičnih uzoraka
  - Podaci unutar uzoraka su međusobno neovisni
  - Uzorci su međusobno neovisni

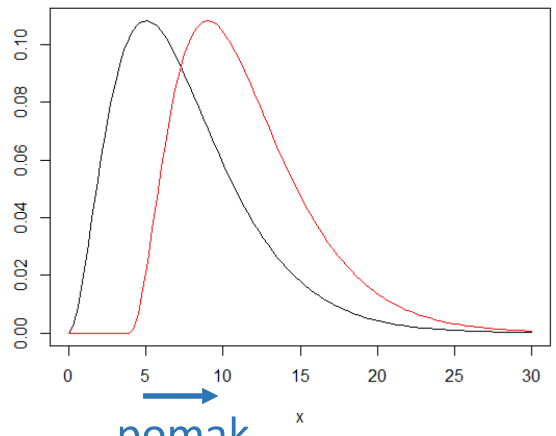
Wilcoxon-Mann-Whitney Rank Sum Test

- Wilcoxon -Mann-Whitney Rank Sum Test pokušava otkriti pomake u lokaciji
  - H1 : A > B (A pomaknuta u desno od B)
  - H1 : A < B (A pomaknuta u lijevo od B)
  - H1 : A ≠B

H_0: A = B



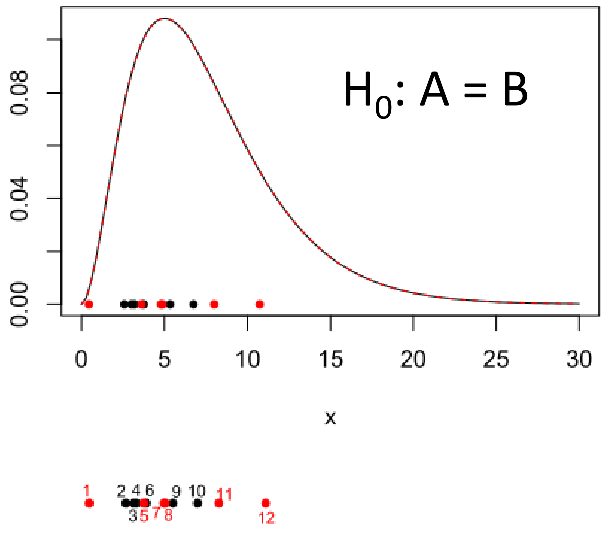
H_1: A > B

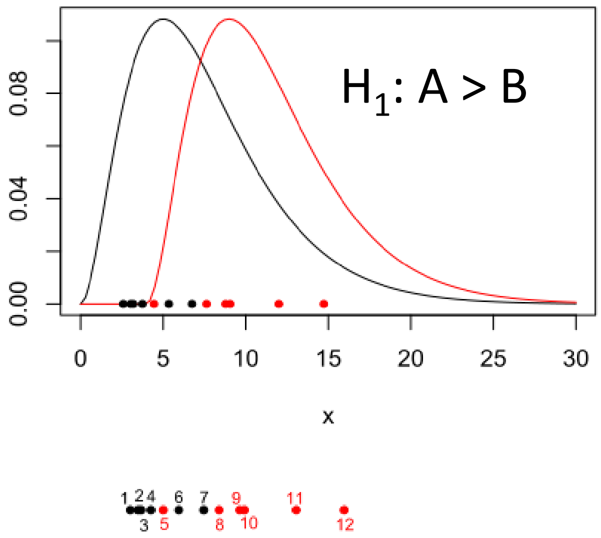


pomak

Wilcoxon-Mann-Whitney Rank Sum Test

- Staviti sve observacije u jedan uzorak i rangirati $n_A + n_B$ observacija kombiniranog uzorka
- Wilcoxon rank-sum statistika – suma rangova observacija iz jednog od uzoraka
- w_A = suma rangova observacija iz uzorka A
- a) H_0:A=B b) H_1: A>B

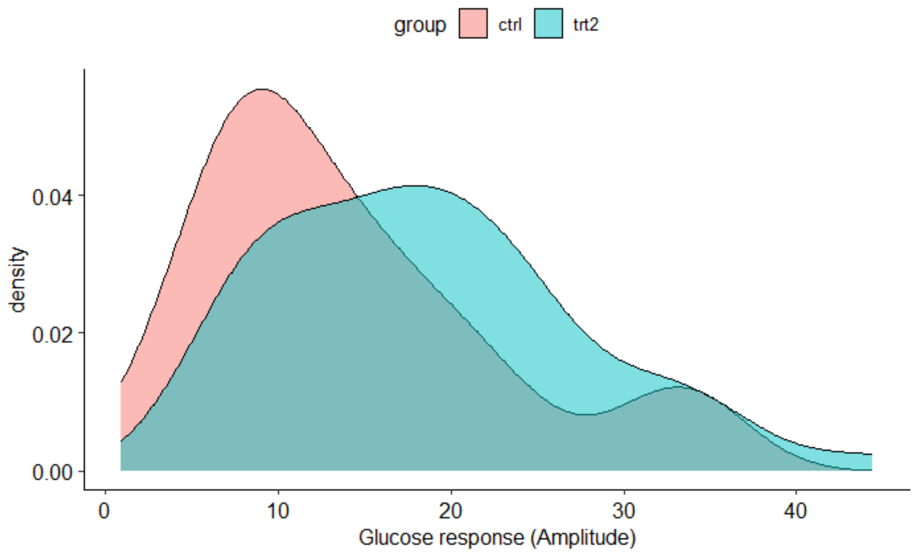




$(H_1 : A > B)$ $P\text{-}value = \mathrm{pr}(W_A \geq w_A).$

$(H_1 : A < B)$ $P\text{-}value = \mathrm{pr}(W_A \leq w_A).$

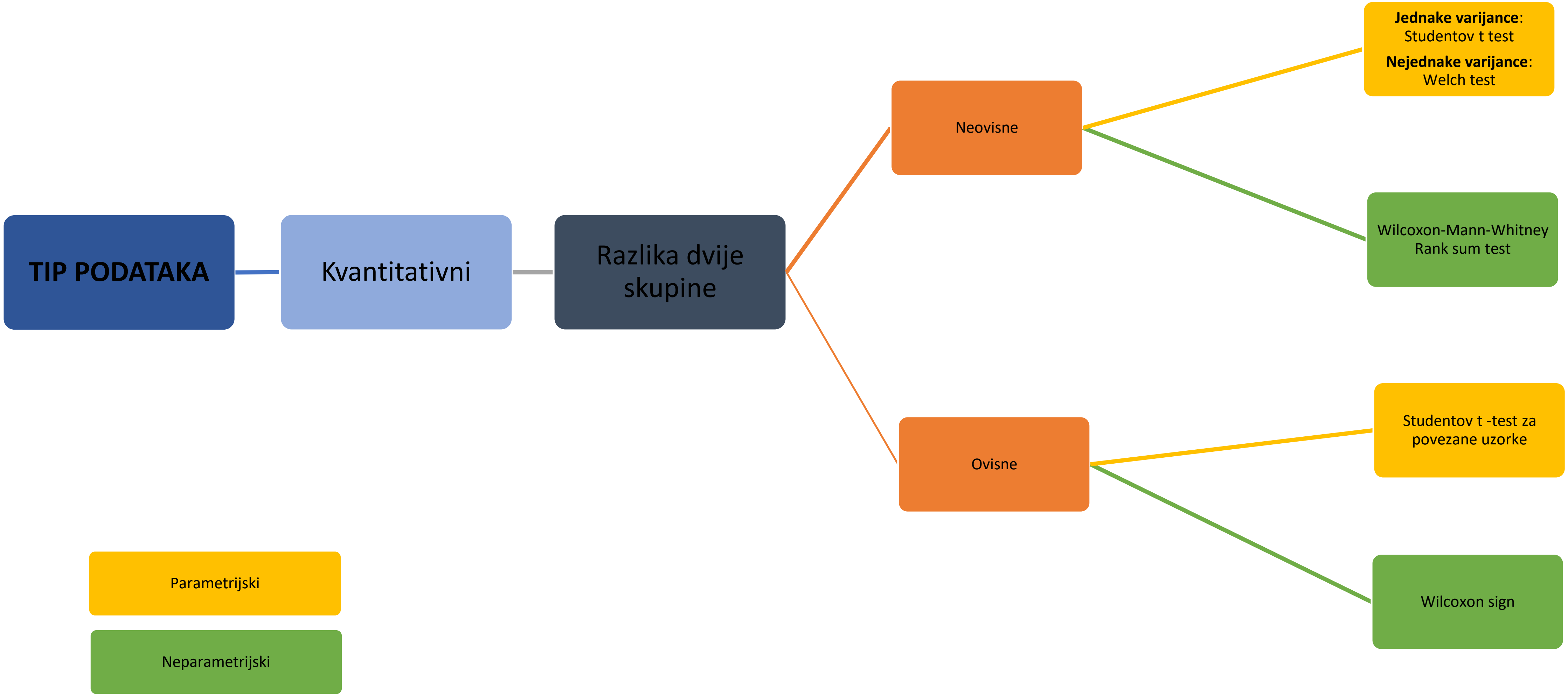
Wilcoxon rank-sum test na našem primjeru




```
> wilcox.test(ctrl, trt2)

        Wilcoxon rank sum test with continuity correction

data:  ctrl and trt2
W = 851, p-value = 0.006011
alternative hypothesis: true location shift is not equal to 0

> t.test(ctrl, trt2, var.equal =  T)

        Two Sample t-test

data:  ctrl and trt2
t = -2.5267, df = 98, p-value = 0.01311
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.931406 -0.953410
sample estimates:
mean of x mean of y
 14.17141  18.61381
```

Wilcoxon signed rank test

- Verzija prethodnog testa gdje se pojedinci mogu mjeriti dvaput ili se mogu uzeti u obzir mjerenja prije i poslije (spareni test)
- Podaci koji imaju isti broj mjerenja.
- Upareni uzorci $(X_i,Y_i) \rightarrow D_i=X_i-Y_i$
- Ako niti jedan tretman nema učinka, tada ne samo da bi razlike trebale biti jednako raspoređene s obje strane 0, nego bi i to koliko su razlike udaljene od 0 trebale biti isto s obje strane.
- Postupak:
  - Izračunajte D_i i rangirati
  - Izračunajte W+ = zbroj rangova s pozitivnim predznakom ili W− = zbroj rangova s negativnim predznakom
  - Idejaako je distribucija X (F(x)) ista kao distribucija Y (F(y)), tada je jednako vjerojatno da će Di biti pozitivni kao i negativni. Dakle, oko pola rangova je pozitivno, a pola je negativno.ako je F(y) veći od F(x), očekujte da većina rangova ima pozitivne predznake i stoga će W+ biti velik
- Za n≥20 približno normalno raspoređen


TIP PODATAKA
Kvantitativni
Razlika dvije skupine
Neovisne
Ovisne
Jednake varijance: Studentov t test
Nejednake varijance: Welch test
Wilcoxon-Mann-Whitney Rank sum test
Studentov t -test za povezane uzorke
Wilcoxon sign
Parametrijski
Neparametrijski

Primjer 1: Distribucije podataka

- Koristeći Excel proizvedite grafičke i numeričke sažetke podataka

Primjer 2: Utjecaj genotipa na fenotip

- Navedite broj i tipove varijabli
- Odredite statistički test i nultu hipotezu
- Provjerite zadovoljavaju li podaci pretpostavke statističkog testa
- Odlučite koji statistički test koristiti
- Provedite statistički test
- Interpretirajte rezultate statističkog testa

Primjer 3: Utjecaj genotipa na fenotip, v2

- Navedite broj i tipove varijabli
- Odredite statistički test i nultu hipotezu
- Provjerite zadovoljavaju li podaci pretpostavke statističkog testa
- Odlučite koji statistički test koristiti
- Provedite statistički test
- Interpretirajte rezultate statističkog testa

Primjer 4: Utjecaj tretmana na krvni tlak

- Koristeći Excel i podatke u datoteci data.practical2.ex3.txt testirajte hipotezu da je srednja vrijednost sistoličkog tlaka jednaka u kontrolnoj skupini i u skupini pacijenata koja je primila tretman.
- Mijenja li se rezultat ako su isti pacijenti u kontrolnoj skupini i skupini koja je primila tretman?