ANOVA TEST

Usporedba više srednjih vrijednosti - ANOVA

Za testiranje hipoteze o jednakosti srednjih vrijednosti dvije populacije:

$$H_0 \quad \mu_X = \mu_Y$$

koristili smo t-test.

Kako usporediti više srednjih vrijednosti? Npr.

$$H_0 \quad \mu_1 = \mu_2 = \mu_3 \quad ?$$

Usporedba više srednjih vrijednosti se javlja u situaciji kada uspoređujemo više tretmana (metoda) i želimo ustanoviti jesu li one različite.

Možemo li hipotezu H_0 testirati tako da testiramo hipoteze

$$\begin{aligned} H_{0,1} &: \quad \mu_1 = \mu_2 \\ H_{0,2} &: \quad \mu_1 = \mu_3 \\ H_{0,3} &: \quad \mu_2 = \mu_3 \end{aligned}$$

Višestruko testiranje

Pretpostavimo da hipotezu

$$H_0 : \quad \mu_1 = \mu_2$$

nezavisno testiramo 5 puta.

(5 puta biramo uzorak i testiramo hipotezu.)

Nakon 5 testiranja hipotezu odbacujemo ukoliko smo je odbacili u barem jednom pojedinačnom testiranju.

Razina značajnosti u pojedinom testiranju je $\bar{\alpha}$.

Kolika je vjerojatnost da ćemo nakon 5 testiranja odbaciti hipotezu ukoliko je ona istinita?

Kolika je pogreška I. vrste za test koji se sastoji od 5 pojedinačnih testiranja?

Primjer. Na uzorku veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$.

Uzorak je generiran pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 42 puta (**4.2%**).

Ovo je u skladu s razinom značajnosti od **5%**.

Primjer. Na **5** uzoraka veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$. Hipotezu odbacujemo ukoliko barem u jednom od 5 testiranja odbacimo hipotezu.

Uzorci su generirani pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 222 puta (**22.2%**).

Ovo nije u skladu s razinom značajnosti od **5%**.

Označimo:

A = Prihvaćanje nul hipoteze

A_1 = Prihvaćanje nul hipoteze u 1. testiranju

A_2 = Prihvaćanje nul hipoteze u 2. testiranju

A_3 = Prihvaćanje nul hipoteze u 3. testiranju

A_4 = Prihvaćanje nul hipoteze u 4. testiranju

A_5 = Prihvaćanje nul hipoteze u 5. testiranju

Sada je

$$A = A_1 \text{ i } A_2 \text{ i } A_3 \text{ i } A_4 \text{ i } A_5$$

Uvjetna vjerojatnost uz uvjet da je H_0 istinita za ove događaje je

$$P(A \mid H_0) = P(A_1 \textbf{ i } A_2 \textbf{ i } A_3 \textbf{ i } A_4 \textbf{ i } A_5 \mid H_0) =$$

nezavisni testovi

$$= P(A_1 \mid H_0) \cdot P(A_2 \mid H_0) \cdot P(A_3 \mid H_0) \cdot P(A_4 \mid H_0) \cdot P(A_5 \mid H_0)$$

Oznake:

$\bar{\alpha}$ - Razina značajnosti u pojedinom testiranju.

α - Razina značajnosti ukupnog testa.

Dakle:

$$1 - \alpha = (1 - \bar{\alpha})^5$$

odnosno

$$\bar{\alpha} = 1 - \sqrt[5]{1 - \alpha}$$

Općenito, kod ponavljanja *m* testova,

$$1 - \alpha = (1 - \bar{\alpha})^m$$

i

$$\bar{\alpha} = 1 - \sqrt[m]{1 - \alpha}.$$

Jer je

$$\sqrt[m]{1 - \alpha} \approx 1 - \frac{\alpha}{m}$$

slijedi da je

$$\bar{\alpha} \approx \frac{\alpha}{m}.$$

Primjer. Na **5** uzoraka veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$. Hipotezu odbacujemo ukoliko barem u jednom od 5 testiranja odbacimo hipotezu.

Uzorci su generirani pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 222 puta (**22.2%**).

Ovo nije u skladu s razinom značajnosti od **5%**.

$$\alpha = 1 - (1 - \bar{\alpha})^m = 1 - (1 - 0.05)^5 = 0.226219 = \mathbf{22.6\%}$$

Ako su testovi potpuno zavisni, rezultat je drugačiji.

Npr., srednje vrijednosti uspoređujemo više varijabli na dva uzorka i varijable su zavisne. Uspoređujemo visinu danu u:

- metrima (X)
- centimetrima (Y)
- inčima (Z)

$$
\begin{aligned}
H_{0,1} &: \quad X_1 = X_2 \\
H_{0,2} &: \quad Y_1 = Y_2 \\
H_{0,3} &: \quad Z_1 = Z_2
\end{aligned}
$$

Označimo:

A = Prihvaćanje nul hipoteze

A_1 = Prihvaćanje nul hipoteze u 1. testiranju ($H_{0,1}$)

A_2 = Prihvaćanje nul hipoteze u 2. testiranju ($H_{0,2}$)

A_3 = Prihvaćanje nul hipoteze u 3. testiranju ($H_{0,3}$)

Sada je

$$A = A_1 \textbf{ i } A_2 \textbf{ i } A_3 = A_1 = A_2 = A_3$$

i

$$P(A \mid H_0) = P(A_1 \textbf{ i } A_2 \textbf{ i } A_3 \mid H_0) = P(A_1 \mid H_0) = P(A_2 \mid H_0) = P(A_3 \mid H_0)$$

odnosno

$$1 - \alpha = 1 - \bar{\alpha} \quad \Leftrightarrow \quad \alpha = \bar{\alpha}$$

Korekcija treba biti između

- $1/m$ za nezavisne testove i
- 1 za potpuno zavisne.

Testiramo m hipoteza: $H_1,\ H_2,\ H_3, \dots, H_m$.

Generalna hipoteza:

H_0: Sve hipoteze $H_1,\ H_2,\ H_3, \dots, H_m$ su istinite.

H_a: Barem jedna od hipoteza $H_1,\ H_2,\ H_3, \dots, H_m$ nije istinita.

Oznake:

$\bar{\alpha}$ - Razina značajnosti u pojedinom testiranju hipoteza $H_1,\ H_2, \dots, H_m$.

α - Razina značajnosti ukupnog testa, testiranja hipoteze H_0.

Kako odrediti $\bar{\alpha}$ tako da vjerojatnost pogreške I. vrste bude manja od α?

'Family-wise error rate (FWER)' je vjerojatnost pojave jednog ili više lažnih otkrića, odnosno, pogrešaka I. vrste među svim hipotezama pri testiranju više hipoteza,

Bonferronijeva korekcija $\bar{\alpha} = \frac{\alpha}{m}.$

Šidákova korekcija $\bar{\alpha} = 1 - \sqrt[m]{1 - \alpha}.$

Holmova korekcija

- $p_{(1)}, \ldots, p_{(m)}$ sortirane p-vrijednosti individualnih testova $(p_{(1)} \leq \ldots \leq p_{(m)})$.
  Neka su $H_1, \ldots, H_m$ odgovarajuće nul hipoteze.
- Holmova procedura je definirana korak po korak (engl. 'stepwise').
- 1. korak. Ako je $p_{(1)} \geq \alpha/m$, prihvati $H_1, \ldots, H_m$ i stani.
  Ako je $p_{(1)} < \alpha/m$, odbaci H_1 i testiraj ostalih $m - 1$ hipoteza na razini $\alpha/(m - 1)$.
- 2. korak. Ako je $p_{(1)} < \alpha/m$, a $p_{(2)} \geq \alpha/(m - 1)$, prihvati $H_2, \ldots, H_m$ i stani. Ako je $p_{(1)} < \alpha/m$ i $p_{(2)} < \alpha/(m - 1)$, odbaci H_2 uz H_1 i testiraj ostalih $m - 2$ hipoteza na razini $\alpha/(m - 2)$.
- ...

- Bonferronijeva (i Šidákova) korekcija je konzervativna.
- Premala pogreška I. vrste.
- Povećani broj prihvaćanja nul hipoteze.
- Veća pogreška II. vrste.
- Mala snaga testa.

Druge metode.

Tukeyeva procedura, Duncanova procedura. - Usporedba svih parova ($m(m-1)/2$ usporedbi).

Dunnettova procedura. - Usporedba kontrolne skupine s preostalim skupinama ($m-1$ usporedbi).

Umjesto FWER-a možemo kontrolirati:

FDR ('False Discovery Rate') - kontrolira se udio krivo odbačenih ('false positive') hipoteza unutar skupa odbačenih hipoteza.

- Benjamini-Hochberg procedura

- Benjamini-Hochberg-Yekutieli procedura

Primjer. Na **5** uzoraka veličine 100 ($n = 100$) iz normalne populacije sa srednjom vrijednošću 0 i standardnom devijacijom 1 testirana je hipoteza $\mu = 0$ uz razinu značajnosti $\alpha = 0.05$. Hipotezu odbacujemo ukoliko barem u jednom od 5 testiranja odbacimo hipotezu uz razinu značajnosti $\bar{\alpha} = 0.05/5 = 0.01$.

Uzorci su generirani pomoću generatora slučajnih brojeva.

Eksperiment je ponovljen 1,000 puta.

Rezultat. Hipoteza je odbačena 52 puta (**5.2%**).

Ovo je u skladu s razinom značajnosti od **5%**.

ANOVA test

Za k **normalnih** populacija želimo provjerilti jesu li srednje vrijednosti obilježja u tim populacijama jednake.

Testiramo hipotezu

$$H_0 \quad \mu_1 = \mu_2 = \ldots = \mu_k.$$

Alternativna hipoteza:

$$H_a \quad \text{barem jedan par srednjih vrijednosti je različit}$$

Želimo izbjeći pristup pomoću višestrukog uspoređivanja srednjih vrijednosti ($k(k-1)/2$ usporedbi).

Najčešće se radi o usporedbi više tretmana.

Svaka populacija odgovara pojedinom tretmanu.

Biramo

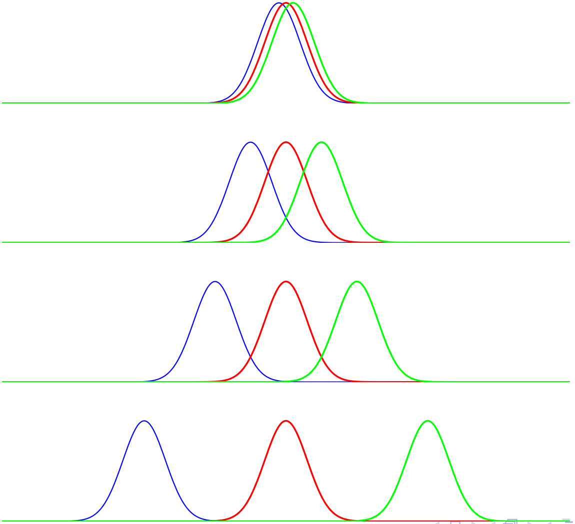
- k uzoraka iz k populacija ili
- k uzoraka iz jedne populacije i na svaki uzorak primijenimo drugi tretman.
- veličina uzoraka ne treba biti ista. Mi ćemo radi jednostavnosti pretpostaviti da su svi uzorci iste veličine (n, ukupno $k \cdot n$ podataka)

Važne pretpostavke na populacije:

- normalna distribucija obilježja u svakoj populaciji
- ista standardna devijacija (varijanca) obilježja u svih k populacija.

Distribucija za 3 (jednako velike) populacije.

Gdje je najveća ukupna varijanca?



ANOVA - Analysis of variance

U prvom koraku izračunamo srednje vrijednosti za svaki uzorak.

Tretman	Slučajni uzorak	Sr. vr. uzorka
Tretman 1	$X_{11}\ X_{12}\ X_{13}\ \ldots\ X_{1j}\ \ldots\ X_{1n}$	$\bar{X}_1$
Tretman 2	$X_{21}\ X_{22}\ X_{23}\ \ldots\ X_{2j}\ \ldots\ X_{2n}$	$\bar{X}_2$
$\vdots$	$\vdots\ \vdots\ \vdots\ \vdots\ \vdots$	$\vdots$
Tretman i	$X_{i1}\ X_{i2}\ X_{i3}\ \ldots\ X_{ij}\ \ldots\ X_{in}$	$\bar{X}_i$
$\vdots$	$\vdots\ \vdots\ \vdots\ \vdots\ \vdots$	$\vdots$
Tretman k	$X_{k1}\ X_{k2}\ X_{k3}\ \ldots\ X_{kj}\ \ldots\ X_{kn}$	$\bar{X}_k$

$$\bar{X}_i = \frac{1}{n}\sum_j X_{ij}$$ - srednja vrijednost uzorka

Ujedno izračunamo i srednju vrijednost cijelog uzorka ($k \cdot n$ podataka)

$$\bar{X} = \frac{1}{k \cdot n} \sum_i \sum_j X_{ij} = \frac{1}{k} \sum_i \bar{X}_i.$$

Suma kvadrata (sum of squares)

Osnova ANOVA testa je usporedba različitih procjena varijance σ^2.

Na osnovu svih uzoraka ($k \cdot n$ podataka) računamo ukupnu varijancu:

$$S^2 = \frac{1}{k \cdot n - 1} \sum_i \sum_j \left(X_{ij} - \bar{X}\right)^2.$$

Posebno ćemo označiti **ukupnu sumu kvadrata**

$$SS(T) = \sum_i \sum_j \left(X_{ij} - \bar{X}\right)^2.$$

$SS(T)$ - total sum of squares

Za svaki tretman (uzorak) izračunamo procjenu varijance

$$S_i^2 = \frac{1}{n-1} \sum_j \left(X_{ij} - \bar{X}_i\right)^2, \quad i = 1, \ldots, k.$$

Na osnovu ovih k procjena konstruiramo jednu zajedničku procjenu:

$$S_P^2 = \frac{1}{k} \sum_i S_i^2 = \frac{1}{k \cdot (n-1)} \sum_i \sum_j \left(X_{ij} - \bar{X}_i\right)^2.$$

Odgovarajuću sumu kvadrata nazivamo **suma kvadrata za pogreške**

$$SS(E) = \sum_i \sum_j \left(X_{ij} - \bar{X}_i\right)^2.$$

$SS(E)$ - sum of squares for errors

Na kraju računamo **sumu kvadrata za tretmane (faktore)**

$$SS(F) = \sum_i n\left(\bar{X}_i - \bar{X}\right)^2.$$

Može se pokazati da je

$$SS(T) = SS(E) + SS(F).$$

Ukoliko su uzorci iz normalne populacije, tada je

$$(k \cdot n - 1)\frac{S^2}{\sigma^2} \sim \chi^2(k \cdot n - 1)$$

tj.

$$\frac{SS(T)}{\sigma^2} \sim \chi^2(k \cdot n - 1).$$

Nadalje,

$$(n-1)\frac{S_i^2}{\sigma^2} \sim \chi^2(n-1), \quad i = 1, \ldots, k.$$

Kako su S_i^2 nezavisni (zbog nezavisnosti uzoraka), tada je

$$\begin{aligned}
(n-1)\frac{S_1^2}{\sigma^2} + (n-1)\frac{S_2^2}{\sigma^2} + \ldots + (n-1)\frac{S_k^2}{\sigma^2} &\sim \chi^2(k \cdot (n-1)) \\
k \cdot (n-1)\frac{S_P^2}{\sigma^2} &\sim \chi^2(k \cdot (n-1)) \\
\frac{SS(E)}{\sigma^2} &\sim \chi^2(k \cdot (n-1))
\end{aligned}$$

Što je sa $SS(F)$?

Može se pokazati da je

- $\dfrac{SS(F)}{\sigma^2} \sim \chi^2(k-1)$
- $SS(E)$ i $SS(F)$ su nezavisne.

Uz

$$\frac{SS(E)}{\sigma^2} \sim \chi^2(k\cdot(n-1))$$

dobijamo statistiku

$$F = \frac{\frac{SS(F)}{k-1}}{\frac{SS(E)}{k\cdot(n-1)}} = \frac{MS(F)}{MS(T)} \sim F(k-1, k\cdot(n-1))$$

Kritično područje za razinu značajnosti α je

$$F \geq F_{1-\alpha}(k-1, k\cdot(n-1)).$$

Rezultati se obično prikazuju pomoću tablice

Izvor varijacije	**Suma kvadrata**	**Br. st. slobode**	**Srednja vr. sume kvadrata**	**Omjer F**
Tretman (faktor)	$SS(F)$	$k-1$	$MS(F) = \dfrac{SS(F)}{k-1}$	$\dfrac{MS(F)}{MS(E)}$
Pogreška	$SS(E)$	$k \cdot (n-1)$	$MS(E) = \dfrac{SS(E)}{k \cdot (n-1)}$	
Ukupno	$SS(T)$	$k \cdot n - 1$		

Primjer. Trener želi usporediti tri različite metode treninga. Svaku od metoda primijenio je na po $n = 4$ studenta. Nakon 30 dana ocijenjena je uspješnost i ocjene su prikazane u tablici

Metoda	**Observacije**			
Metoda 1	3	6	4	7
Metoda 2	11	8	10	7
Metoda 3	6	9	5	8

Jesu li sve tri metode jednako uspješne? Hipotezu testirajte uz razinu značajnosti $\alpha = 0.05$.

Rješenje. 3 uzorka ($k = 3$) s po 4 ispitanika u svakom uzorku ($n = 4$).

Prvo računamo srednje vrijednosti uzoraka:

$$\begin{aligned}
\bar{X}_1 &= \frac{3+6+4+7}{4} = \frac{20}{4} = 5 \\
\bar{X}_2 &= \frac{11+8+10+7}{4} = \frac{36}{4} = 9 \\
\bar{X}_3 &= \frac{6+9+5+8}{4} = \frac{28}{4} = 7 \\
\bar{X} &= \frac{20+36+28}{12} \left(= \frac{5+9+7}{3} \right) = \frac{84}{12} = 7
\end{aligned}$$

Metoda	Observacije				Sr. vr.
Metoda 1	3	6	4	7	5
Metoda 2	11	8	10	7	9
Metoda 3	6	9	5	8	7
Ukupno					7

Zatim računamo sume kvadrata.

Suma kvadrata za faktore:

$$\begin{aligned} SS(F) &= \sum_i 4\left(\bar{X}_i - \bar{X}\right)^2 = \\ &= 4\cdot(5-7)^2 + 4\cdot(9-7)^2 + 4\cdot(7-7)^2 = \\ &= 16+16+0 = \mathbf{32} \end{aligned}$$

Suma kvadrata za pogreške:

$$\begin{aligned} SS(E) &= \sum_i \sum_j \left(X_{ij} - \bar{X}_i\right)^2 = \\ &= (3-5)^2 + (6-5)^2 + (4-5)^2 + (7-5)^2 + \\ &\quad (11-9)^2 + (8-9)^2 + (10-9)^2 + (7-9)^2 + \\ &\quad (6-7)^2 + (9-7)^2 + (5-7)^2 + (8-7)^2 = \\ &= 4+1+1+4+4+1+1+4+1+4+4+1 = \\ &= \mathbf{30} \end{aligned}$$

Za ilustraciju računamo i ukupnu sumu kvadrata:

$$
\begin{aligned}
SS(T) &= \sum_i \sum_j \left(X_{ij} - \bar{X}\right)^2 = \\
&= (3-7)^2 + (6-7)^2 + (4-7)^2 + (7-7)^2 + \\
&\quad (11-7)^2 + (8-7)^2 + (10-7)^2 + (7-7)^2 + \\
&\quad (6-7)^2 + (9-7)^2 + (5-7)^2 + (8-7)^2 = \\
&= 16 + 1 + 9 + 0 + 16 + 1 + 9 + 0 + 1 + 4 + 4 + 1 = \\
&= \mathbf{62}
\end{aligned}
$$

Provjerimo

$$
\begin{aligned}
SS(T) &= SS(F) + SS(E) \\
62 &= 32 + 30
\end{aligned}
$$

Računamo srednje sume kvadrata:

$$\begin{aligned} MS(F) &= \frac{SS(F)}{k-1} = \frac{32}{2} = 16 \\ MS(E) &= \frac{SS(E)}{k\cdot(n-1)} = \frac{30}{3\cdot(4-1)} = \frac{30}{9} = 3.333 \end{aligned}$$

i statistiku

$$F = \frac{MS(F)}{MS(E)} = \frac{16}{3.33} = 4.80$$

Sve ove podatke pregledno prikazujemo u ANOVA tablici:

Izvor varijacije	**Suma kvadrata**	**Br. st. slobode**	**Srednja vr. sume kvadrata**	**Omjer F**
Tretman	32	2	16	4.80
Pogreška	30	9	3.33	
Ukupno	62	11		

Pomoću kalkulatora izračunamo $F_{0.95}(2,9) = 4.26$.

Jer je

$$F = 4.80 > 4.26 = F_{0.95}(2,9)$$

hipotezu o jednakosti srednjih vrijednosti odbacujemo uz razinu značajnosti $\alpha = 0.95$.

Veličina efekta

- značajan F omjer znači da postoji razlika između srednjih vrijednosti.
- razlika može biti zanemarivog iznosa ali statistički značajna zbog dovoljno velikog uzorka ili male varijacije unutar populacija.
- ponekad razlika može biti velika ali ne i statistički značajna (mali uzorak ili velika standardna devijacija).
- veličina efekta - metoda određivanja efekta tretmana bez utjecaja veličine uzorka

R^2 (η^2, eta kvadrat)

$$R^2 = \frac{SS(F)}{SS(T)}$$

$$0 \leq R^2 \leq 1$$

R^2 - dio varijance koji je opisan efektom tretmana.
Podaci iz primjera:

$$R^2 = \frac{SS(F)}{SS(T)} = \frac{32}{62} = 0.516$$

R^2 je pristrani procjenitelj

ω^2

$$\omega^2 = \frac{SS(F) - (k-1)MS(E)}{SS(T) + MS(E)}$$

U našem primjeru:

$$\omega^2 = \frac{SS(F) - (k-1)MS(E)}{SS(T) + MS(E)3} = \frac{32 - 2 \cdot 3.33}{62 + 3.33} = 0.388$$

$$\omega^2 \leq R^2$$

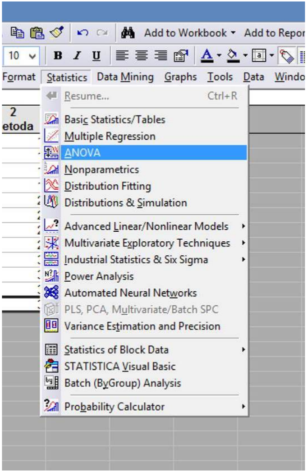
Također mjeri udio objašnjne varijance u ukupnoj varijanci.

ω^2 je nepristrani procjenitelj (bolji izbor nego R^2)

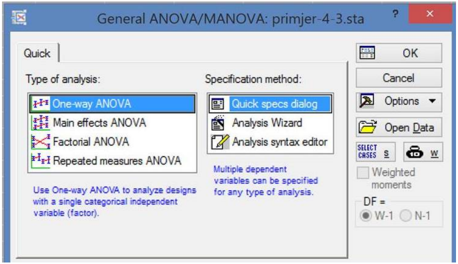
Rješenje primjera pomoću Statistice.

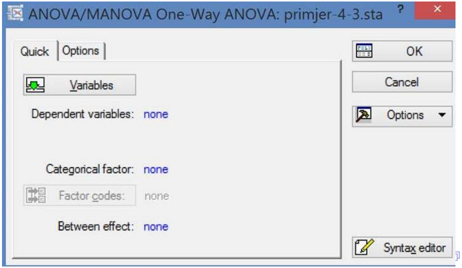
Podaci:

	1 Ocjena	2 Metoda
1	3	1
2	6	1
3	4	1
4	7	1
5	11	2
6	8	2
7	10	2
8	7	2
9	6	3
10	9	3
11	5	3
12	8	3

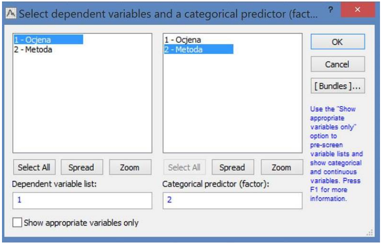


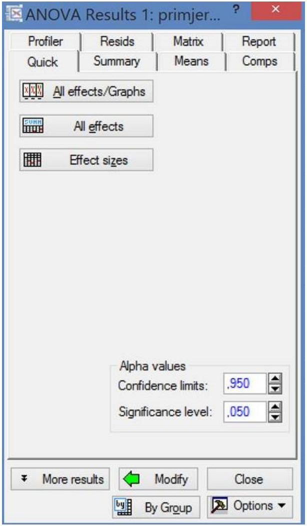
Biramo ANOVA test za usporedbu *k* srednjih vrijednosti:





Izaberemo varijable i pokrenemo analizu:





Izaberemo 'All effects'.

ANOVA tablica:

Effect	Univariate Tests of Significance for Ocjena (primjer-4-3.sta) Sigma-restricted parameterization Effective hypothesis decomposition SS	Degr. of Freedom	MS	F	p
Intercept	588,0000	1	588,0000	176,4000	0,000000
Metoda	32,0000	2	16,0000	4,8000	0,038131
Error	30,0000	9	3,3333		

Napomena. Prvi redak je dodatak.

Zbroj suma kvadrata je 640, zbroj kvadrata svih podataka.

ANOVA test nam je pokazao da su srednje vrijednosti različite, tj. da se metode razlikuju.

Koje od tih tri metoda se razlikuju?

Za odgovor na ovo pitanje treba napraviti međusobnu usporedbu svih metoda.

$\rightarrow$ višestruka usporedba.

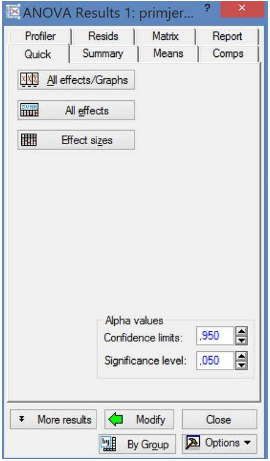
Nije poželjno koristiti *t*-test za nezavisne uzorke s Bonferronijevom korekcijom.

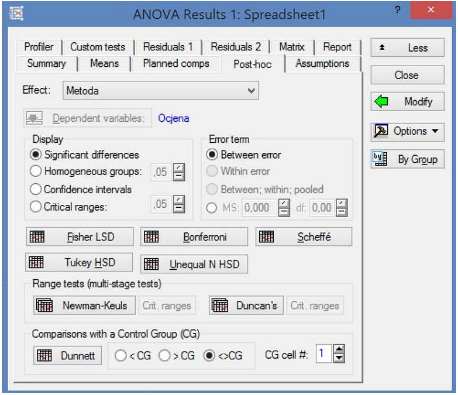
Ukoliko ANOVA test pokaže razliku između tretmana, standardno se koriste sljedeći testovi za višestruku usporedbu ('post hoc' testovi):

- **Duncanov test** - kontrolira pogrešku II. vrste na štetu pogreške I. vrste.
- **Tukeyev test** - u principu se radi od t-testu uz kontrolu FWER-a.
- **Dunnettov test** - usporedba samo s kontrolnom grupom ($k-1$ usporedbi) za razliku od Tukeyovog testa gdje se sve grupe međusobno uspoređuju ($k \cdot (k-1)$ usporedbi).

Za 'post hoc' testiranje u Statistici izaberemo opciju 'More results':

Za ’post hoc’ testiranje u Statistici izaberemo opciju ’More results’ te izbornik ’Post hoc’





Izaberemo ’Tukey HSD’

Tukey HSD test; variable Ocjena (Spreadsheet1)
Approximate Probabilities for Post Hoc Tests
Error: Between MS = 3,3333, df = 9,0000

Cell No.	Metoda	{1} 5,0000	{2} 9,0000	{3} 7,0000
1	1		0,031141	0,315092
2	2	0,031141		0,315092
3	3	0,315092	0,315092	

Razlikuju se Metoda 1 i Metoda 2.

Metoda	**Sr. vr.**
Metoda 1	5
Metoda 2	9
Metoda 3	7

ANOVA i 'post hoc' testovi pretpostavlja da

- uzorci su nezavisni
  to uvijek pretpostavljamo ukoliko nije drugačije napomenuto
- uzorci su iz normalne populacije
  Ova pretpostavka vrijedi za skoro sve do sada spomenute testove
- Standardna devijacija populacija je jednaka.
  Ova pretpostavka se u slučaju dvije populacije testira F-testom.
  Što u slučaju više populacija?

Levene-ov test - usporedba više varijanci

Za testiranje hipoteze o jednakosti varijanci k populacija:

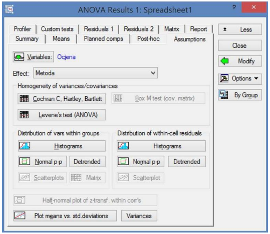
$$H_0 : \quad \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$$

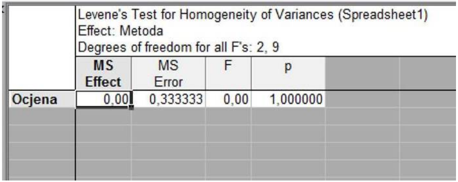
koristi se **Levene-ov test**.

Homogenost varijanci = jednake varijance = homoscedastičnost

heteroscedastičnost - različite varijance

U izborniku 'More results' izaberemo izbornik 'Assumptions':





Levene's Test for Homogeneity of Variances (Spreadsheet1)
Effect: Metoda
Degrees of freedom for all F's: 2, 9

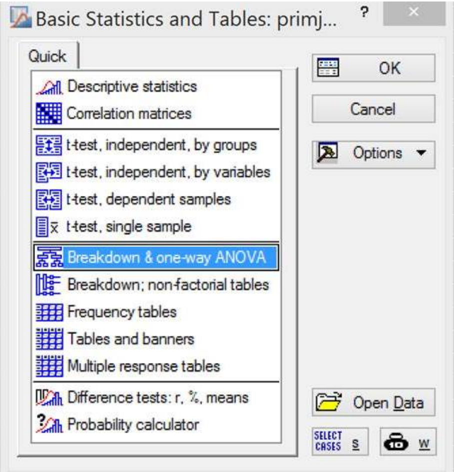
	MS Effect	MS Error	F	p
Ocjena	0,00	0,333333	0,00	1,000000

Ukoliko se varijance razlikuju, ne možemo koristiti ANOVA test.

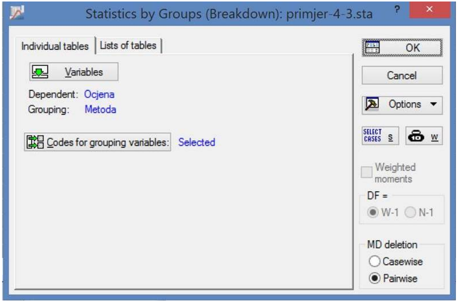
Rješenje: **Welchov test** - sve je isto kao u ANOVA testu jedino se koristi drugačija statistika

Za Welchov test u Statistici treba koristiti drugu (osnovnu) verziju ANOVA testa.

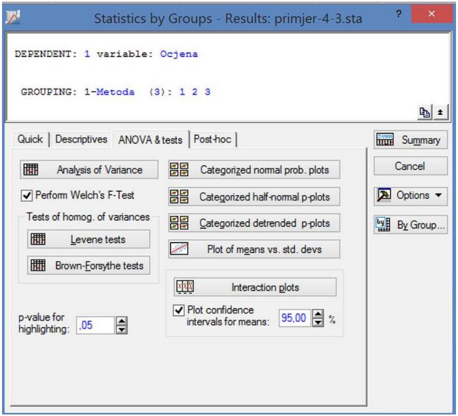
U basic Statistics and Tables izaberemo Breakdown & one-way ANOVA



Definiramo varijable



Izbornik ANOVA & tests



Leveneov i Welchow test

Levene's Test for Homogeneity of Variances (Spreadsheet1)
Effect: Metoda
Degrees of freedom for all F's: 2, 9

	MS Effect	MS Error	F	p
Ocjena	0,00	0,333333	0,00	1,000000

Workbook6* - Analysis of Variance (primjer-4-3.sta)

Analysis of Variance (primjer-4-3.sta)
Marked effects are significant at p < ,05000

Variable	SS Effect	df Effect	MS Effect	SS Error	df Error	MS Error	F	p	Welch df Effect	Welch df Error	Welch F	Welch p
Ocjena	32,00000	2	16,00000	30,00000	9	3,333333	4,800000	0,038131	2	6,000000	4,320000	0,068838

Napomena. Ako se uspoređuju dvije populacije, ANOVA i t-test daju isti rezultat.

Za dvije populacije (X_1 i X_2) je

$$F = \frac{\left(\bar{X}_1 - \bar{X}_2\right)^2}{\left(S_1^2 + S_2^2\right)/n}.$$

S druge strane, u t-testu za dvije populacije s jednakim varijancama, statistika je dana s

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(S_1^2 + S_2^2\right)/n}}.$$

Vrijedi:

$$F = t^2.$$

ANOVA - višestruka klasifikacija

Primjer. Istraživač želi provjeriti da li nova vježba utječe na rezultat treninga (npr. povećanje eksplozivne snage).

Kako potvrditi pretpostavku?

Slučajno izabere dvije grupe ispitanika.

- Jedna grupa trenira mjesec dana po starom programu.

- Jedna grupa trenira mjesec dana po novom programu.

t-test - usporedba rezultata kontrolne i eksperimentalne grupe.

Istraživač želi još i provjeriti kvalitetu nove opreme za vježbanje.

Kontrolna i eksperimentalna grupa + t-test?

Može iskoristiti kontrolnu grupu iz prethodnog istraživanja.

Testiranje: t-test?

Dva testa $\rightarrow$ bolje jedan test $\rightarrow$ ANOVA

Usporedba 3 tretmana. - Jedna grupa trenira mjesec dana po starom programu i na staroj opremi.

- Jedna grupa trenira mjesec dana po novom programu.

- Jedna grupa trenira mjesec dana na novoj opremi.

Napomena. Usporedba samo s kontrolnom grupom.
$\rightarrow$ Dunnett-ov 'post hoc' test.

Može li izabrati bolji dizajn eksperimenta?

- Jedna grupa trenira mjesec dana po starom programu.

- Jedna grupa trenira mjesec dana po novom programu i na novoj opremi.

Ukoliko postoji razlika između dviju grupa ne može se utvrditi je li razlika posljedica novog programa ili nove opreme.

Koristimo 4 grupe:

- Jedna grupa trenira mjesec dana po starom programu i na staroj opremi.
- Jedna grupa trenira mjesec dana po novom programu.
- Jedna grupa trenira mjesec dana na novoj opremi.
- Jedna grupa trenira mjesec dana po novom programu i na novoj opremi.

Varijable:

Program - novi, stari

Oprema - nova, stara

Za svaki uzorak izračunamo srednju vrijednost:

		Oprema	
		Stara	**Nova**
Program	**Stari**	$\bar{X}_{11}$	$\bar{X}_{12}$
	Novi	$\bar{X}_{21}$	$\bar{X}_{22}$

Zanima nas razlika

- između starog i novog programa
- između stare i nove opreme

Ne uspoređujemo srednje vrijednosti ovih grupa.

Za svaki program i svaku opremu izračunamo srednje vrijednosti:

		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	$\bar{X}_{11}$	$\bar{X}_{12}$	$\bar{X}_{p1}$
	Novi	$\bar{X}_{21}$	$\bar{X}_{22}$	$\bar{X}_{p2}$
Marginalna sr. vr.		$\bar{X}_{o1}$	$\bar{X}_{o2}$	$\bar{X}$

Sume kvadrata računaju se pomoću marginalnih srednjih vrijednosti.

Testiraju se hipoteze:

$$\begin{aligned} H_{p0} &: \quad \mu_{p1} = \mu_{p2} \\ H_{o0} &: \quad \mu_{o1} = \mu_{o2} \end{aligned}$$

- Grupe smo definirali prema vrijednostima varijabli program i oprema.
- Ove varijable se nazivaju **klasifikacijske varijable**.
- Koristili smo dvije klasifikacijske varijable $\rightarrow$ **dvostruka klasifikacija**.
- 'Obična' ANOVA - samo jedna klasifikacijska varijabla $\rightarrow$ **ANOVA s jednostrukom klasifikacijom**
- Još se koristi i naziv 'one-way' ANOVA
- Može se koristiti i više klasifikacijskih varijabli $\rightarrow$ **višestruka klasifikacija**.

- Klasifikacijske varijable se nazivaju i
  - **faktori** $\rightarrow$ **faktorska ANOVA**.
  - **efekti**
  - nezavisne varijable
  - prediktori
  - grupirajuće varijable
- Neprekidna varijabla (koju analiziramo) često se naziva
  - **kriterijska varijabla**
  - zavisna varijabla
  - varijabla odziva ('response variable')

- Vrijednosti klasifikacijskih varijabli se nazivaju **nivoi** ('level')
- Klasifikacijske varijable program i oprema imaju po dva nivoa.
- U našem primjeru govorimo o 2×2 (faktorskom) dizajnu eksperimenta.
- Npr., ukoliko klasifikacijske varijable imaju po 4, 2 i 3 nivoa, tada se radi o $4 \times 2 \times 3$ dizajnu eksperimenta.
- **Na istom broju ispitanika faktorska ANOVA ima veću snagu od ostalih testova.**
- Uz to, možemo testirati i međudjelovanje faktora (**interakcija**)

Interakcija

Interakciju ćemo ilustrirati na izmišljenim podacima za naš primjer.

		Oprema		Marginalna
		Stara	Nova	sr. vr.
Program	Stari	60	70	65
	Novi	80	90	85
Marginalna sr. vr.		70	80	75

		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	60	70	65
	Novi	80	90	85
Marginalna sr. vr.		70	80	75

Efekt opreme: $80 > 70$

Nova oprema povećava eksplozivnu snagu za 10.

Efekt programa: $85 > 65$

Novi program povećava eksplozivnu snagu za 20.

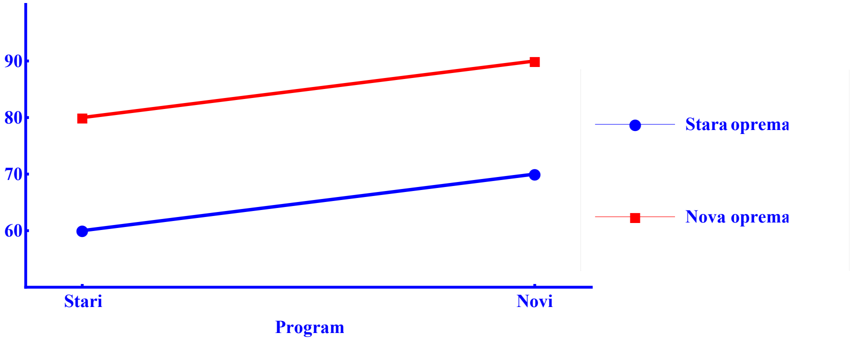
		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	60	70	65
	Novi	80	90	85
Marginalna sr. vr.		70	80	75

Efekt interakcije:

Program=stari
Nova oprema povećava eksplozivnu snagu za 70-60= 10.

Program=novi
Nova oprema povećava eksplozivnu snagu za 90-80= 10.

Za svaki nivo programa povećanje je isto $\rightarrow$ **Nema interakcije**



Pravci su paralelni $\rightarrow$ **Nema interakcije**

Primjer 2. Drugi oblik interakcije ćemo ilustrirati s drugim skupom simuliranih podataka.

		Oprema		Marginalna
		Stara	Nova	sr. vr.
Program	Stari	60	70	65
	Novi	80	100	90
Marginalna sr. vr.		70	85	77.5

		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	60	70	65
	Novi	80	100	90
Marginalna sr. vr.		70	85	77.5

Efekt opreme: $85 > 70$

Nova oprema povećava eksplozivnu snagu za 15.

Efekt programa: $90 > 65$

Novi program povećava eksplozivnu snagu za 25.

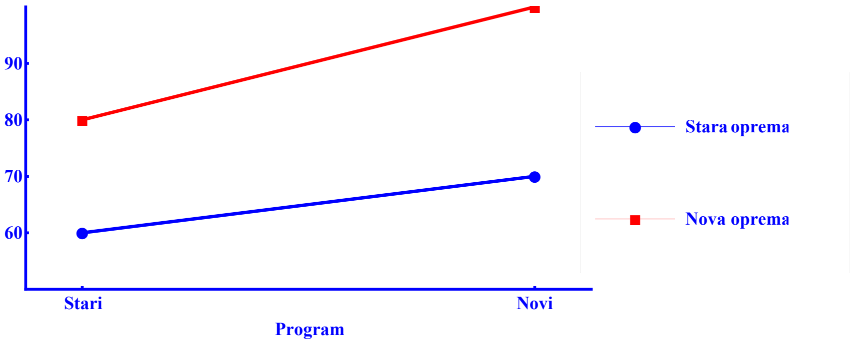
		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	60	70	65
	Novi	80	100	90
Marginalna sr. vr.		70	85	77.5

Efekt interakcije:

Program=stari
Nova oprema povećava eksplozivnu snagu za 70-60= 10.

Program=novi
Nova oprema povećava eksplozivnu snagu za 100-80= 20.

Za različite nivoe povećanje nije isto → **Postoji interakcija**



Pravci nisu paralelni $\rightarrow$ **Postoji interakcija**

Primjer 3. Još jedan oblik interakcije ćemo ilustrirati s još jednim skupom simuliranih podataka.

		Oprema		Marginalna
		Stara	Nova	sr. vr.
Program	Stari	60	80	70
	Novi	90	70	80
Marginalna sr. vr.		75	75	75

		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	60	80	70
	Novi	90	70	80
Marginalna sr. vr.		75	75	75

Efekt opreme: $75 = 75$

Nova oprema ne povećava eksplozivnu snagu.

Efekt programa: $80 > 70$

Novi program povećava eksplozivnu snagu za 10.

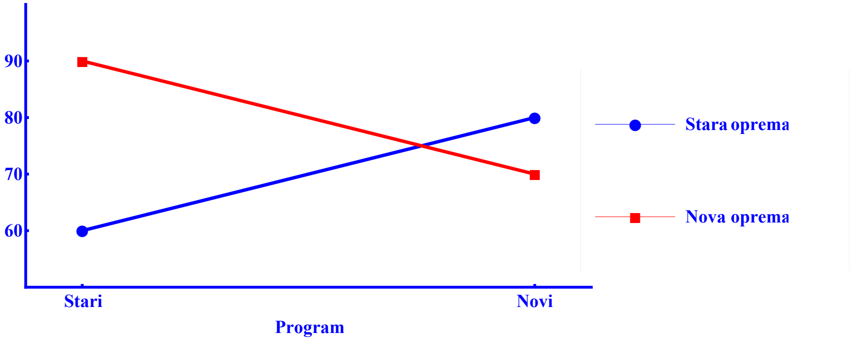
		Oprema		Marginalna sr. vr.
		Stara	Nova	
Program	Stari	60	80	70
	Novi	90	70	80
Marginalna sr. vr.		75	75	75

Efekt interakcije:

Program=stari
Nova oprema povećava eksplozivnu snagu za 80-60= 20.

Program=novi
Nova oprema smanjuje eksplozivnu snagu za 90-70= 20.

Različit efekt opreme za različite nivoe programa $\rightarrow$ **Postoji interakcija**



Pravci nisu paralelni $\rightarrow$ **Postoji interakcija**

Sume kvadrata

$$
\begin{aligned}
SS(F_1) &= n \cdot k_1 \sum_i \left(\bar{X}_{F_1,i} - \bar{X}\right)^2 \\
SS(F_2) &= n \cdot k_2 \sum_i \left(\bar{X}_{F_2,i} - \bar{X}\right)^2 \\
SS(I_{12}) &= n \sum_i \sum_j \left(\bar{X}_{ij} - \bar{X}_{F_1,i} - \bar{X}_{F_2,i} + \bar{X}\right)^2 \\
SS(E) &= \sum_i \sum_j \sum_k \left(X_{ijk} - \bar{X}_{ij}\right)^2 \\
SS(T) &= \sum_i \sum_j \sum_k \left(X_{ijk} - \bar{X}\right)^2
\end{aligned}
$$

		F_1		Marginalna sr. vr.
		1	**2**	
F_2	**1**	$\bar{X}_{11}$	$\bar{X}_{12}$	$\bar{X}_{F_2,1}$
	2	$\bar{X}_{21}$	$\bar{X}_{22}$	$\bar{X}_{F_2,2}$
Marginalna sr. vr.		$\bar{X}_{F_1,1}$	$\bar{X}_{F_1,2}$	$\bar{X}$

Tablica za faktorski ANOVA test

Izvor varijacije	Suma kvadrata	Br. st. slobode	Srednja vr. sume kvadrata	Omjer F
Faktor 1	$SS(F_1)$	$k_1 - 1$	$MS(F_1)$	$\frac{MS(F_1)}{MS(E)}$
Faktor 2	$SS(F_2)$	$k_2 - 1$	$MS(F_2)$	$\frac{MS(F_2)}{MS(E)}$
Interakcija	$SS(I_{12})$	$(k_1 - 1)(k_2 - 1)$	$MS(I_{12})$	$\frac{MS(I_{12})}{MS(E)}$
Pogreška	$SS(E)$	$k_1 \cdot k_2 \cdot (n - 1)$	$MS(E)$	
Ukupno	$SS(T)$	$k_1 \cdot k_2 \cdot n - 1$		

- k_1 - broj nivoa 1. faktora
- k_2 - broj nivoa 2. faktora
- n - broj podataka (veličina uzorka) u svakoj kombinaciji tretmana
- $MS(F_1) = \dfrac{SS(F_1)}{k_1 - 1}$
- $MS(F_2) = \dfrac{SS(F_2)}{k_2 - 1}$
- $MS(E) = \dfrac{SS(E)}{k_1 \cdot k_2 \cdot (n - 1)}$

Primjer. Istraživač želi utvrditi efekt učestalosti vježbanja (jednom, tri i pet puta tjedno po 20 minuta) na preciznost bacanja nedominantnom rukom.

Uz to, istraživač bi htio utvrditi da li je efekt vježbanja isti kod ispitanika s iskustvom u sportu i onih bez iskustva u sportu.

3×2 dizajn eksperimenta:

- Vježbanje (3 nivoa: $1\times$ tjedno, $3\times$ tjedno, $5\times$ tjedno)
- Iskustvo (2 nivoa: sportaši, nesportaši)

Za svaku grupu izabran je uzorak iste veličine (3 ispitanika)

Izabrano je 9 sportaša i 9 studenata bez iskustva u sportu.

Nakon 6 tjedana vježbanja ispitanici su testirani.

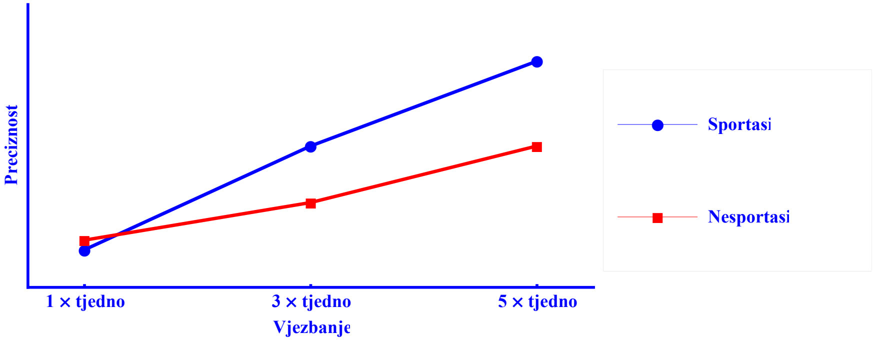
Podaci:

		Iskustvo	
		Sportaši	**Nesportaši**
Vježbanje	**1 $\times$ tjedno**	1	2
		2	2
		1	1
	3 $\times$ tjedno	3	3
		5	2
		7	4
	5 $\times$ tjedno	8	5
		7	6
		9	4

Srednje vrijednosti:

		Iskustvo		Marginalna sr. vr.
		Sportaši	Nesportaši	
Vježbanje	$1 \times$ tjedno	1.33	1.67	1.50
	$3 \times$ tjedno	5.00	3.00	4.00
	$5 \times$ tjedno	8.00	5.00	6.50
Marginalna sr. vr.		4.78	3.22	**4.00**

- Učestalost treninga utječe na preciznost.
- Iskustvo utječe na preciznost.
- Povećanje učestalosti treninga više utječe na preciznost kod sportaša.



- Postoji razlika izmeđnivoa za svaki faktor.
- Pravci nisu paralelni → postoji interakcija.

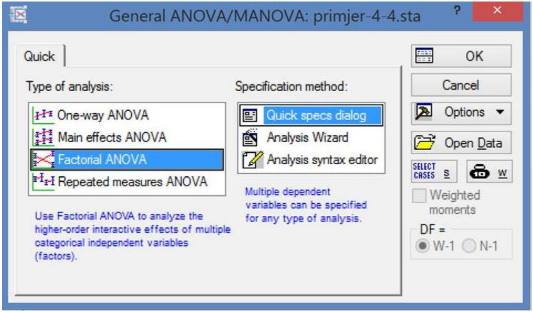
Analiza pomoću Statistice.

Podaci:

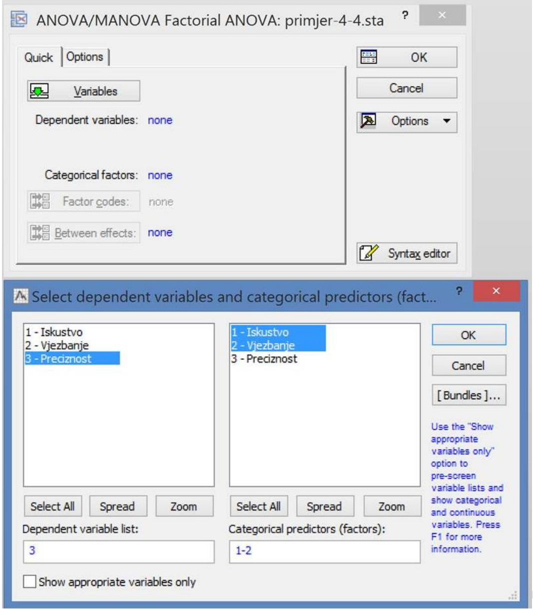
Data: primjer-4-4.sta (3v by ...

	1 Iskustvo	2 Vjezbanje	3 Preciznost
1	1	1	1
2	1	1	2
3	1	1	1
4	1	2	3
5	1	2	5
6	1	2	7
7	1	3	8
8	1	3	7
9	1	3	9
10	2	1	2
11	2	1	2
12	2	1	1
13	2	2	3
14	2	2	2
15	2	2	4
16	2	3	5
17	2	3	6
18	2	3	4

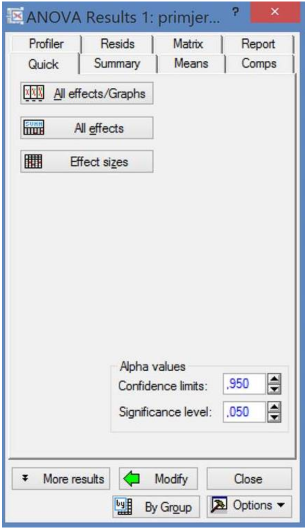
Izbor faktorskog ANOVA testa



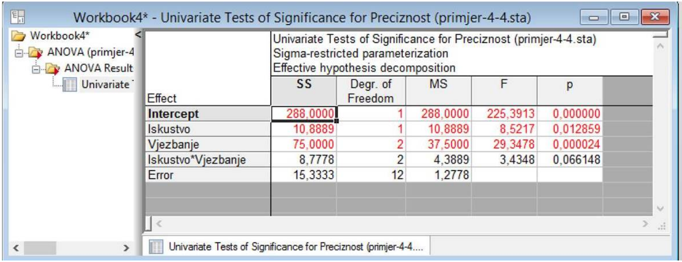
Definiranje varijabli



Izbor prikaza rezultata



ANOVA tablica

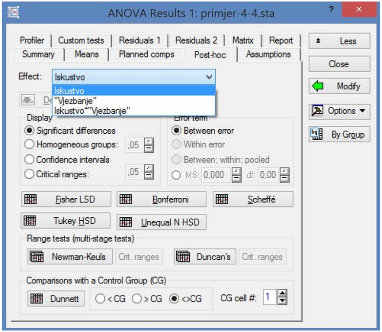


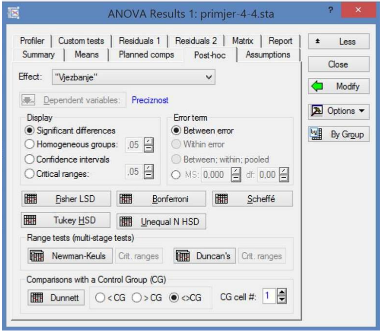
Univariate Tests of Significance for Preciznost (primjer-4-4.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

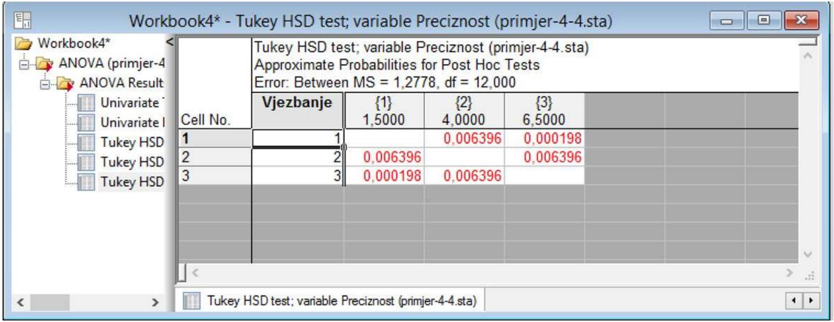
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	288,0000	1	288,0000	225,3913	0,000000
Iskustvo	10,8889	1	10,8889	8,5217	0,012859
Vjezbanje	75,0000	2	37,5000	29,3478	0,000024
Iskustvo*Vjezbanje	8,7778	2	4,3889	3,4348	0,066148
Error	15,3333	12	1,2778		

’Post-hoc’ test

Testiramo faktor Vježbanje





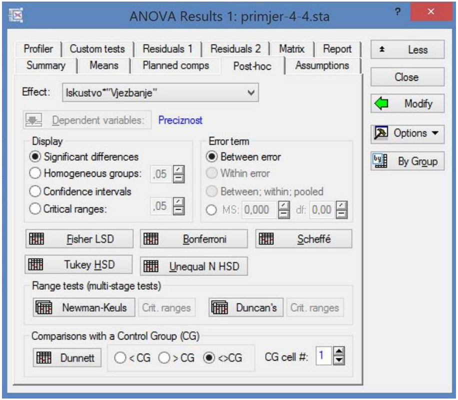


Workbook4* - Tukey HSD test; variable Preciznost (primjer-4-4.sta)

Tukey HSD test; variable Preciznost (primjer-4-4.sta)
Approximate Probabilities for Post Hoc Tests
Error: Between MS = 1,2778, df = 12,000

Cell No.	Vjezbanje	{1} 1,5000	{2} 4,0000	{3} 6,5000
1	1		0,006396	0,000198
2	2	0,006396		0,006396
3	3	0,000198	0,006396	

Testiramo interakciju



Workbook4* - Tukey HSD test; variable Preciznost (primjer-4-4.sta)

Workbook4*
ANOVA (primjer-4
ANOVA Result
Univariate
Univariate
Tukey HSD
Tukey HSD
Tukey HSD
Tukey HSD

Tukey HSD test; variable Preciznost (primjer-4-4.sta)
Approximate Probabilities for Post Hoc Tests
Error: Between MS = 1,2778, df = 12,000

Cell No.	Iskustvo	Vjezbanje	{1} 1,3333	{2} 5,0000	{3} 8,0000	{4} 1,6667	{5} 3,0000	{6} 5,0000
1	1	1		0,017827	0,000253	0,999008	0,497075	0,017827
2	1	2	0,017827		0,060017	0,032759	0,318844	1,000000
3	1	3	0,000253	0,060017		0,000322	0,001772	0,060017
4	2	1	0,999008	0,032759	0,000322		0,702054	0,032759
5	2	2	0,497075	0,318844	0,001772	0,702054		0,318844
6	2	3	0,017827	1,000000	0,060017	0,032759	0,318844	

Tukey HSD test; variable Preciznost (primjer-4-4.sta)

Veličina efekta

$$R^2 = \frac{SS(F)}{SS(T)}$$

Parcijalni η^2 (η_P^2)

$$\eta_P^2 = \frac{SS(F)}{SS(F) + SS(E)}$$

$$\omega^2 = \frac{SS(F) - (k-1)MS(E)}{SS(T) + MS(E)}$$

U našem primjeru:

$$R^2(\text{Iskustvo}) = \frac{SS(F)}{SS(T)} = \frac{10.89}{110} = 0.09899$$

$$R^2(\text{Vježbanje}) = \frac{SS(F)}{SS(T)} = \frac{75}{110} = 0.68182$$

$$R^2(\text{Iskustvo} \times \text{Vježbanje}) = \frac{SS(F)}{SS(T)} = \frac{8.78}{110} = 0.07980$$

9.9% varijance je posljedica varijabilnosti faktora Iskustvo

68.2% varijance je posljedica varijabilnosti faktora Vježbanje

8.0% varijance je posljedica varijabilnosti interakcije faktora Iskustvo i Vježbanje

ANOVA - ponovljena mjerenja

Istraživač želi ispitati smanjenje osjećaja ravnoteže koji biciklisti osjete povećanjem umora tijekom utrke.

Da bi izmjerio gubitak ravnoteže, istraživač je postavio trkaći bicikl na ergometar s valjcima.

Na sredini prednjeg cilindra obojana je bijela traka širine 10 cm. Zadnji valjak je povezan s kočionim sustavom da osigura otpor zadnjem kotaču. Pogreške u ravnoteži su mjerene brojanjem skretanja prednjeg kotača s bijele trake širine 10 cm.

Povećanjem otpora, povećava se umor i postaje sve teže održati prednji kotač na bijeloj traci.

Ispitanik vozi bicikl 15 minuta. Taj interval je podijeljen u 3-minutne periode za prikupljanje podataka.

Broj pogrešaka ravnoteže je mjeren u zadnjoj minuti 3-minutnog perioda i na kraju 3-minutnog perioda je povećan otpor.

Zavisna varijabla je pogreška u ravnoteži (broj pogrešaka u minuti).

Faktor je varijabla otpor (umor).

U istraživanju je provedeno 5 skupova mjerenja.

Na svakom ispitaniku provedeno je 5 mjerenja.

Uzorci su **zavisni**.

Ne možemo koristiti standardni ANOVA test.

Koristimo ANOVA test za ponovljena mjerenja.

Podaci.

Ispitanik	Minuta 3	Minuta 6	Minuta 9	Minuta 12	Minuta 15
1	7	7	23	36	70
2	12	22	26	26	20
3	11	6	9	31	30
4	10	18	16	40	25
5	6	12	9	28	37
6	13	21	30	55	65
7	5	0	2	10	11
8	15	18	22	37	42
9	0	2	0	16	11
10	6	8	27	32	54

Srednje vrijednosti.

Ispitanik	T_1	T_2	$\dots$	T_k	Sr. vr.
1	X_{11}	X_{12}	$\dots$	X_{1k}	$\bar{X}_{I1}$
2	X_{21}	X_{22}	$\dots$	X_{2k}	$\bar{X}_{I2}$
3	X_{31}	X_{32}	$\dots$	X_{3k}	$\bar{X}_{I3}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
n	X_{n1}	X_{n2}	$\dots$	X_{nk}	$\bar{X}_{In}$
Sr. vr.	$\bar{X}_1$	$\bar{X}_2$	$\dots$	$\bar{X}_k$	$\bar{X}$

Sume kvadrata se računaju slično kao kod faktorskog ANOVA testa.

Sume kvadrata.

$$
\begin{aligned}
SS(F) &= n \cdot \sum_i (\bar{X}_i - \bar{X})^2 \\
SS(I) &= k \cdot \sum_i (\bar{X}_{I,i} - \bar{X})^2 \\
SS(T) &= \sum_i \sum_j (X_{ij} - \bar{X})^2 \\
SS(E) &= SS(T) - SS(F) - SS(I)
\end{aligned}
$$

- $SS(F)$ - suma kvadrata za efekt (faktor)
- $SS(I)$ - suma kvadrata za ispitanike
- $SS(E)$ - suma kvadrata za pogreške
- $SS(T)$ - ukupna suma kvadrata

Tablica za ANOVA test za ponovljena mjerenja

Izvor varijacije	Suma kvadrata	Br. st. slobode	Srednja vr. sume kvadrata	Omjer F
Faktor	$SS(F)$	$k-1$	$MS(F)$	$\frac{MS(F)}{MS(E)}$
Pogreška	$SS(E)$	$(k-1)\cdot(n-1)$	$MS(E)$	
Ispitanici	$SS(I)$	$n-1$		
Ukupno	$SS(T)$	$n\cdot k-1$		

Može se prikazati samo

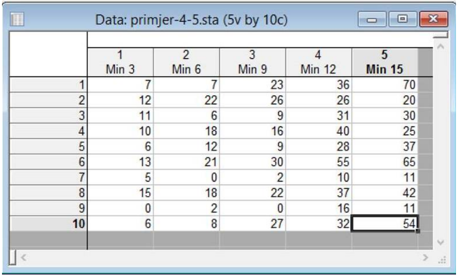
Izvor varijacije	**Suma kvadrata**	**Br. st. slobode**	**Srednja vr. sume kvadrata**	**Omjer F**
Faktor	$SS(F)$	$k-1$	$MS(F)$	$\frac{MS(F)}{MS(E)}$
Pogreška	$SS(E)$	$(k-1)\cdot(n-1)$	$MS(E)$	

Za naš primjer:

Izvor varijacije	**Suma kvadrata**	**Br. st. slobode**	**Srednja vr. sume kvadrata**	**Omjer F**
Umor	6,155.88	4	1,528.97	18.36
Pogreška	2,998.12	36	83.28	
Ispitanici	4,242.58	9		
Ukupno	13,356.58	49		

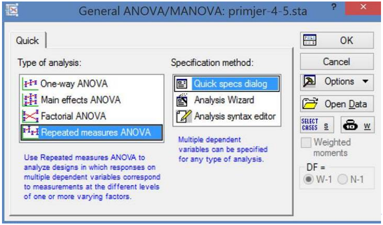
Rješenje pomoću Statistice

Podaci:

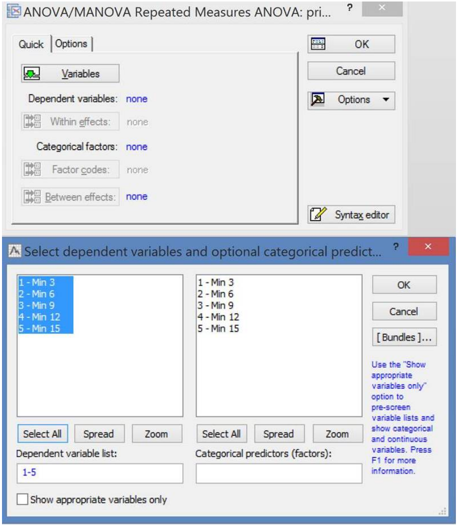

Data: primjer-4-5.sta (5v by 10c)

	1 Min 3	2 Min 6	3 Min 9	4 Min 12	5 Min 15
1	7	7	23	36	70
2	12	22	26	26	20
3	11	6	9	31	30
4	10	18	16	40	25
5	6	12	9	28	37
6	13	21	30	55	65
7	5	0	2	10	11
8	15	18	22	37	42
9	0	2	0	16	11
10	6	8	27	32	54

Izbor testa: ANOVA za ponovljena mjerenja

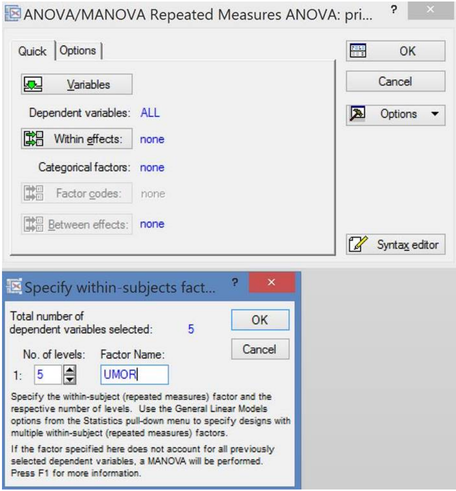


Definiranje varijabli



Faktor se ne definira ovdje.

Definiranje faktora (zavisnih mjerenja)



ANOVA tablica

Repeated Measures Analysis of Variance (primjer-4-5.sta)

Repeated Measures Analysis of Variance (primjer-4-5.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	SS	Degr. of Freedom	MS	F	p
Intercept	21590,42	1	21590,42	45,80085	0,000082
Error	4242,58	9	471,40		
UMOR	6115,88	4	1528,97	18,35915	0,000000
Error	2998,12	36	83,28		

Repeated Measures Analysis of Variance (primjer-4-5.sta)

'Post-hoc' test

ook5* - Tukey HSD test; variable DV_1 (primjer-4-5.sta)

Tukey HSD test; variable DV_1 (primjer-4-5.sta)
Approximate Probabilities for Post Hoc Tests
Error: Within MS = 83,281, df = 36,000

Cell No.	UMOR	{1} 8,5000	{2} 11,400	{3} 16,400	{4} 31,100	{5} 36,500
1	Min 3		0,952787	0,317695	0,000146	0,000125
2	Min 6	0,952787		0,737043	0,000358	0,000127
3	Min 9	0,317695	0,737043		0,008017	0,000297
4	Min 12	0,000146	0,000358	0,008017		0,678964
5	Min 15	0,000125	0,000127	0,000297	0,678964	

Repeated Measures Analysis of Variance (primjer-4-5.sta) | Tukey HSD test; variable DV_1 (primj

Pretpostavke ANOVA testa za ponovljena mjerenja

- normalnost podataka
- **sferičnost**
  - homogenost varijanci
  - **homogenost kovarijanci** između tretmana
- testiranje sferičnosti: Mauchlyjev test

Korekcije za narušene pretpostavke o sferičnosti

Greenhouse-Geisserova korekcija

- GG procjena sferičnosti ε
- ε poprima vrijednosti između 0 (maksimalno narušena sferičnost) i 1 (sferičnost nije narušena).
- računanje procjene sferičnosti ε je složeno $\rightarrow$ primjena računala
- Korekcija: stupnjevi slobode za faktor i pogreške se množe s faktorom ε.
- Primjena GG korekcije pretpostavlja maksimalno kršenje pretpostavke o sferičnosti.
- ukoliko je kršenje pretpostavke o sferičnosti malo, tada je smanjenje broja stupnjeva slobode preveliko.
- $\rightarrow$ pogreška II. vrste

Huynh-Feldtova korekcija

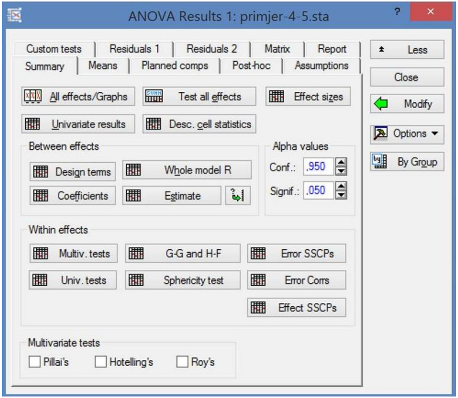
- HF procjena sferičnosti ε
- ε vrijednost za GG je konzervativnija nego ona za HF $\rightarrow$ GG pruža bolju zaštitu od pogreške I. vrste.
- HF GG pruža bolju zaštitu od pogreške II. vrste.

Strategija za određivanje značajnosti omjera F

- Izračunati p-vrijednost za F pomoću GG korekcije.
- Ako je F s GG korekcijom značajan ($p < \alpha$), odbaciti H_0.
- Ako F s GG korekcijom nije značajan, izračunati p-vrijednost za F bez korekcije
- Ako F bez korekcije nije značajan, prihvatititi H_0
- Ako F s GG korekcijom nije značajan ali je F bez korekcije značajan, koristiti HF korekciju za krajnju odluku.

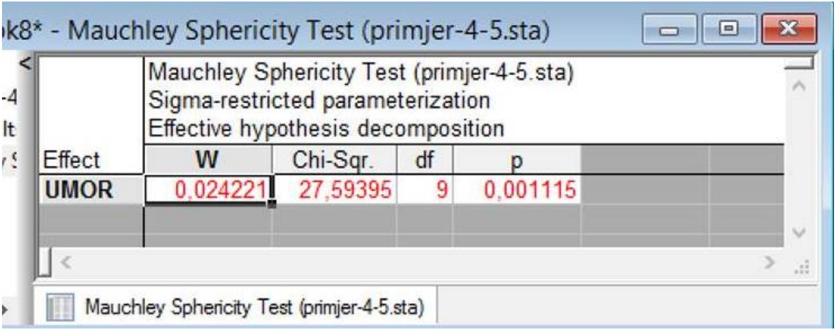
Testiranje sferičnosti i korekcije pomoću Statistice

Izbornik Summary



'Sphericity test' i 'G-G and H-F'

Mauchlyjev test



Mauchley Sphericity Test (primjer-4-5.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	W	Chi-Sqr.	df	p
UMOR	0,024221	27,59395	9	0,001115

Značajno narušena sferičnost ($p = 0.00115$)

GG i HF korekcija

Workbook8* - Adjusted Univariate Tests for Repeated Measure: DV_1 (primjer-4-5.sta)

Adjusted Univariate Tests for Repeated Measure: DV_1 (primjer-4-5.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	Degr. of Freedom	F	p	G-G Epsilon	G-G Adj. df1	G-G Adj. df2	G-G Adj. p	H-F Epsilon	H-F Adj. df1	H-F Adj. df2	H-F Adj. p	Lowr.Bnd Epsilon
UMOR	4	18,35915	0,000000	0,371312	1,485246	13,36722	0,000316	0,427574	1,710297	15,39268	0,000135	0,250000
Error	36											

Mauchley Sphericity Test (primjer-4-5.sta) | Adjusted Univariate Tests for Repeated Measure: DV_1 (primjer...

Određivanje značajnosti omjera F

- p-vrijednost za F pomoću GG korekcije.

  $p = 0.000316$

- Ako je F s GG korekcijom značajan ($p < \alpha$), odbaciti H_0.

  Odbacujemo H_0. Srednje vrijednosti su različite.