

ANALIZA VARIJANCE

ANOVA

Analysis of variance

Analysis of means using variance

ANALIZA VARIJANCE

- ANOVA je tehnika usporedbe razlika između prosjeka uzoraka u zaključivanju o postojanju (ili ne) razlika između prosjeka populacija
- Jedinствени postupak kojim je moguće raščlaniti i procijeniti varijabilnosti uvjetovane različitim čimbenicima – izvorima

◆ t-test kad su samo dva uzorka

◆ H_0 :

- t-test vrijedi samo za **slučajne** usporedbe

- gubi se na preciznosti i pouzdanosti računanja **varijanci**
koja je uvjetovana varijabilnošću svih grupa

Pretpostavke za ANOVA-u

- Opažaji (*observations*) su **normalno** distribuirani unutar svake populacije
- Varijance populacija su jednake – **homogene** (*homogeneity of variance or homoscedasticity*)
- Opažaji su **neovisni**

ANALIZA VARIJANCE

Usporedba većeg broja
prosječnih vrijednosti
odjednom

Nulta hipoteza H_0 :

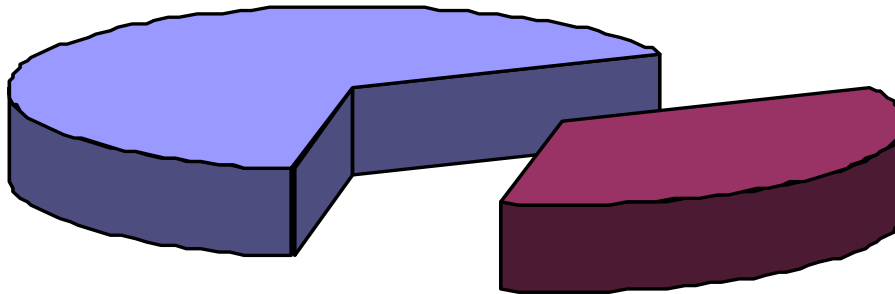
prosjeci svih populacija su
jednaki

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$

$$H_0 : \bar{x}_1 = \bar{x}_2 = \bar{x}_3 = \dots = \bar{x}_n$$

ANALIZA VARIJANCE

UKUPNA VARIJABILNOST



- između grupa
- unutar grupa (error)

ANALIZA VARIJANCE

- ◆ **si Ronald Fisher je razradio postupak:ANOVA**
- ◆ **Fisherova analiza varijance**
- ◆ **Ukupna varijabilnost se rastavlja: na varijabilnost između prosječnih vrijednosti grupa i na varijabilnost unutar grupa**
- ◆ **Varijabilnost između srednjih vrijednosti grupa v.s. varijabilnost unutar grupa**
- ◆ **Postupak za proučavanje varijabilnosti između prosječnih vrijednosti, mjeri se varijancom**

ANALIZA VARIJANCE

-test za testiranje varijanci (**F test**)

-test za testiranje razlika između prosječnih vrijednosti (**t - test**)

Kalkulacija:

- Definirati Sume kvadrata (*Sum of Squares* - **SS**)
 - SS_{total}
 - SS_{grupa}
 - SS_{error}
- Izračunati stupnjeve slobode (*degrees of freedom* - **df**)
- Izračunati srednje kvadrate tj. varijance (*mean squares* **MS**) i **F** vrijednost

Kalkulacija SS:

$$SS_{total} = \Sigma (X - \bar{X}_{..})^2$$

$$SS_{grupa} = n \Sigma (\bar{X}_j - \bar{X}_{..})^2$$

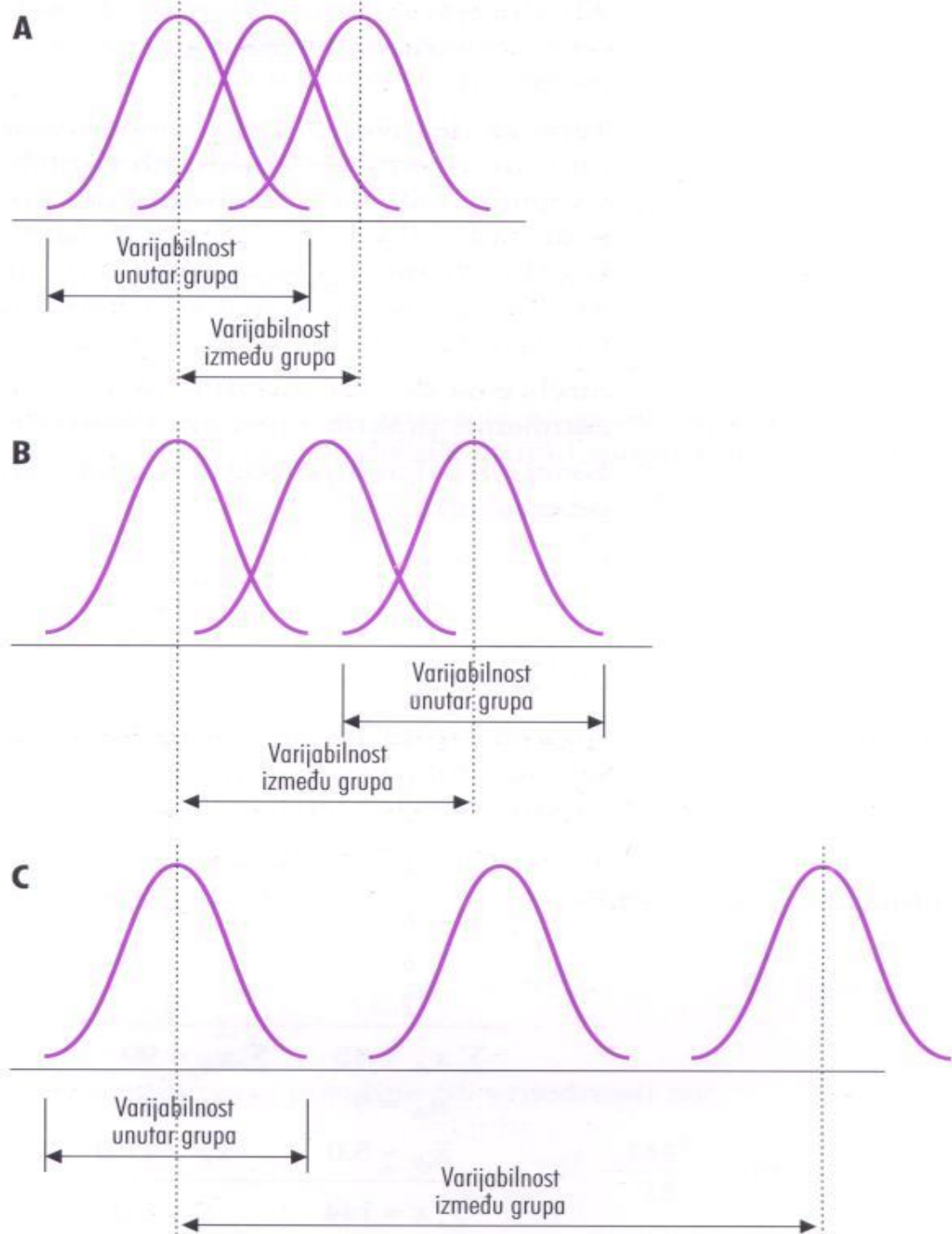
$$SS_{error} = SS_{total} - SS_{grupa}$$

- ukoliko je H_0 točna, tada varijabilnost između uzoraka nije veća od varijabilnosti unutar uzoraka (pogreške)
- procjena varijabilnosti (sume kvadrata):

$$SS_{UKUPNO} = \sum x^2 - \frac{(\sum x)^2}{n} \qquad \frac{(\sum x)^2}{n} = C.F.$$

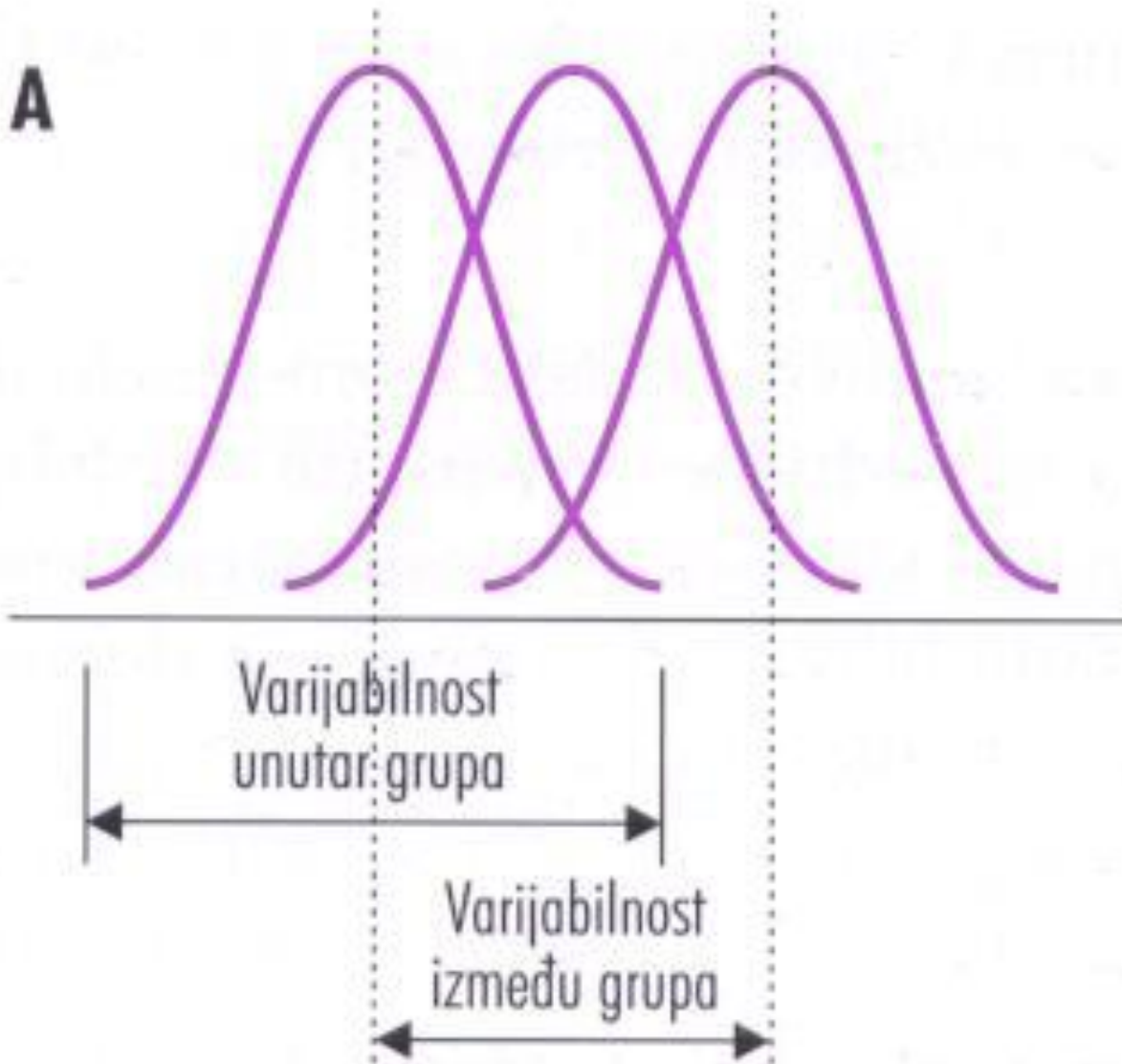
$$SS_{IZMEDU UZORAKA} = \frac{(\sum x_1)^2}{n_1} + \frac{(\sum x_2)^2}{n_2} + \dots + \frac{(\sum x_k)^2}{n_k} - C.F.$$

$$SS_{UNUTAR UZORAKA} = SS_{UKUPNO} - SS_{IZMEDU UZORAKA}$$



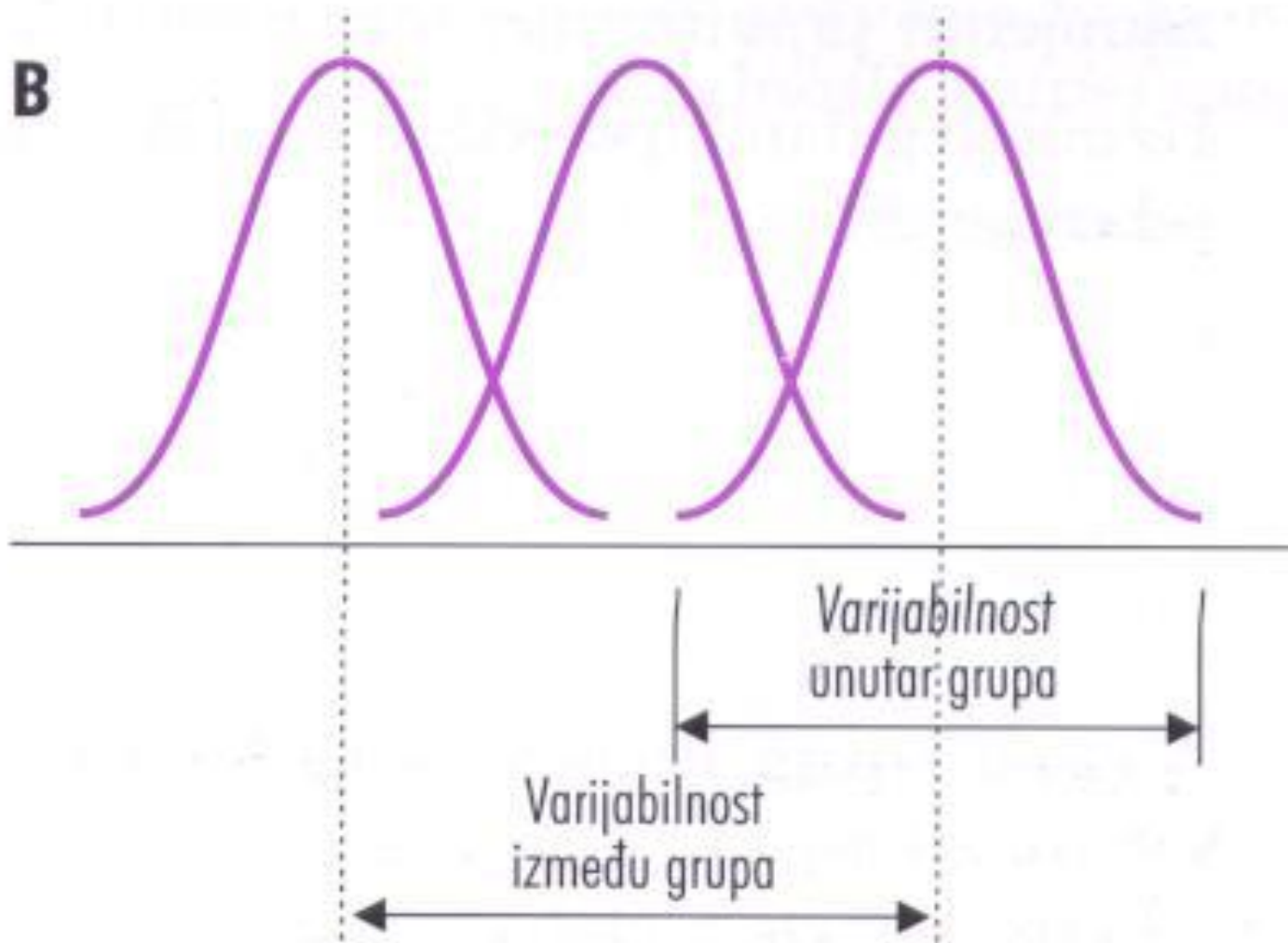
SLIKA 22.
Varijabilnost između grupa u odnosu na varijabilnost unutar grupa

između grupa < unutar grupa

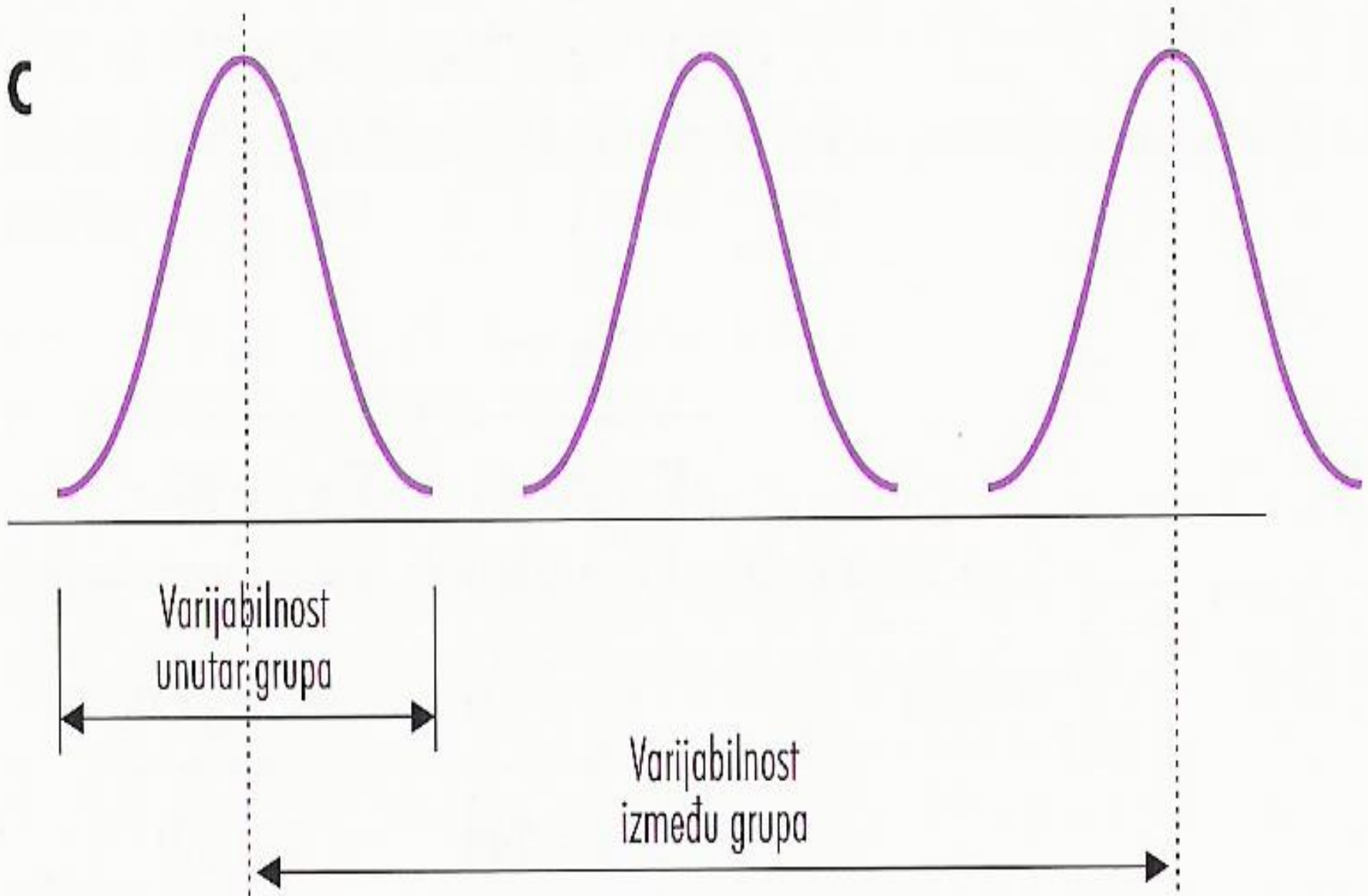


između grupa > unutar grupa

B



između grupa > unutar grupa
signifikantno!



Grafičko predstavljnje varijabilnosti

Slika A:

- varijabilnost između grupa $<$ varijabilnosti unutar grupa

Slika B:

- varijabilnost između grupa $>$ varijabilnosti unutar grupa

Slika C:

- varijabilnost između grupa $>$ varijabilnosti unutar grupa
sve varijante jedne razlikuju se od svih varijanata drugih grupa

Hipotetičan primjer: uзорak sa dvije grupe

	GRUPE			
	A	B		
	19	23		
	14	19		
	15	19		
	17	21		
	20	18		
suma	85	100	185	ukupno
n	5	5	10	
x	17	20	18,5	

Zadatak: Jednosmjerna analiza varijance > ANOVA 1

Nulta hipoteza: $H_0 : \mu_A = \mu_B$
 $H_0 : \bar{x}_A = \bar{x}_B$

	grupa A	grupa B	Ukupno
	19	23	
	14	19	
	15	19	
	17	21	
	20	18	
Σx	85,0	100,0	185,0
\bar{x}	17,0	20,0	18,5
n	5	5	10

Provesti analizu varijance:

$$H_0 : \bar{X}_A = \bar{X}_B$$

ili $H_0: D = 0$

- Odstupanja se izražavaju sumama kvadratnih odstupanja (SS)
- Svaka varijanta uzorka odstupa od prosječne vrijednosti grupe i prosječne vrijednosti cijelog uzorka
- Prosječna vrijednost grupe odstupa od prosječne vrijednosti cijelog uzorka

Ukupna varijabilnost:
odstupanje svake varijante uzorka od prosječne vrijednosti
uzorka

- $SS = \text{Suma kvadrata} = \text{Sum of Squares}$
- $SS = (19-18,5)^2 + (14-18,5)^2 + \dots + (18-18,5)^2 = 64,5$
- $SS = (19^2 + 14^2 + \dots + 18^2) - (185)^2/10 = 64,5$

Varijabilnost između grupa:

odstupanje prosjeka svake grupe od prosjeka uzorka, ali da se svaka varijanta svake grupe predstavi svojom srednjom vrijednošću

$$\blacksquare SS = 5(17-18,5)^2 + 5(20-18,5)^2 = 22,5$$

grupa A grupa B

$$\blacksquare SS = (85^2/5 + 100^2/5) - 185^2/10 = 22,5$$

grupa A + grupa B

Varijabilnost unutar grupa:
odstupanje svake varijante u grupi od prosječne vrijednosti
grupe

■ $SS = SS \text{ unutar grupe A} + SS \text{ unutar grupe B}$

■ $SS = (19-17)^2 + \dots + (20-17)^2 +$

$+ (23-20)^2 + \dots + (18-20)^2 = 42,0$

■ $SS = SS \text{ ukupno} - SS \text{ između grupa} =$

$64,5 - 22,5 = 42,0$

$$SS = \sum (x - \bar{x})^2$$

	grupa A	grupa B	Ukupno
	19	23	
	14	19	
	15	19	
	17	21	
	20	18	
Σx	85,0	100,0	185,0
\bar{x}	17,0	20,0	18,5
n	5	5	10

$$SS_{ukupno} = (19 - 18,5)^2 + (14 - 18,5)^2 + \dots + (18 - 18,5)^2 = 64,5$$

$$SS_{izmedu.grupa} = 5 \cdot (17 - 18,5)^2 + 5 \cdot (20 - 18,5)^2 = 22,5$$

$$SS_{unutar.grupa} = (19 - 17)^2 + \dots + (20 - 17)^2 + (23 - 20)^2 + \dots + (18 - 20)^2 = 42,0$$

$$SS = \sum x^2 - \frac{(\sum x)^2}{n}$$

	grupa A	grupa B	Ukupno
	19	23	
	14	19	
	15	19	
	17	21	
	20	18	
Σ x	85,0	100,0	185,0
?	17,0	20,0	18,5
n	5	5	10

$$SS_{ukupno} = (19^2 + 14^2 + \dots + 18^2) - \frac{185^2}{10} = 64,5$$

$$SS_{izmedu . grupa} = \left(\frac{85^2}{5} + \frac{100^2}{5} \right) - \frac{185^2}{10} = 22,5$$

$$SS_{unutar . grupa} = SS_{ukupno} - SS_{izmedu . grupa} = 64,5 - 22,5 = 42,0$$

ANOVA

(jednosmjerna analiza varijance)

$$F = \frac{\text{varijanca između uzoraka}}{\text{varijanca unutar uzoraka}}$$

TABLICA ANALIZE VARIJANCE:

Izvori varijabilnosti	n-1	SS	varijanca	Fexp	Ftab (za 1 i 8)
Ukupna varijabilnost	10-1=9	64,5			
Varijabilnost između grupa	2-1=1	22,5	22,5	4,29 n.s.	5,32 (5%) 11,26 (1%)
Varijabilnost unutar grupa	9-1=8	42	5,25		

Tablica analize varijance:

$$= SS : df$$

<i>Izvori varijabilnosti</i>	<i>df</i>	<i>SS</i>	<i>s²</i>	<i>F_{exp}</i>	<i>F_{tab} 0.05</i>	<i>F_{tab} 0.01</i>
Ukupno	9	64.5				
Između grupa	1	22.5	22.50	4.29 ns	5.32	11.26
Unutar grupa	8	42.0	5.25			

$$= 22.50 : 5.25$$

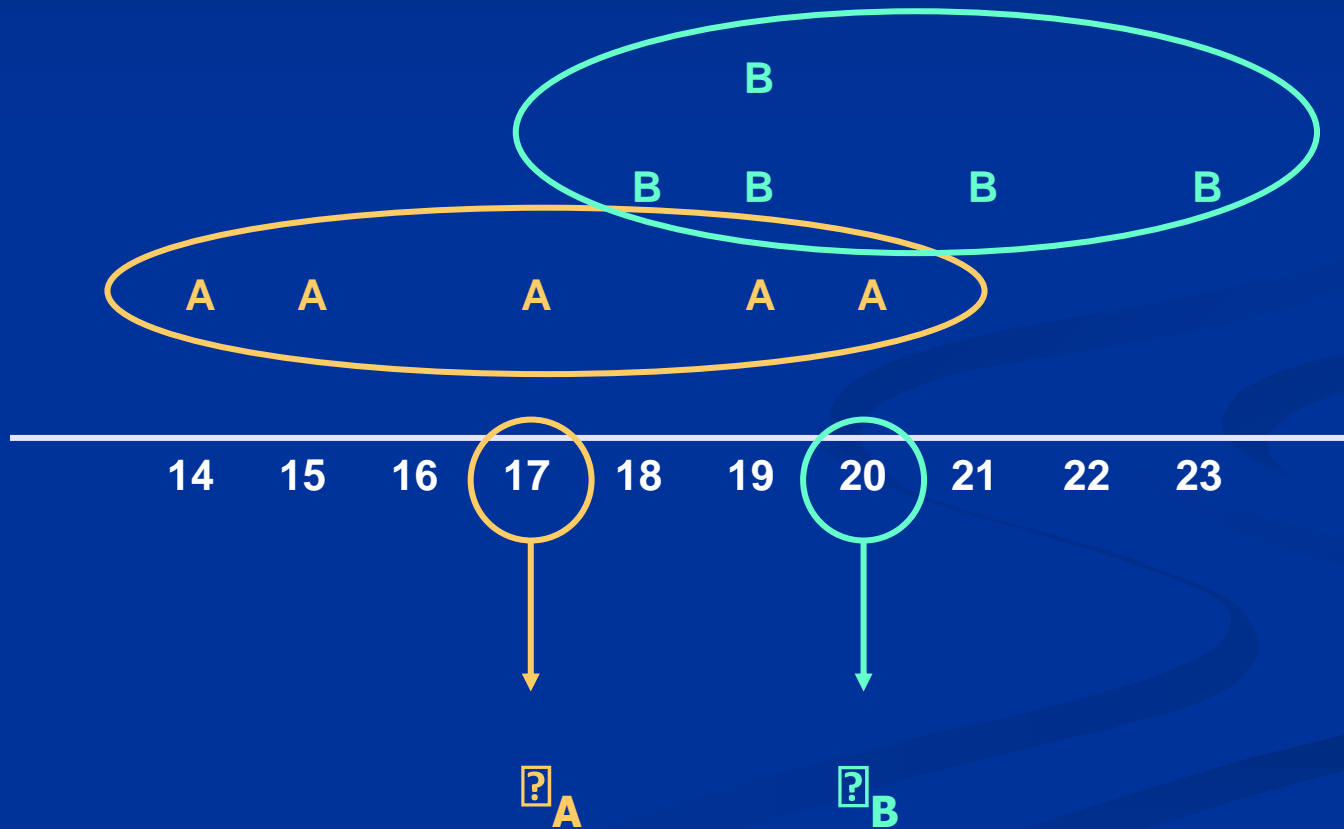
očitati iz prvog
stupca i osmog
reda F tablica

Zaključak iza F testa u ANOVA-i

1:

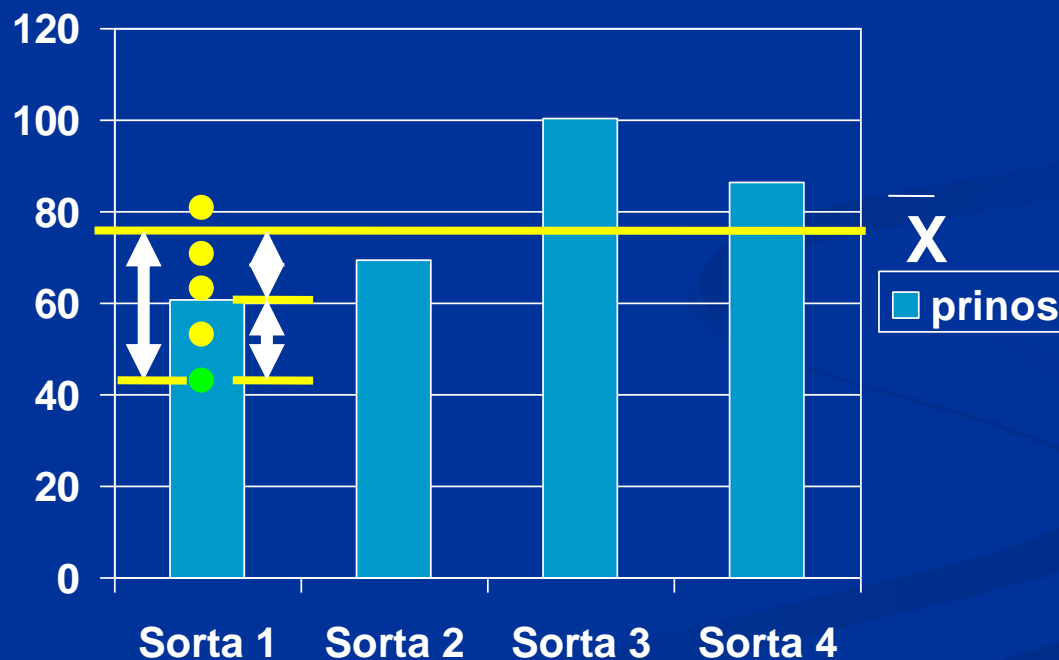
- $F_{\text{exp}} < F_{\text{tab}}$
- H_0 se prihvaća
- grupe A i B pripadaju istoj uzorku
- prosjeci grupa A i B se ne razlikuju signifikantno

grupe **A** i **B** grafički

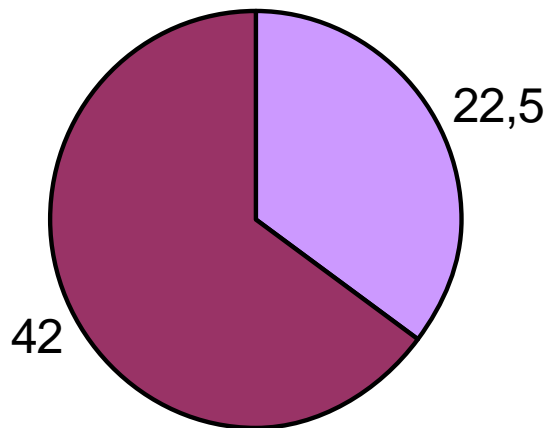


$$?_A - ?_B = D_{\text{exp}}^{\text{ns}}$$

Jedan dio **ukupne** varijabilnosti pripada razlikama **između grupa** (tretiranje, gnojidba...), a ostatak varijabilnosti pripada razlikama **unutar grupa** (error)



Ukupna varijabilnost:



ukupna varijabilnost 64,5

između grupa

unutar grupa

Ishrana gusjenica različitim travama

C	J	L	M
206	114	90	140
178	253	270	210
255	313	173	354
300	112	97	124
168	130	62	199
142	157	193	273
296	198	211	221

TABLICA ANALIZE VARIJANCE:

Izvori varijabilnosti	n-1	SS	varijanca	Fexp	Ftab (za 3i 24)
Ukupna varijabilnost	27	150.209,25			
Varijabilnost između grupa	3	19.532,96	6.510,99	1,20 n.s.	3,01 (5%) 4,72 (1%)
Varijabilnost unutar grupa	24	130.676,29	5.444,84		

A	B	C	A-Veliki jorkšir
36,6	31	34,9	B-F1 generacija švedskog landrasa i velikog jorkšira
36,1	27,2	32,6	C-nizozemski landras i veliki jorkšir
36,1	26,9	35,4	
38,1	31,3	34,8	
36,6	29	34,1	
39,2	30,5	38,2	
38,3	30	33,5	
38,6	30,8	34	
37,3	-	36,9	
38,1	-	30,7	
-	-	34	
-	-	32,2	

TABLICA ANALIZE VARIJANCE:

Izvori varijabilnosti	n-1	SS	varijanca	Fexp	Ftab (za 2 i 27)
Ukupna varijabilnost	29	355,48			
Varijabilnost između grupa	2	278,87	139,43	49,09 **	3,35 (5%) 5,49 (1%)
Varijabilnost unutar grupa	27	76,61	2,84		

Tipovi masti

1	2	3	4
64	78	75	55
72	91	93	66
68	97	78	49
77	82	71	64
56	85	63	70
95	77	76	68
432	510	456	372
72	85	76	62

TABLICA ANALIZE VARIJANCE:

Izvori varijabilnosti	n-1	SS	varijanca	Fexp	Ftab (za 3 i 20)
Ukupna varijabilnost	23	3654			
Varijabilnost između grupa	3	1636	545,33	5,40 **	3,10 (5%) 4,94 (1%)
Varijabilnost unutar grupa	20	2018	100,9		