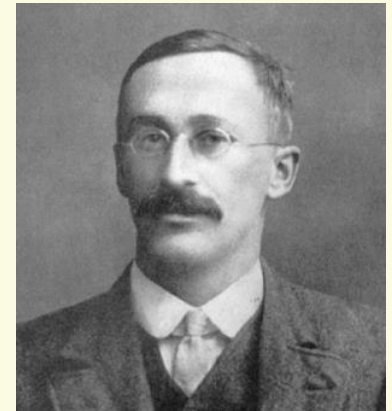


# T-TEST

*T-test* je statistički postupak koji je osmislio Wiliam Sealy Gosset-Student, a kojim se utvrđuje statistička značajnost razlike između dviju aritmetičkih sredina.

*Statistički značajna razlika* je razlika utvrđena na uzorku entiteta, a za koju s visokim stupnjem sigurnosti (95% ili 99%) možemo tvrditi da se nije dogodila slučajno (kao posljedica slučajnog variranja aritmetičkih sredina uzoraka u odnosu na aritmetičku sredinu populacije).



Wiliam Sealy Gosset  
(1876. – 1937.)

# T-TEST

---

*T-test za nezavisne uzorke* je postupak kojim se utvrđuje statistička značajnost razlike između aritmetičkih sredina dvaju uzoraka (dvaju nezavisnih skupova entiteta).

Pri tome je moguće postaviti sljedeću alternativnu (H1) odnosno nultu (H0) hipotezu:

**H1:**  $\bar{x}_1 \neq \bar{x}_2$  - Razlika između aritmetičkih sredina analiziranih uzoraka je statistički značajna uz pogrešku  $p$ .

**H0:**  $\bar{x}_1 = \bar{x}_2$  - Uz pogrešku  $p$  ne možemo tvrditi da je razlika između aritmetičkih sredina analiziranih uzoraka statistički značajna.

# T-TEST

---

*T-test za zavisne uzorke* je postupak kojim se utvrđuje statistička značajnost razlike između aritmetičkih sredina jednog uzorka mjenenog u dvije vremenske točke.

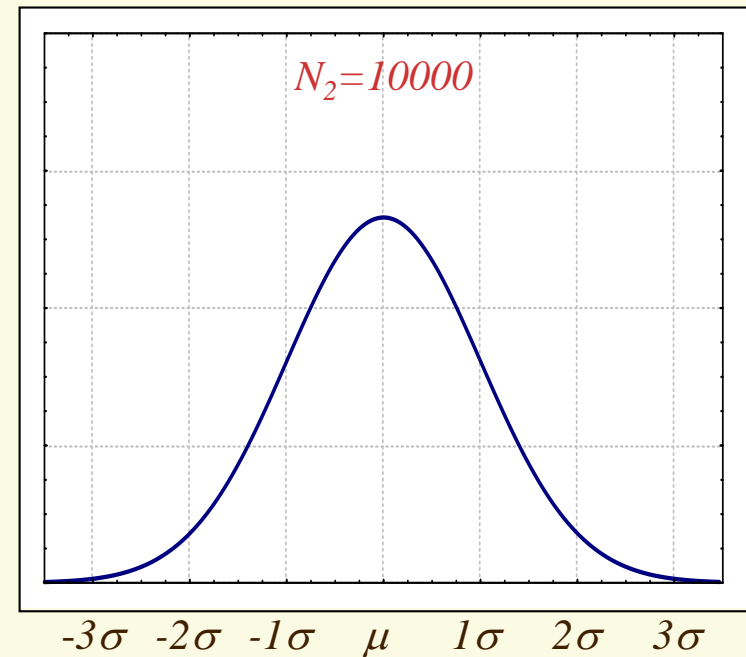
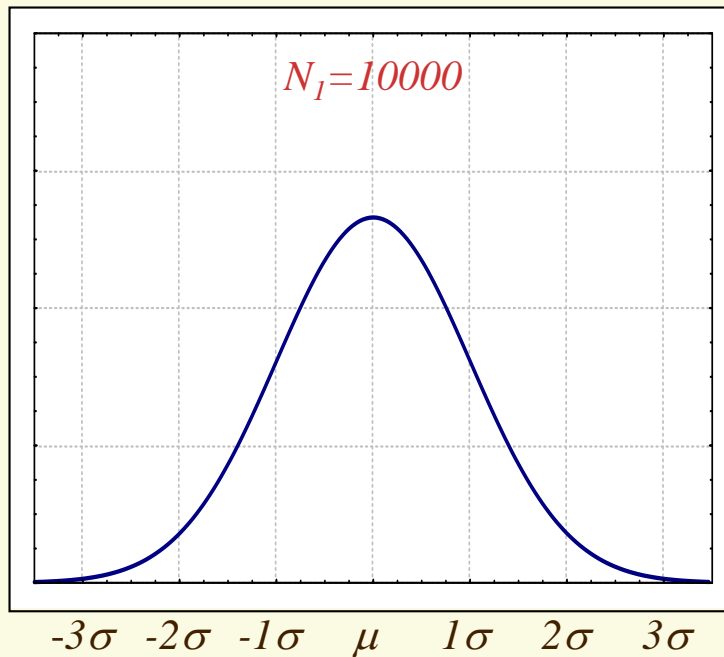
Pri tome je moguće postaviti sljedeću alternativnu (H1) odnosno nultu (H0) hipotezu:

**H1:**  $\bar{x}_1 \neq \bar{x}_2$  - Razlika između aritmetičkih sredina prvog i drugog mjerenja je statistički značajna uz pogrešku  $p$ .

**H0:**  $\bar{x}_1 = \bar{x}_2$  - Uz pogrešku  $p$  ne možemo tvrditi da je razlika između aritmetičkih sredina prvog i drugog mjerenja statistički značajna.

# T-TEST

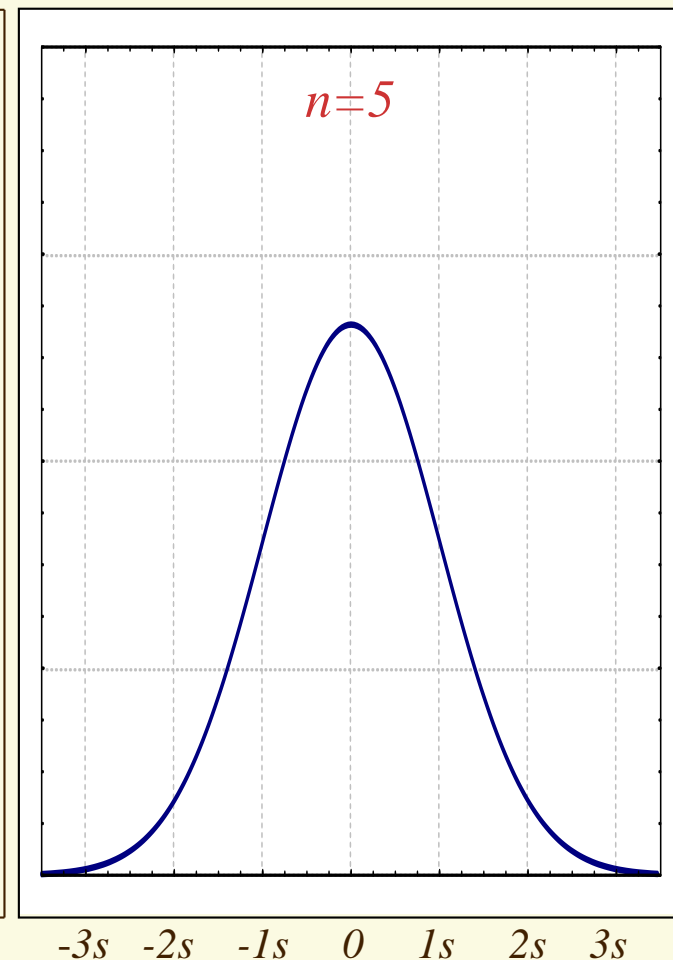
Pretpostavimo da su u neke dvije konačne populacije ( $N_1=10000$  i  $N_2 = 10000$ ) aritmetičke sredine ( $\mu_1$  i  $\mu_2$ ) neke normalno distribuirane varijable  $X$  potpuno jednake, tj. razlika između aritmetičkih sredina ne postoji.



# T-TEST

Pretpostavimo da iz prethodno opisanih populacija formiramo veliki broj parova slučajnih uzoraka (npr. 10000) veličine 5 entiteta te izračunamo razliku aritmetičkih sredina svakog para. Da li su sve razlike jednake nuli? Kako su distribuirane “slučajne” razlike između aritmetičkih sredina?

$$\begin{array}{l} \bar{x}_1 - \bar{x}_2 = d_1 \\ \bar{x}_1 - \bar{x}_2 = d_2 \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \bar{x}_1 - \bar{x}_2 = d_{10000} \end{array}$$



# T-TEST

---

U slučaju da između aritmetičkih sredina populacija ne postoji razlika na temelju prethodno izračunatih razlika između aritmetičkih sredina slučajnih uzoraka možemo zaključiti sljedeće:

- razlika između aritmetičkih sredina slučajnih uzoraka nije uvijek jednaka nuli
- razlike između aritmetičkih sredina slučajnih uzoraka variraju
- aritmetička sredina “slučajnih” razlika, tj. razlika aritmetičkih sredina uzoraka tendirat će nuli
- distribucija razlika između aritmetičkih sredina slučajno odabranih uzoraka iste veličine bit će *normalna*.

# T-TEST

Pretpostavimo da iz istih populacija formiramo 10000 parova slučajnih uzoraka veličine 50 entiteta te izračunamo razliku aritmetičkih sredina svakog para. Da li razlike aritmetičkih sredina uzoraka veličine 50 entiteta variraju više ili manje nego razlike aritmetičkih sredina uzoraka veličine 5 entiteta?

$$\bar{x}_1 - \bar{x}_2 = d_1$$

$$\bar{x}_1 - \bar{x}_2 = d_2$$

• • •

• • •

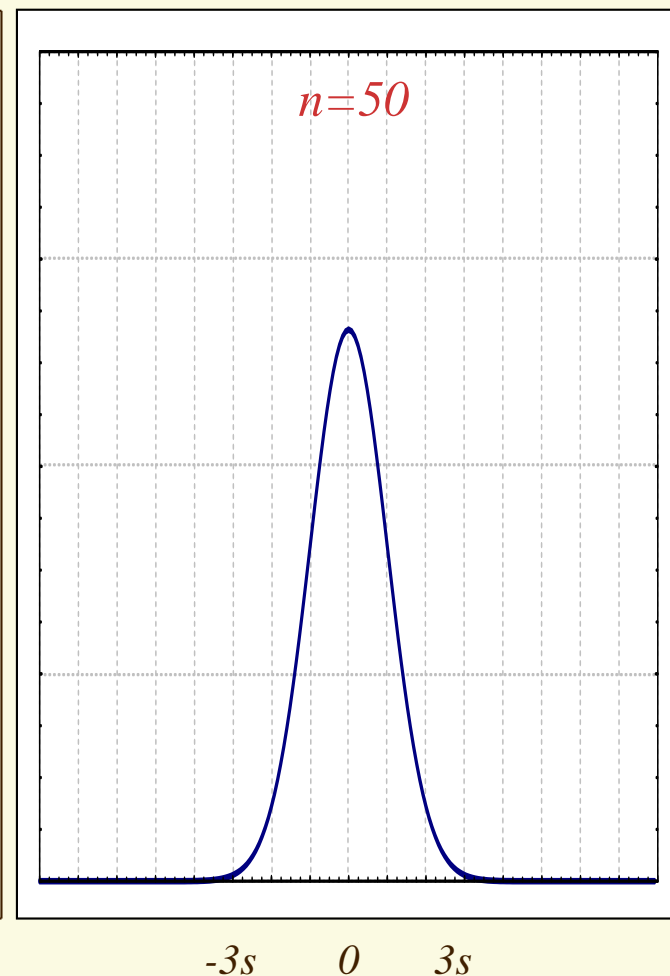
• • •

• • •

• • •

• • •

$$\bar{x}_1 - \bar{x}_2 = d_{10000}$$



# T-TEST

Usporedbom varijabli razlika između aritmetičkih sredina parova slučajnih uzoraka veličine 5 entiteta i parova slučajnih uzoraka veličine 50 entiteta moguće je zaključiti da je standardna devijacija “slučajnih” razlika manja što je broj entiteta u uzorku veći.

Osim o broju entiteta u uzorku, standardna devijacija “slučajnih” razlika zavisi i o varijabilnosti istraživanog obilježja (varijable) u populaciji. Što je varijabilnost obilježja u populaciji veća to će i “slučajne” razlike više varirati.

Standardna devijacija “slučajnih” razlika naziva se *standardna pogreška razlika aritmetičkih sredina*, a označava se simbolom  $s_{\bar{x}_1 - \bar{x}_2}$ .



# T-TEST

Standardna pogreška razlika aritmetičkih sredina u *t-testu za nezavisne uzorke* izračunava se formulom

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

gdje je

- ➔  $s_1$  - standardna devijacija prvog uzorka
- ➔  $s_2$  - standardna devijacija drugog uzorka
- ➔  $n_1$  - broj entiteta u prvom uzorku
- ➔  $n_2$  - broj entiteta u drugom uzorku

# T-TEST

Standardna pogreška razlika aritmetičkih sredina u *t-testu za zavisne uzorke* izračunava se formulom

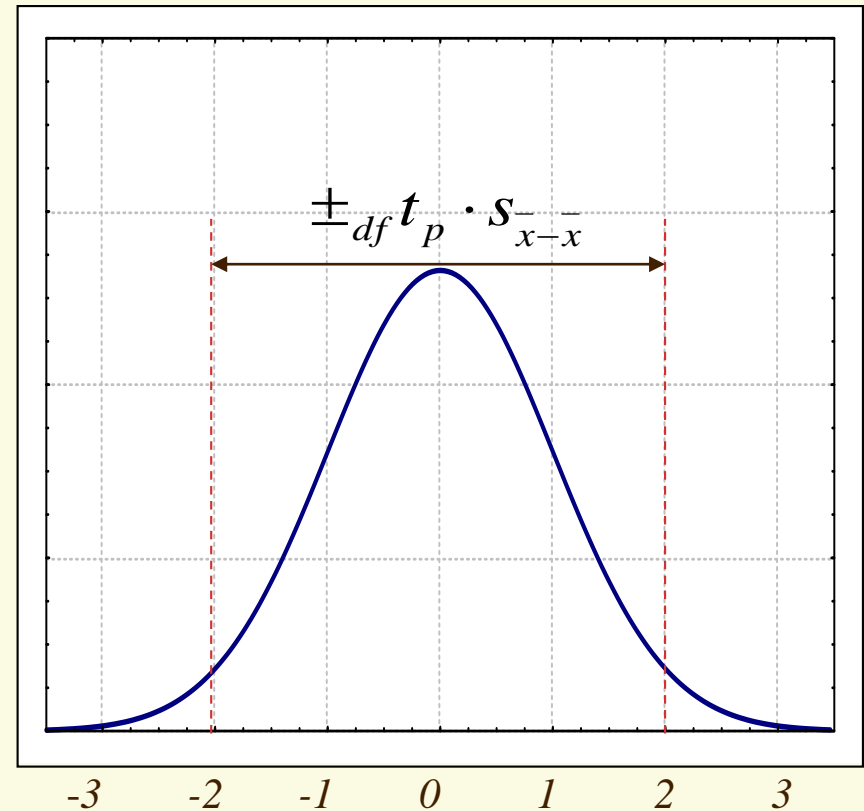
$$s_{x_1 - x_2} = \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2 \cdot r \cdot \frac{s_1}{\sqrt{n}} \cdot \frac{s_2}{\sqrt{n}}}$$

gdje je

- ➔  $s_1$  - standardna devijacija prvog mjerenja
- ➔  $s_2$  - standardna devijacija drugog mjerenja
- ➔  $n$  - broj entiteta u uzorku
- ➔  $r$  - korelacija između varijabli prvog i drugog mjerenja

# T-TEST

Ako je razlika između dviju aritmetičkih sredina  $d_f t_p$  puta veća od standardne pogreške razlika aritmetičkih sredina onda je vjerojatnost da razlika ne postoji u populaciji manja od  $p$ . Omjer razlike između dviju aritmetičkih sredina i standardne pogreške razlika naziva se *t-omjer* ili *t-vrijednost*, označava se s  $t$ , a testira se putem t-distribucije.



# T-TEST

*T-omjer*, tj. *t-vrijednost* ( $t$ ) izračunava se jednako u t-testu za nezavisne uzorke i u t-testu za zavisne uzorke i to pomoću sljedeće formule

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

gdje je

- ➔  $\bar{x}_1$  - aritmetička sredina prvog uzorka/mjerenja
- ➔  $\bar{x}_2$  - aritmetička sredina drugog uzorka/mjerenja
- ➔  $s_{\bar{x}_1 - \bar{x}_2}$  - standardna pogreška razlika aritmetičkih sredina

# T-TEST

U *t*-testu za nezavisne uzorke kritična *t*-vrijednost ( $df t_p$ ) određuje se na temelju broja stupnjeva slobode  $df=n_1-1+n_2-1$  i pogreške statističkog zaključka *p* (npr.  $p<0,05$  ili  $p<0,01$ ).

Testiranje hipoteza vrši se na sljedeći način:

$$|t| < df t_p \Rightarrow H_0: \bar{x}_1 = \bar{x}_2$$

( $H_0$ : Uz pogrešku *p* ne možemo tvrditi da je razlika između aritmetičkih sredina analiziranih uzoraka statistički značajna)

$$|t| > df t_p \Rightarrow H_1: \bar{x}_1 \neq \bar{x}_2$$

( $H_1$ : Razlika između aritmetičkih sredina analiziranih uzoraka statistički je značajna uz pogrešku *p*)

# T-TEST

U *t*-testu za zavisne uzorke kritična *t*-vrijednost ( $_{df}t_p$ ) određuje se na temelju broja stupnjeva slobode  $df=n-1$  i pogreške statističkog zaključka  $p$  (npr.  $p<0,05$  ili  $p<0,01$ ).

Testiranje hipoteza vrši se na sljedeći način:

$$|t| < {}_{df}t_p \Rightarrow H_0: \bar{x}_1 = \bar{x}_2$$

( $H_0$ : Uz pogrešku  $p$  ne možemo tvrditi da je razlika između aritmetičkih sredina prvog i drugog mjerenja statistički značajna)

$$|t| > {}_{df}t_p \Rightarrow H_1: \bar{x}_1 \neq \bar{x}_2$$

( $H_1$ : Razlika između aritmetičkih sredina prvog i drugog mjerenja statistički je značajna uz pogrešku  $p$ )

# T-TEST

**Primjer:** Učenici 1.a i 1.b razreda testirani su na kraju školske godine testom za procjenu fleksibilnosti *Pretklon raskoračno* i dobiveni su sljedeći rezultati:

$$\Rightarrow n_{1a} = 25 \quad n_{1b} = 27$$

$$\Rightarrow \bar{x}_{1a} = 34 \quad \bar{x}_{1b} = 38$$

$$\Rightarrow s_{1a} = 10 \quad s_{1b} = 18$$

Utvrđi da li se aritmetičke sredine 1.a i 1.b razreda statistički značajno razlikuju uz pogrešku  $p < 0,01$ !

# T-TEST

*Primjer - nastavak:*

$$\Rightarrow s_{x_1 - x_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{10^2}{25} + \frac{18^2}{27}} = 4$$

$$\Rightarrow |t| = \left| \frac{\bar{x}_1 - \bar{x}_2}{s_{x_1 - x_2}} \right| = \left| \frac{34 - 38}{4} \right| = |-1| = 1$$

$$\Rightarrow df = n_1 - 1 + n_2 - 1 = 25 - 1 + 27 - 1 = 50 \quad {}_{df}t_p = {}_{50}t_{0,01} = 2,68$$

$$\Rightarrow |t| < {}_{df}t_p \Rightarrow H_0: \bar{x}_1 = \bar{x}_2$$

( $H_0$ : Uz pogrešku 0,01 ne možemo tvrditi da je razlika između aritmetičkih sredina 1.a i 1.b razreda statistički značajna)



# MICROSOFT EXCEL

## *Utvrđivanje kritične t-vrijednosti*

Utvrđivanje kritične t-vrijednosti vrši se pomoću funkcije *Tinv*. Funkcija se unosi u označeno polje matrice odabirom opcije *Function...* padajućeg izbornika *Insert*. U traku *Probability* dijaloškog okvira za unos ove funkcije potrebno je upisati pogrešku  $p$ , a u traku *Deg\_freedom* broj stupnjeva slobode.

**Zadatak** -  $n_1=55$ ,  $\bar{x}_1=220$ ,  $s_1=15$ ,  $n_2=61$ ,  $\bar{x}_2=195$ ,  $s_2=10$

Uz pogrešku  $p=0,05$  testirajte da li se aritmetičke sredine ovih uzoraka statistički značajno razlikuju.

# MICROSOFT EXCEL

## *T-test*

Utvrdjivanje minimalne pogreške s kojom je moguće zaključiti da se dvije aritmetičke sredine statistički značajno razlikuju vrši se pomoću funkcije *Ttest*. Putem traka *Array 1* i *Array 2* potrebno je definirati niz podataka prvog odnosno drugog uzorka/mjerenja, u traku *Tails* unijeti vrijednost 2, a u traku *Type* vrijednost 1 ako se provodi t-test za zavisne uzorke, odnosno 2 ako se provodi t-test za nezavisne uzorke.

**Zadatak** - U datoteci *Pejcic-318.xls* uz pogrešku  $p=0,01$  testirajte da li se aritmetičke sredine učenika 1. i 2. razreda u varijabli *MPOL* statistički značajno razlikuju!

# STATISTICA 7

## *Utvrđivanje kritične t-vrijednosti*

Izračunavanje površina ispod *Studentove t-distribucije* vrši se pomoću dijaloškog okvira *Probability Distribution Calculator* koji se pokreće slijedom koraka: padajući izbornik *Statistics* → *Probability Calculator* → *Distributions* → *t (Student)*. Za izračunavanje *t-vrijednosti* potrebno je označiti opcije *Inverse*, *Two-tailed* i *(1-Cumulative p)*, u traku *p* upisati pogrešku *p*, a u traku *df* broj stupnjeva slobode.

**Zadatak** - Izračunajte koliko iznosi t-vrijednost ako je broj stupnjeva slobode  $df=44$ , a pogreška  $p=0,01$ ?

# STATISTICA 7

## *T-test za nezavisne uzorke*

T-test za nezavisne uzorke na prethodno utvrđenim deskriptivnim parametrima izvodi se slijedom koraka: padajući izbornik *Statistics* → *Basic Statistics/Tables* → *Difference tests: r, %, means* → *Difference between two means (normal distribution)*. U trake *M 1* i *M 2* potrebno je upisati aritmetičku sredinu prvog odnosno drugog uzorka, u trake *StDv 1* i *StDv 2* standardne devijacije, a u trake *N1* i *N2* veličine uzorka.

**Zadatak** -  $n_1=30$ ,  $\bar{x}_1=17$ ,  $s_1=1$ ,  $n_2=34$ ,  $\bar{x}_2=22$ ,  $s_2=2$

Uz pogrešku  $p=0,05$  testirajte da li se aritmetičke sredine ovih uzorka statistički značajno razlikuju.

# STATISTICA 7

## *T-test za nezavisne uzorke*

T-test za nezavisne uzorke na temelju originalnih podataka izvodi se slijedom koraka: padajući izbornik *Statistics* → *Basic Statistics/Tables* → *t-test, independent, by groups*. U dijaloškom okviru koji se pokreće odabirom opcije *Variables* potrebno je označiti selektorsku varijablu (*Grouping variable*) i jednu ili više zavisnih varijabli (*Dependent variables*).

**Zadatak** - U datoteci *Pejcic-318.sta* uz pogrešku  $p=0,01$  testirajte da li se aritmetičke sredine učenika 3. i 4. razreda u varijabli *MPOL* statistički značajno razlikuju!

# STATISTICA 7

## *T-test za zavisne uzorke*

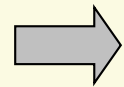
T-test za zavisne uzorke na temelju originalnih podataka izvodi se slijedom koraka: padajući izbornik *Statistics* → *Basic Statistics/Tables* → *t-test, dependent samples*. U dijaloškom okviru koji se pokreće odabirom opcije *Variables* potrebno je označiti varijablu prvog mjerenja (*First variable list*) i varijablu drugog mjerenja (*Second variable list*).

**Zadatak** - U datoteci *POD.sta* uz pogrešku  $p=0,01$  testirajte da li se aritmetičke sredine prvog mjerenja (*MKUS\_I*) i drugog mjerenja (*MKUS\_F*) statistički značajno razlikuju!

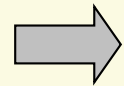
# T-TEST

---

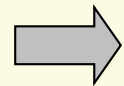
## *Literatura za pripremanje kolokvija*



Dizdar, D. (2006). *Kvantitativne metode*. Zagreb: Kineziološki fakultet, str. 135-149.



Petz, B. (2002). *Osnovne statističke metode za nematematičare*. Jastrebarsko: Naklada Slap, str. 126-151.



Langer, M. (2004). *Brzi vizualni vodič Microsoft Excel 2003 za Windows*. Zagreb: Miš, str. 75-103.