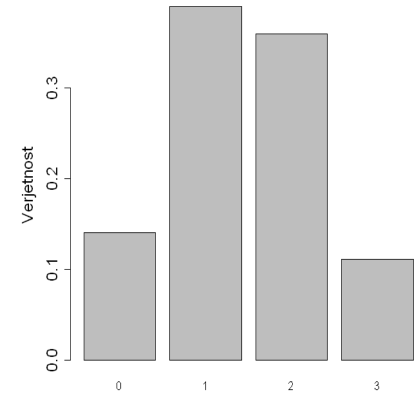
Inferenčna statistika

Inštitut za biostatistiko in medicinsko informatiko
Medicinska fakulteta, Univerza v Ljubljani

Binomska porazdelitev

$$P(K = k \mid n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

parametra Bin(n,p)



n – število enot (neodvisne)

k – slučajna spremenljivka (število uspešnih dogodkov pri n enotah)

p – verjetnost uspeha

Povprečna vrednost (pričakovana vrednost, expected value)

E(K)=ΣP(K=k)*k=n*p

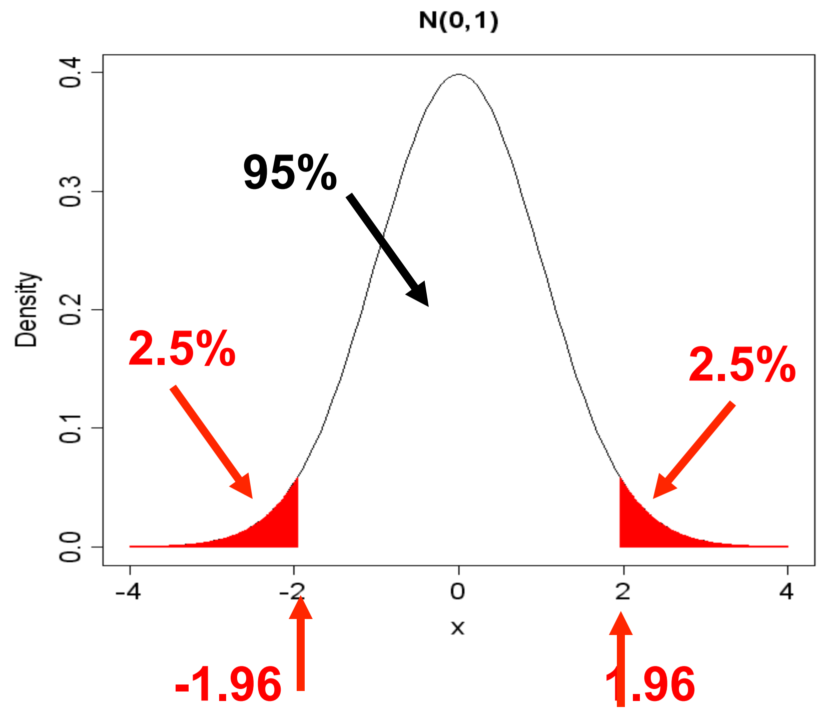
Varianca

Var(K)=n*p*(1-p)

Normalna porazdelitev

$$f(x \mid \mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

parametra N(μ, σ)



2,5 percentil

$z_{0,025}=-z_{0,975}$

97,5 percentil

$z_{0.975}=z_{1-0.05/2}$

STANDARDIZACIJA

- $X \sim N(\mu, \sigma)$, $Z=(X-\mu)/\sigma$
- $Z \sim N(0, 1)$

$z_{0,025}$: $P(Z <= z_{0,025}) = 0{,}025$

Tabele za N(0,1) – npr. glej

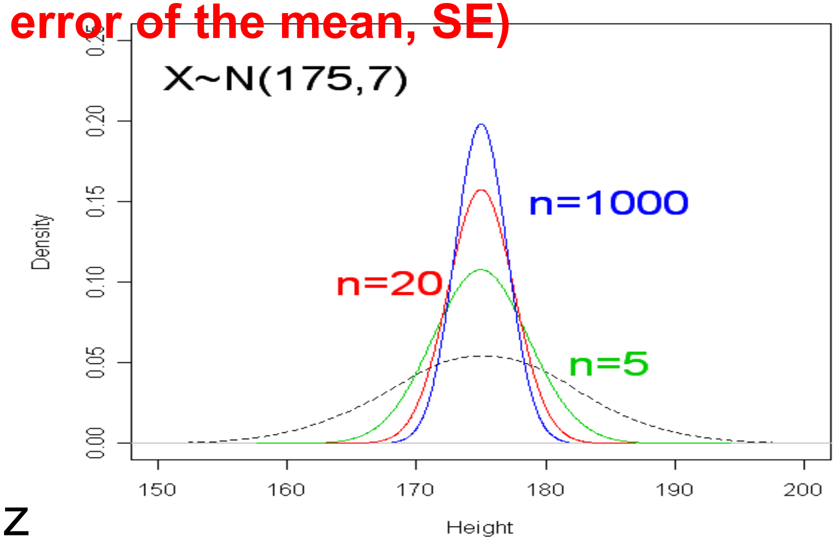
Distributions.xls

Pomembnost N(μ, σ) - CLI

- spremenljivka X merjena na neodvisnih enotah
- Ponavljaj velikokrat:
  - izberemo vzorec velikosti n, izračun $\bar{X}$
- Dobimo veliko povprečij $\bar{X}_i$
- Ta povprečja so porazdeljena po porazdelitvi
- ~ N(μ, σ/√n)

Standardna napaka (standard error of the mean, SE)

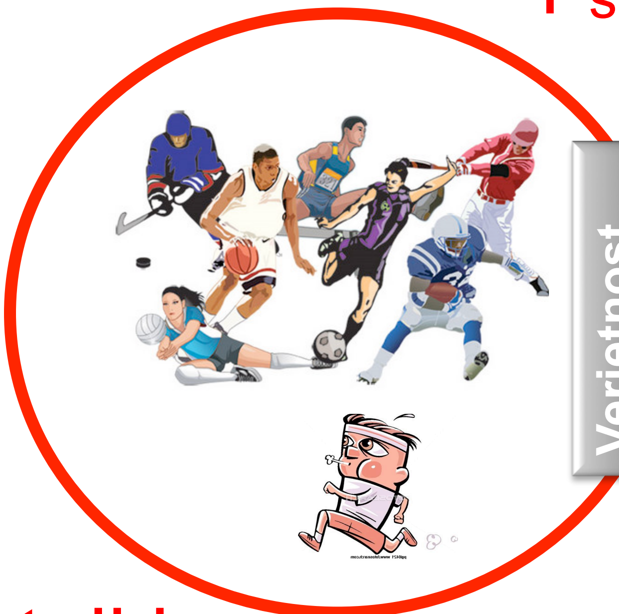
Povprečje spremenljivke X
je torej v populaciji
enako μ.
SE je odvisna od velikosti vzorca!



PRIMER intervali zaupanja za pulz

Raziskovalno vprašanje

??? $\mu_s = 60$???



Verjetnost
Inferenčna
statistika

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

športniki

Vzorčno povprečje: $\bar{X}_s$

Namen: Določiti, ali je povprečni pulz športnikov enak 60 udarcev na minuto.

Statistično sklepanje

SE

$$\bar{X}_s \sim N(\mu_s, \sigma_s/\sqrt{n_s})$$

- Ničelna hipoteza/domneva (H0): μ_s = **60**
- Alternativna hipoteza/domenva (Ha): $\mu_s \neq$ **60**
- Če H0 drži (in je populacijska varianca znana)

$$\bar{\mathbf{X}}_s \sim \mathbf{N(60}, \sigma_s/\sqrt{n_s}\mathbf{)} \qquad Z = \frac{\bar{X} - 60}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

- Kolikšna je **verjetnost**, da bi lahko ta standardizirani odmik od 0 (ali pa še večji odmik) nastal po naključju (če velja ničelna domneva)? p

Statistično sklepanje - primer

$\bar{X}_s \sim N(60, \sigma_s/\sqrt{n_s})$ $$Z = \frac{\overline{X} - 60}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$p = P(\overline{X} \leq -\bar{x} \bigcup \overline{X} \geq \bar{x} \mid H_0) = P(Z \leq -z \bigcup Z \geq z)$$

Primer:
67, 42, 53, 40, 55, σ=5

Vzorčno povprečje=51,4, n=5; SE= σ/√n =5/ √5=2,24

z=(51,4-60)/2,24=-3,85: testna statistika

Kakšen je p (verjetnost) pri tej testni statistiki?

Primer - nadaljevanje

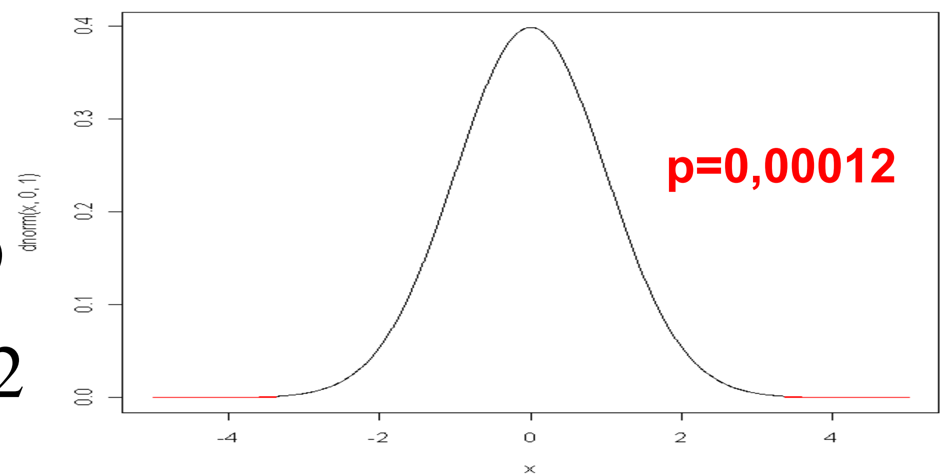
$$p = P(\overline{X} \le -\overline{x} \bigcup \overline{X} \ge \overline{x} \mid H_0) = P(Z \le -z \bigcup Z \ge z)$$

$$P(Z \le -z \bigcup Z \ge z) = p$$

$$= 2P(Z \le -z)$$

$$P(Z \le -3{,}85) = 0{,}000059$$

$$\text{p} = 2 * 0{,}000059 = 0{,}00012$$



Sklep:

zavrnemo H0: $\mu_s = 60$

μ_s povprečni pulz športnikov v populaciji

Interpretacija rezultatov

- **Kaj pomeni P=0,00012?**
  - če bi bil povprečen pulz v populaciji športnikov enak 60 (i.e. ničelna hipoteza drži) bi bilo zelo malo verjetno, da bi opazili na vzorcu tak ali večji odmik od 60, kot smo ga opazili.
- **Kaj naredimo?**
  - Zavrnemo ničelno domnevo in rečemo, da imajo športniki povprečen pulz, ki je različen od 60.
  - Ponavadi pravimo, da če je $p<0,05$ ali $p<0,01$, potem je rezultat "statistično značilen"

Interpretacija rezultatov

- **Kaj pomeni P=0,25?**
  - *"absence of evidence is not evidence of absence"*
- **Kaj bi naredili?**
  - Ne bi zavrnili ničelne domneve (≠sprejeli!!!). Rekli bi, da na podlagi naših podatkov, ne moremo trditi, da je povprečen pulz različen od 60. Rezultat ni statistično značilen.

Kaj lahko naredimo, če ne poznamo populacijske variance?

En vzorec: $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Varianco ocenimo na podlagi vzorca

Prej:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

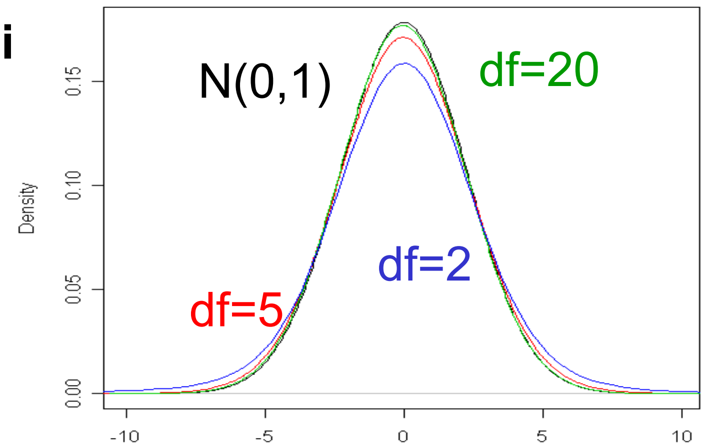
Potem:

$$\frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

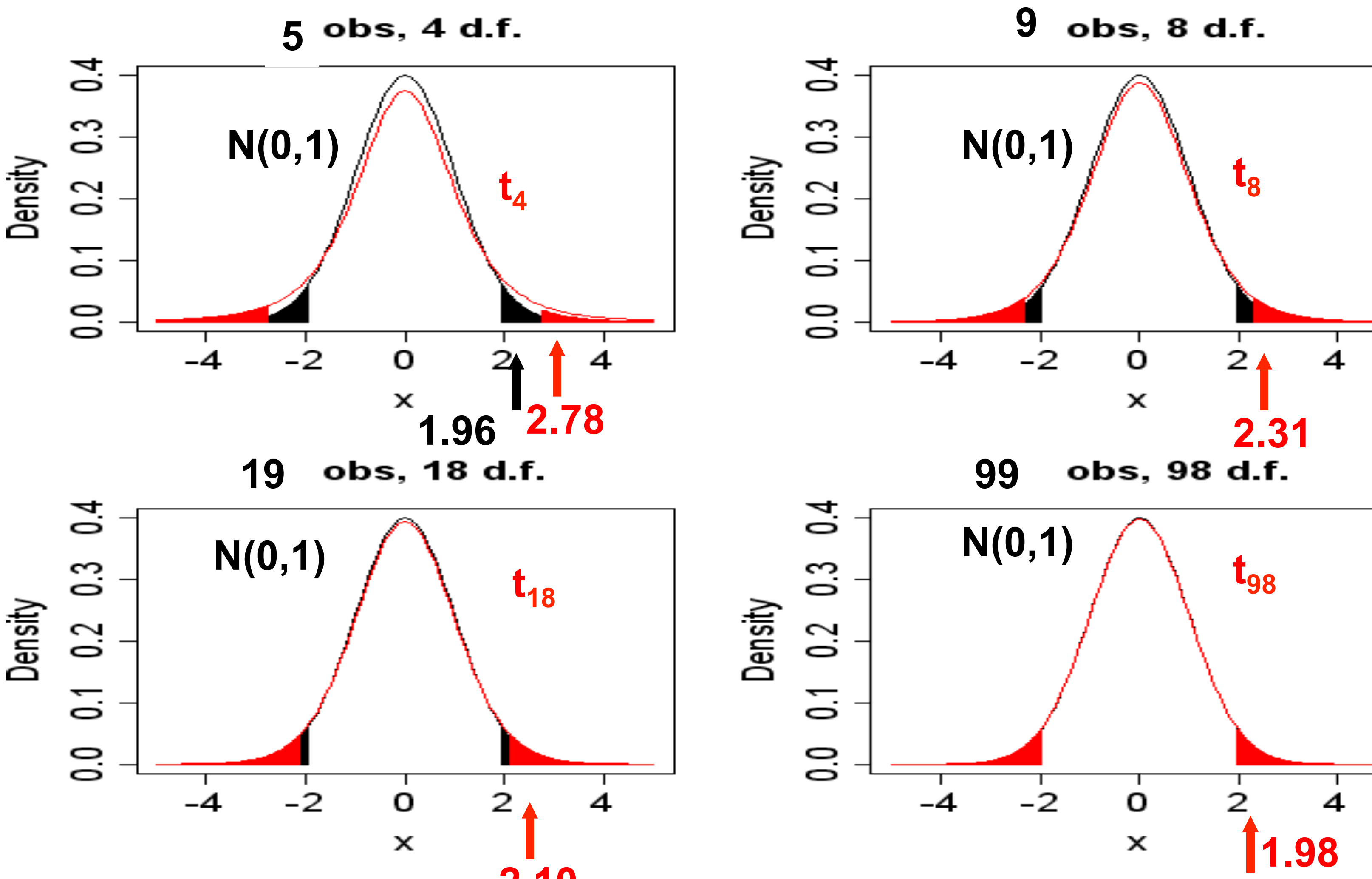
t-porazdelitev z n-1 stopinjami prostosti

Stopinje prostosti

(degrees of freedom, df)



Z in t... Ali sta si podobni?



Tudi za t-porazdelitev si bomo pomagali s tabelami za izračun verjetnosti

Raziskovalno vprašanje

??? $\mu_s = \mu_{ns}$???

Verjetnost
Inferenčna
statistika

športniki

ostali

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Vzorčno povprečje: $\overline{X}_s$

Vzorčno povprečje: $\overline{X}_{ns}$

Namen: Primerjati povprečni pulz športnikov in ostalih ljudi v populaciji

Statistično sklepanje

$\mu_s-\mu_{ns}=0$ v populaciji $\cancel{\xrightarrow{??}}$ na vzorcu $\underbrace{\bar{x}_s - \bar{x}_{ns}}^{\text{odmik na vzorcu}}=0$

Ničelna domneva:

Kolikšna je **verjetnost**, da bi lahko ta odmik* (ali pa še večji odmik) nastal po naključju?

P-vrednost (p) (stopnja tveganja)

Če je verjetnost majhna, bomo zavrnili (ničelno) domnevo, da v populaciji ni razlike

Tipična odločitev: $p< 0{,}05$ ---> $\mu_s - \mu_{ns} \neq 0$ ("statistično značilna razlika")

(zavrnemo domnevo, da $\mu_s - \mu_{ns} = 0$ in s 95 % gotovostjo trdimo, da gre za dejansko razliko v povprečnem pulzu; razlika med pulzoma ni 0)

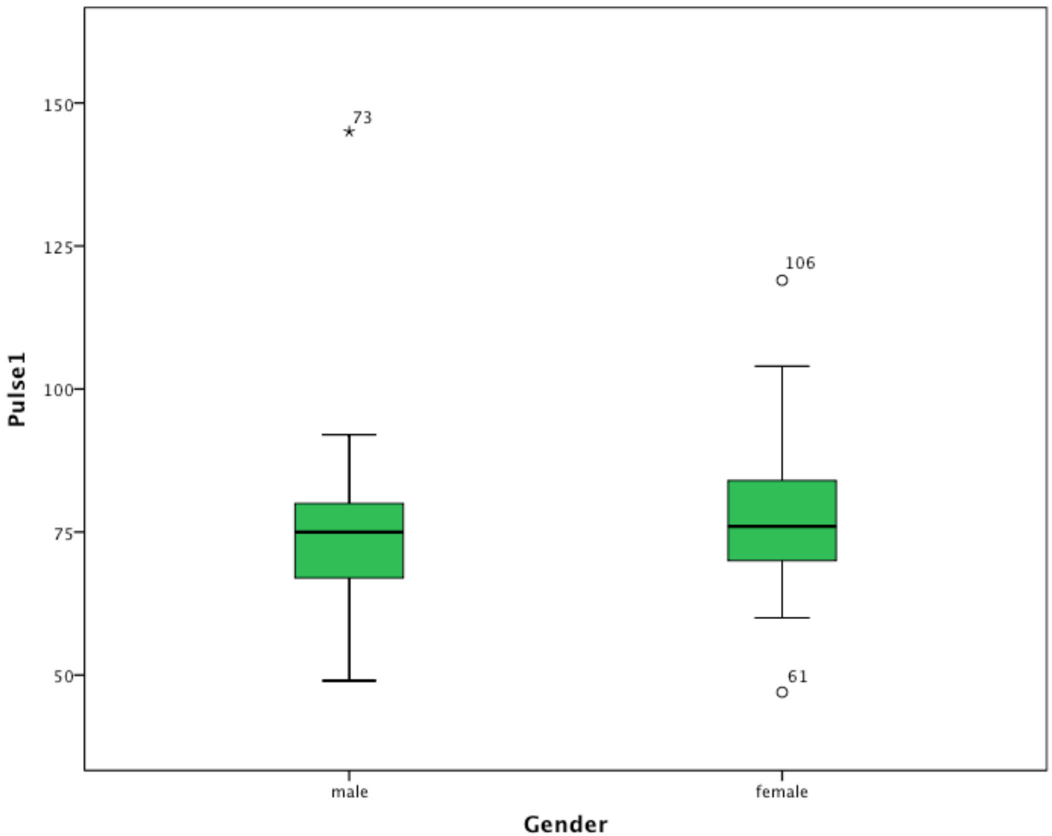
$\alpha=0{,}05$ je **stopnja značilnosti**

* Odmik bomo standardizirali, glede na velikosti vzorca in na razpršenosti spremenljivke v populaciji

Ali imajo v povprečju dekleta in fantje v mirovanju isti pulz?

Povprečje: **74,1 (m)** **77,5 (f)**

Standardni odklon: **13,8 (m)** **12,6 (f)**



Populacija: študentje UQ

$\mu_f - \mu_m = 0$

Ničelna domneva

Vzorec: študentje Osnov stat.

$\bar{x}_f - \bar{x}_m = 3{,}4$

Kako bomo primerjali vzorčna povprečja dveh vzorcev?

En vzorec: $\bar{X}_s \sim N(\mu_s , \sigma_s/\sqrt{n_s})$

Dva vzorca, če sta populacijski varianci znani

SE

$$\bar{X}_s - \bar{X}_{ns} \sim N\left(\mu_s - \mu_{ns} , \sqrt{\sigma^2_s/n_s + \sigma^2_{ns}/n_{ns}}\right)$$

Dva vzorca, če populacijski varianci nista znani in sta enaki– bolj realistična situacija

Porazdelitev

Testna statistika t

df

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n1+n2-2}$$

Standardna napaka

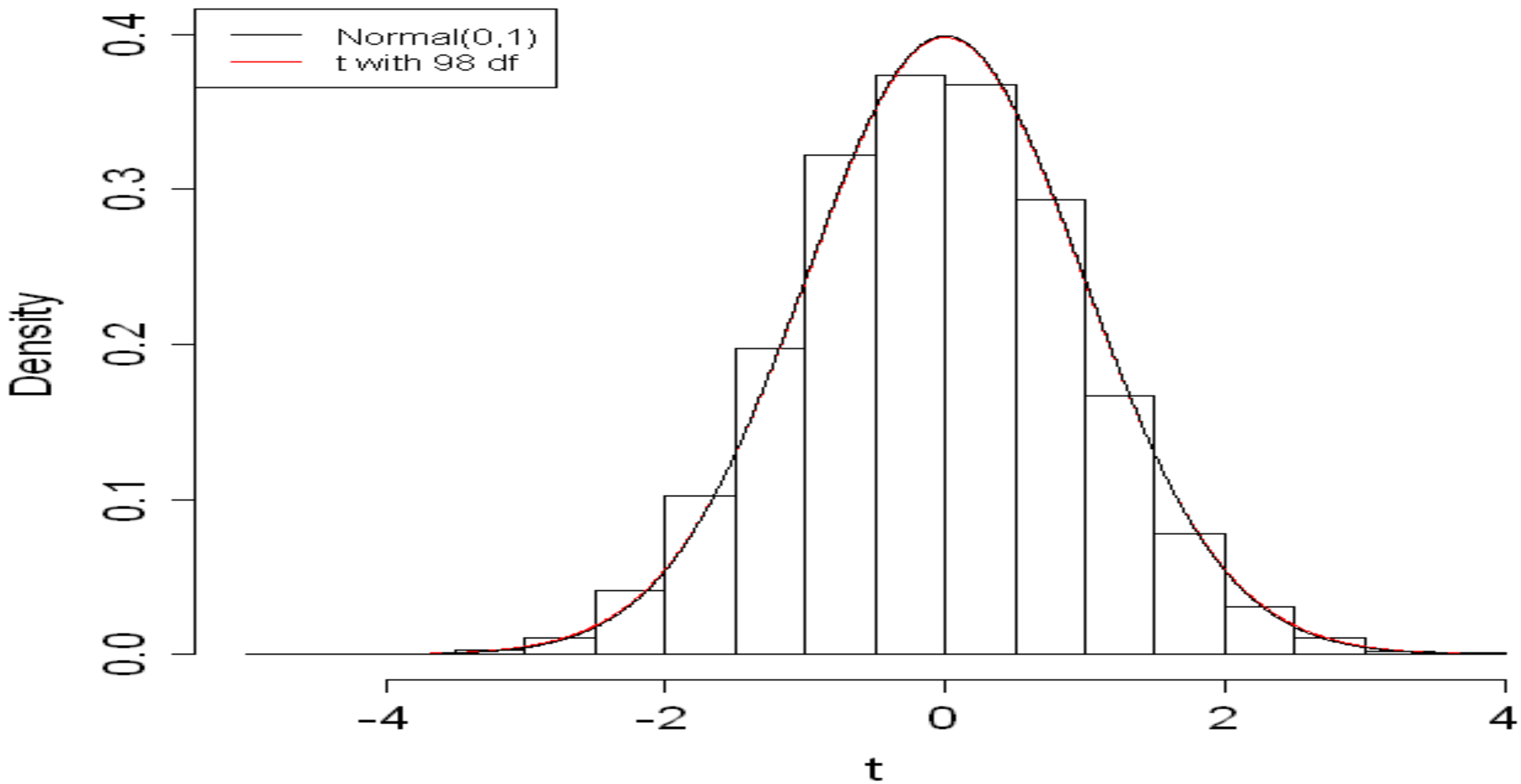
(SE: standard error)

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

Skupna varianca

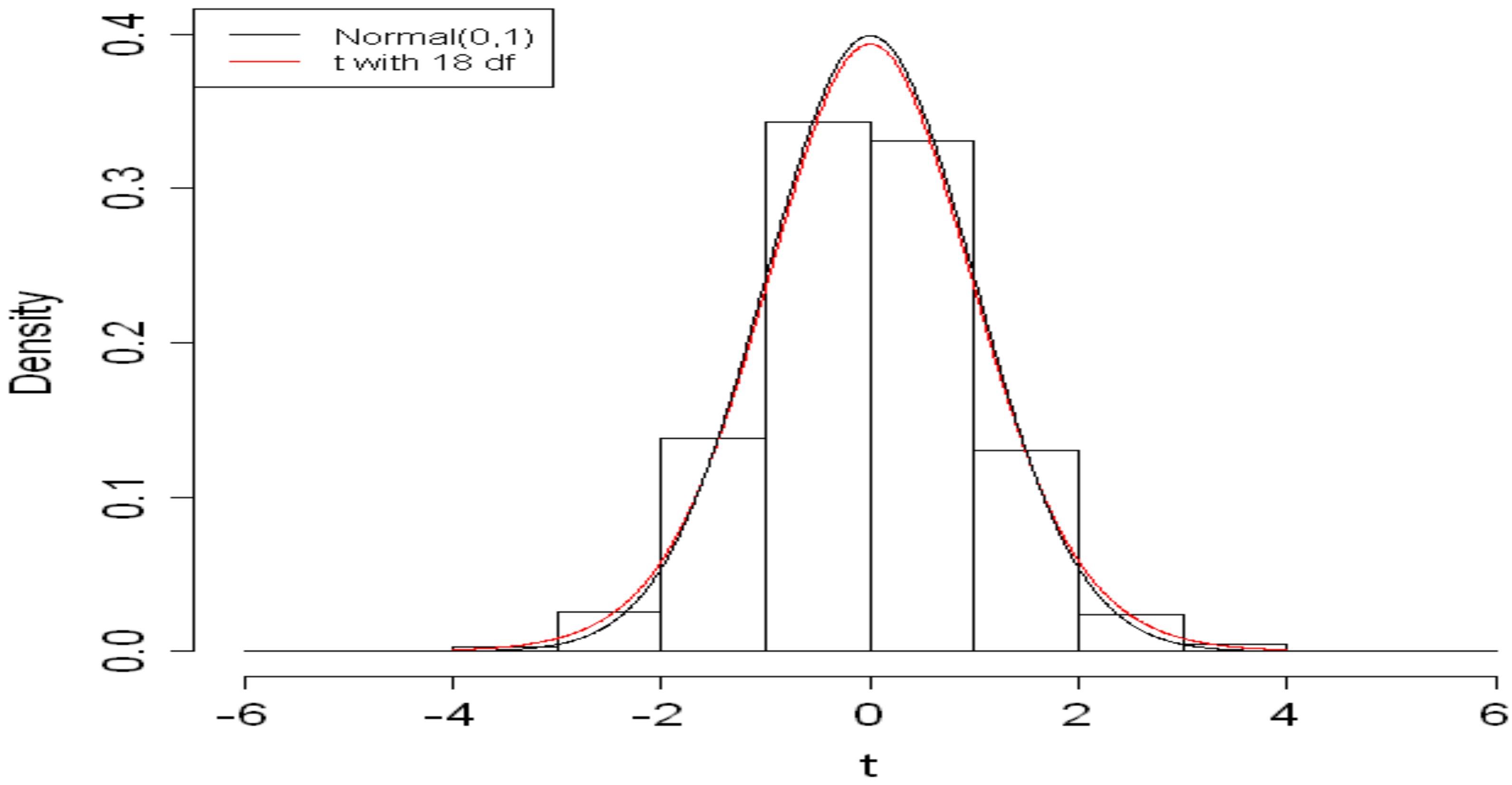
(pooled variance)

Kako se porazdeli t, če $\mu_1 = \mu_2$ in $n_1 = n_2 = 50$?

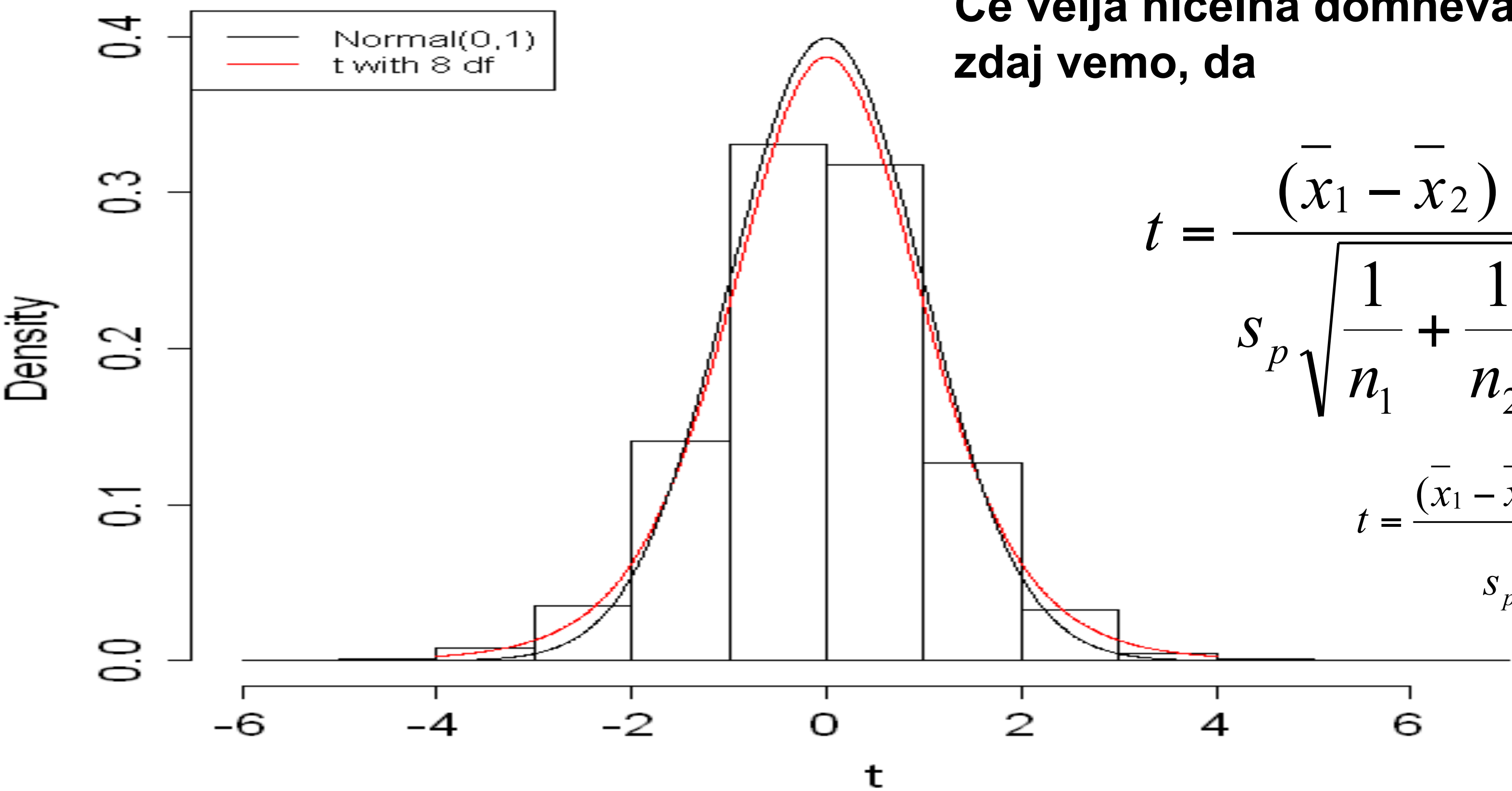


X~N(μ, σ) -> Z=(X- μ)/ σ ~ N(0, 1) – standardna normalna porazdelitev

Kako se porazdeli t, če $\mu_1 = \mu_2$ in $n_1 = n_2 = 10$?



Kako se porazdeli t, če $\mu_1 = \mu_2$ in $n_1 = n_2 = 5$?



Če velja ničelna domneva ($\mu_1 = \mu_2$) zdaj vemo, da

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

	povpr.	SD	n (število študentov)
študentke	77,5	12,6	50
študenti	74,1	13,8	59

Skupna varianca (1)

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

$$= \sqrt{\frac{(50-1)\cdot 12{,}6^2+(59-1)\cdot 13{,}8^2}{(50-1)+(59-1)}} = 13{,}3$$

Standardna napaka (2)

$$SE = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 13{,}3\sqrt{\frac{1}{50}+\frac{1}{59}} = 2{,}56$$

Razlika povprečij (3)

$$\bar{x}_1 - \bar{x}_2 = 77{,}5 - 74{,}1 = 3{,}4$$

Stopinje prostosti (4)

$$df = n_1 + n_2 - 2 = 50 + 59 - 2 = 107$$

Testna statistika (5)

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE} \sim t_{n_1+n_2-2}$$

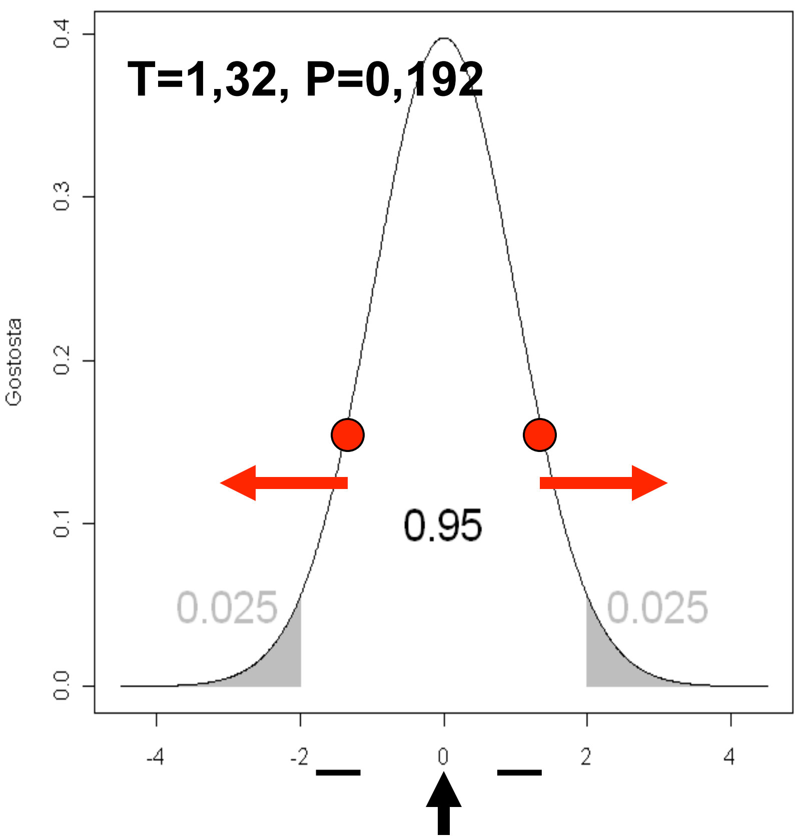
H_0: $\mu_1 = \mu_2$ ničelna domneva

H_a: $\mu_1 \neq \mu_2$ alternativna domneva

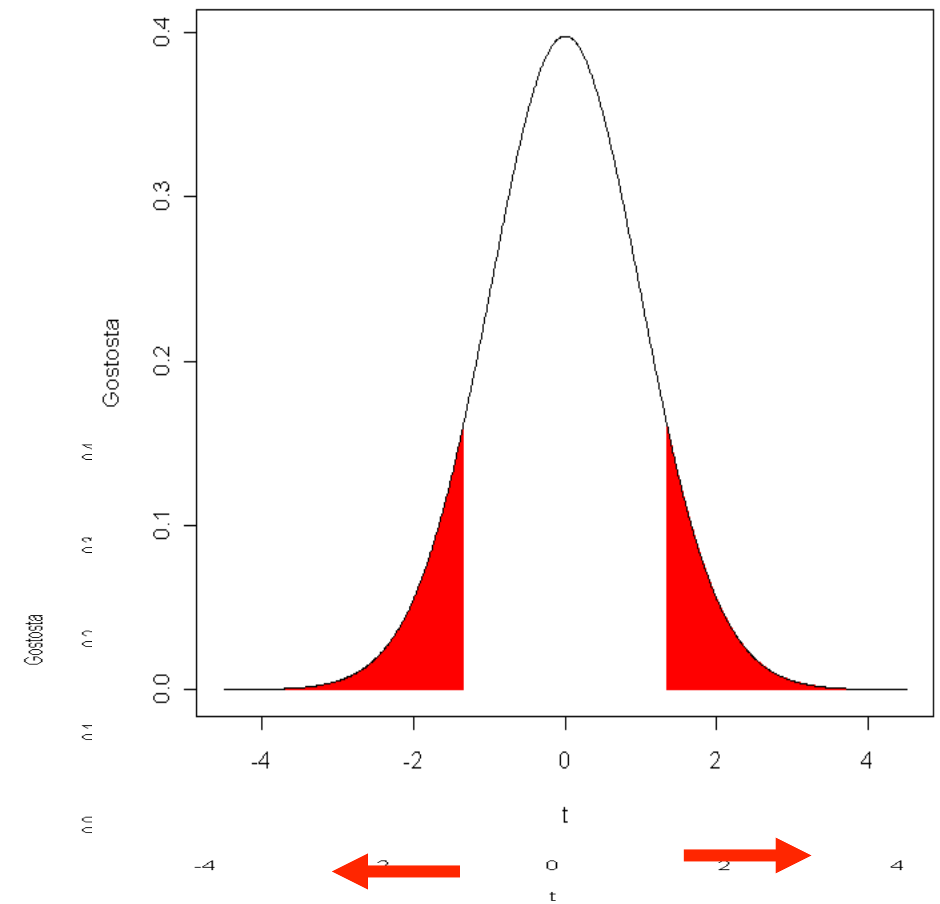
$$t = \frac{3{,}4-0}{2{,}56} = 1{,}32 \overset{H_0}{\sim} t_{107}$$

Kako pridemo iz testne statistike do p-vrednosti?

t_{107}



$$\bar{x}_1 = \bar{x}_2$$



Pričakujemo 95% srednjih vrednosti med -1.99 in 1.99

Predpostavke za t-test za neodvisna vzorca

1. Porazdelitev spremenljivke v populaciji
   - **Majhen vzorec**: spremenljivke so normalno porazdeljene – če niso obstajajo neparametrični testi (Mann-Whitney U / Wilcoxon rank-sum test)
   - **Velik vzorec**: lahko uporabljamo test ne glede na porazdelitev
2. Statistične enote so neodvisne
3. Populacijski varianci sta enaki $\sigma_1 = \sigma_2$
   - Lahko testiramo H0: $\sigma_1 = \sigma_2$ (**Levenov test**)
   - Če nista lahko uporabljamo **Welchev t-test**

$$T = \frac{\bar{Y_1} - \bar{Y_2}}{\sqrt{s_1^2/N_1 + s_2^2/N_2}} \sim t_v$$

$$v = \frac{(s_1^2/N_1 + s_2^2/N_2)^2}{(s_1^2/N_1)^2/(N_1-1) + (s_2^2/N_2)^2/(N_2-1)} \quad \textbf{d.f.}$$

H0: $\sigma_f^2 / \sigma_m^2 = 1$ Naš primer: $s_f^2 / s_m^2 = 1{,}09$, P=0,798 (Levenov test)

Postopek statističnega sklepanja

- Izberite znanstveno vprašanje
  - Populacija
  - Vzorec
  - Spremenljivka/e
  - Kaj primerjati
- Postavite ničelno domnevo in alternativno domnevo
- Izberite testno statistiko, (statistični test, ki ga boste boste uporabili)
  - Določite porazdelitev testne statistike, če velja ničelna domneva
- Izberite stopnjo tveganja (α)
- Izračunajte vrednost testne statistike iz vašega vzorca
- Izračunajte P-vrednost in odgovorite na znanstveno vprašanje

- Ali imajo študentke višji pulz v mirovanju kot študenti?
  - Populacija : študentje UQ
  - Vzorec : študenti pri Osnovah stat.
  - Spremenljvke
    - spol
    - pulz v mirovanju
  - Primerjava: povprečen pulz

H_0: $\mu_f = \mu_m$ ničelna domneva

H_a: $\mu_f \neq \mu_m$ alternativna domneva

$$t = \frac{(\bar{x}_f - \bar{x}_m) - (\mu_f - \mu_m)}{SE} \sim t_{n_f + n_m - 2}$$

$$t = \frac{3,4 - 0}{2,56} = 1,32 \overset{H_0}{\sim} t_{107}$$

P=0,192, ne zavrnemo ničelne domneve

Primer 2: t-test za dva neodvisna vzorca

Bobath or Motor Relearning Programme? A comparison of two different approaches of physiotherapy in stroke rehabilitation: a randomized controlled study

Birgitta Langhammer Department of Medicine and Physiotherapy, Bærum Hospital and **Johan K Stanghelle** Sunnaas Rehabilitation Hospital, Nesoddtangen, Norway

Received 28th October 1999; returned for revisions 17th December 1999; revised manuscript accepted 2nd February 2000.

Objective: To examine whether two different physiotherapy regimes caused any differences in outcome in rehabilitation after acute stroke.
Design: A double-blind study of patients with acute first-ever stroke. Sixty-one patients were consecutively included, block randomized into two groups, and stratified according to gender and hemiplegic site. Group 1 (33 patients) and group 2 (28 patients) had physiotherapy according to Motor Relearning Programme (MRP) and Bobath, respectively. The supplemental treatment did not differ in the two groups.
Main outcome measures: The Motor Assessment Scale (MAS), the Sødring Motor Evaluation Scale (SMES), the Barthel ADL Index and the Nottingham Health Profile (NHP) were used. The following parameters were also registered: length of stay in the hospital, use of assistive devices for mobility, and the patient's accommodation after discharge from the hospital.
Results: Patients treated according to MRP stayed fewer days in hospital than those treated according to Bobath (mean 21 days versus 34 days, $p = 0.008$).

Table 1 MAS scores (mean and range) at the three tests

	Test 1 Mean/median	Test 2 Mean/median	Test 3 Mean/median	Test 1 versus test 2, *p*-value Student's *t*-test	Test 1 versus test 3, *p*-value Student's *t*-test
Group 1 – MRP	24/29	32/40	37/42	0.0001	0.0001
Range	0–41	3–47	7–48		
SD	14	15	12		
Group 2 – Bobath	19/20	23/25	33/39	0.0001	0.0001
Range	0–42	0–46	2–48		
SD	15	16	15		
Group 1/group 2, *p*-value Student's *t*-test	0.44	0.05	0.31		

Test 2	Mean	Standard deviation	n
Group 1	32	15	30
Group 2	23	16	27

Test 2	Mean	Standard deviation	n
Group 1	32	15	30
Group 2	23	16	27

$$H0: \mu_{G1} = \mu_{G2}$$

$$Ha: \mu_{G1} \neq \mu_{G2}$$

$$s_p = \sqrt{\frac{s_{G1}^2(n_{G1}-1)+s_{G2}^2(n_{G2}-1)}{n_{G1}+n_{G2}-2}}$$

$$SE = s_p\sqrt{\frac{1}{n_{G1}}+\frac{1}{n_{G2}}}$$

$$\frac{(\overline{X}_{G1}-\overline{X}_{G2})-(\mu_{G1}-\mu_{G2})}{SE} \sim t_{n_{G1}+n_{G2}-2}$$

$$pri \cdot H0 \cdot velja$$

$$\frac{\overline{X}_{G1}-\overline{X}_{G2}}{SE} \sim t_{n_{G1}+n_{G2}-2}$$

$$s_p = \sqrt{\frac{15^2(30-1)+16^2(27-1)}{30+27-2}} = 15{,}52$$

$$SE = 15{,}52\sqrt{\frac{1}{30}+\frac{1}{27}} = 4{,}19$$

$$pri \cdot H0 \cdot velja$$

$$t = \frac{32 \;\; -23}{4{,}19} = \frac{9}{4{,}19} = 2{,}15 \sim t_{30+27-2}$$

$$p = P(t_{55} < -|t| \bigcup t_{55} > |t|) = 2P(t_{55} > |t|) =$$

$$2(t_{55} > 2{,}15) = 0{,}036$$

Kaj lahko sklepamo?

Ali bi lahko izračunali tudi 95% interval zaupanja za razliko povprečij?

95% interval zaupanja za $\mu_1 - \mu_2$

Interval za katerega imamo 95% zaupanje, da vsebuje neznano populacijsko razliko povprečij

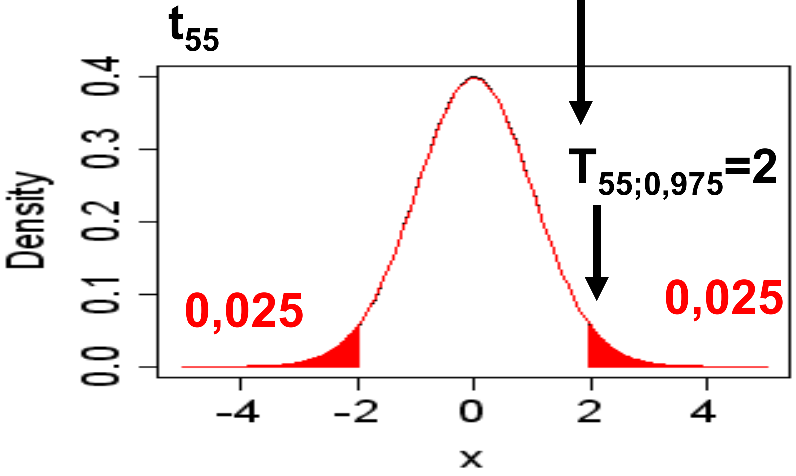
$$95\%\ \text{IZ za}\ \ \mu_{G1} - \mu_{G2}:$$

$$(\bar{x}_{G1} - \bar{x}_{G2}) - t_{55;1-0,05/2} \cdot SE, (\bar{x}_{G1} - \bar{x}_{G2}) + t_{55;1-0,05/2} \cdot SE$$

$$\bar{x}_{G1} - \bar{x}_{G2} = 9 \quad SE = 4,19$$

$$9 - 2 \cdot 4,19 \ \text{do}\ 9 + 2 \cdot 4,19$$

$$0,62 \ \text{do}\ \ 17,38$$



$$t_{df,1-\alpha/2} : P(t_{df} \leq 1 - \frac{\alpha}{2}) = 1 - \alpha$$

S 95% zaupanjem lahko trdimo, da je razlika populacijskih povprečij

($\mu_{G1} - \mu_{G2}$) med 0,62 in 17,38

Predpostavka 2.
Statistične enote so (ne)odvisne?

Table 1 MAS scores (mean and range) at the three tests

	Test 1 Mean/median	Test 2 Mean/median	Test 3 Mean/median	Test 1 versus test 2, *p*-value Student´s *t*-test	Test 1 versus test 3, *p*-value Student´s *t*-test
Group 1 – MRP	24/29	32/40	37/42	0.0001	0.0001
Range	0–41	3–47	7–48		
SD	14	15	12		
Group 2 – Bobath	19/20	23/25	33/39	0.0001	0.0001
Range	0–42	0–46	2–48		
SD	15	16	15		
Group 1/group 2, *p*-value Student´s *t*-test	0.44	0.05	0.31		

Za isto skupino (iste enote) primerjamo 2 izhodna testa (takoj po kapi in 2 tedna po njej.

Uporabimo t-test za odvisna vzorca (parni t-test).

Parni t-test (t-test za dva odvisna vzorca)

Ali se pulza pred in po obremenitvi statistično značilno razlikujeta.

	Razlika D (po – pred)
s1	68
s2	80
s3	63
s4	12
...	...
s46	26
Povprečje	51,4
SD	21,1
SE=SD/√n	3,1

H0: $\mu_{po} - \mu_{pred} = \delta = 0$

$$\frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1} \qquad \frac{\overline{D} - \delta}{\frac{s_d}{\sqrt{n}}} \sim t_{n-1}$$

Testna statistika

T=51,4/3,1=16,52, df=45

p= $9{,}8 \cdot 10^{-21}$

$t_{45;\ 1-0,05/2}$=2,01

95% IZ za δ: 51,4-2,01*3,1 do 51,4+2,01*3,1 =45,1 do 57,7 udarcev na minuto

Predpostavke za t-test za odvisna vzorca

1. Porazdelitev spremenljivke v populaciji
   - **Majhen vzorec**: spremenljivke so normalno porazdeljene – če niso obstajajo neparametrični testi (Wilcoxon signed-rank test)
   - **Velik vzorec**: lahko uporabljamo test ne glede na porazdelitev

- Statistične enote so odvisne

Postopek statističnega sklepanja

- Izberite znanstveno vprašanje
  - Populacija
  - Vzorec
  - Spremenljivka/e
  - Kaj primerjati
- Postavite ničelno domnevo in alternativno domnevo
- Izberite testno statistiko, (statistični test, ki ga boste boste uporabili)
  - Določite porazdelitev testne statistike, če velja ničelna domneva
- Izberite stopnjo tveganja (α)
- Izračunajte vrednost testne statistike iz vašega vzorca
- Izračunajte P-vrednost in odgovorite na znanstveno vprašanje

- Ali imajo študenti višji pulz po obremenitvi?
  - Populacija : študentje UQ
  - Vzorec : študenti pri Osnovah stat.
  - Spremenljivke
    - pulz v mirovanju
    - pulz po obremenitvi
  - Primerjava:
    - povprečna pulza istih študentov

H_0: $\mu_{pred} = \mu_{po}$ ničelna domneva

H_a: $\mu_{pred} \neq \mu_{po}$ alternativna domneva

$$\frac{\overline{D} - \delta}{\frac{s_d}{\sqrt{n}}} \sim t_{n-1}$$

$$t = \frac{51,4 - 0}{3,1} = 16,52 \overset{H_0}{\sim} t_{45}$$

P<0,001, zavrnemo ničelno domnevo