

# Osnovni testovi u SPSS-u

Cilj ovog teksta je da se **ukratko** opiše put za realizaciju nekih statističkih metoda i da se daju osnovna tumačenja dobijenih rezultata.

## Frekvencije i procenti

Analyze – Descriptive Statistics – Frequencies

Varujablu prebaciti u prozor Variable(s). Kliknuti **OK**.

## Aritmetička sredina, standardna devijacija i druge statistike

Analyze – Descriptive Statistics – Descriptives

Varujablu prebaciti u prozor Variable(s) i kliknuti **OK**.

Ako se želi više statistike, onda se klikne na **Options** i štiklira željena statistika (Range, Variance, S.E. mean i slično). Kliknuti **Continue** i **OK**.

## Aritmetička sredina, standardna devijacija, medijana i druge statistike

Analyze – Descriptive Statistics – Explore

Varujablu prebaciti u prozor Variable(s). Kliknuti **OK**.

Kada se u prozor **Factor List** prebaci varijabla koja određuje populacije (grupna varijabla), onda program daje aritmetičku sredinu, standardnu devijaciju, medijanu i ostalo, za svaki uzorak posebno.

Ako se žele percentili (samim tim medijana i kvartili), onda se klikne **Statistics** i štiklira **Percentiles**. Kliknuti **Continue** i **OK**.

## Ispitivanje normalnosti raspodele

Analyze – Descriptive statistics – Explore

Varujablu prebaciti u prozor Variable(s), kliknuti na **Plots**, zatim štiklirati **Normality plots with tests** zatim kliknuti na **Continue** i **OK**.

Ako je obim uzorka veći od 50 gleda se rezultat testa Kolmogorov-Smirnov. U suprotnom gleda se rezultat testa Shapiro-Wilk.

Ako se u prozor **Factor List** prebaci varijabla koja određuje populacije (grupna varijabla), onda se dobija rezultat ispitivanja normalnosti promenljive u

svakoj populaciji. Ovde ima onoliko nultih hipoteza koliko ima populacija i svaka od njih tvrdi da varijabla ima normalnu raspodelu. Ako je Sig. veće od 0,05 hipoteza se prihvata, dok se u suprotnom hipoteza odbacuje.

### Primer.

#### Tests of Normality

	INFARKT	Kolmogoro v-Smirnov			Shapiro- Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
GODINE	<b>0,00</b>	,304	16	,000	,840	16	<b>,010</b>
	<b>1,00</b>	,156	17	,200	,953	17	<b>,489</b>

\*\* This is an upper bound of the true significance.

\* This is a lower bound of the true significance.

a Lilliefors Significance Correction

U našem primeru, za populaciju 0 se dobija Sig. = 0,010 < 0,05 pa se odbacuje nulta hipoteza da varijabla GODINE ima normalnu raspodelu u toj populaciji. Za populaciju 1 se dobija Sig. = 0,489 > 0,05 pa se prihvata hipoteza da varijabla GODINE ima normalnu raspodelu u toj populaciji.

### Testiranje hipoteze o aritmetičkoj sredini $H_0(\mu = \mu_0)$

Prvo se proveriti da li varijabla ima normalnu raspodelu a zatim

#### Analyze – Compare Means – One Sample T Test

Neprekidnu varijablu prebaciti u prozor **Test Variable(s)**. Vrednost  $\mu_0$  upisati u prozor **Test Value** (gde stoji 0) a zatim kliknuti **OK**.

### Primer.

#### One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
GODINE	<b>33</b>	<b>53,6364</b>	<b>14,2079</b>	2,4733

#### One-Sample Test

Test Value = <b>60</b>						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
GODINE	-2,573	32	<b>,015</b>	-6,3636	-11,4015	-1,3257

U prvoj tabeli se daje broj elemenata uzorka, aritmetička sredina i standardna devijacija varijable u uzorku. Ako je u drugoj tabeli Sig. (2-tailed) > 0,05 nulta hipoteza se prihvata, dok se u suprotnom hipoteza odbacuje. U našem primeru je Sig. (2-tailed) = 0,015 < 0,05 pa se nulta hipoteza (da je srednja vrednost obeležja GODINE u populaciji jednaka 60) odbacuje.

### Testiranje hipoteze o jednakosti aritmetičkih sredina

$$H_0(\mu_1 = \mu_2)$$

Prvo se proveriti da li varijabla ima normalnu raspodelu u obe populacije, pa ako ima onda

#### Analyze – Compare Means – Independent-Samples T Test

Neprekidnu varijablu prebaciti u prozor **Test Variable(s)**. Varijablu koja određuje grupe prebaciti u prozor **Grouping Variable**. Kliknuti na **Define Groups** a zatim upisati brojeve grupa (na primer, 1 i 3). Kliknuti **Continue** i **OK**.

#### Primer.

##### Group Statistics

	INFARKT	N	Mean	Std. Deviation	Std. Error Mean
TEZINA	,00	<b>16</b>	<b>68,7500</b>	<b>13,2690</b>	3,3173
	1,00	<b>17</b>	<b>73,7647</b>	<b>13,1887</b>	3,1987

##### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
TEZINA	Equal variances assumed	,002	<b>,963</b>	-1,088	31	<b>,285</b>	-5,0147	4,6074
	Equal variances not assumed			-1,088	30,855	,285	-5,0147	4,6083

U prvoj tabeli se daje broj elemenata uzorka, aritmetička sredina i standardna devijacija varijable, za svaki uzorak. Ako je u drugoj tabeli Sig. > 0,05 uzima se prva vrednost u koloni Sig. (2-tailed), dok se u suprotnom uzima druga vrednost. Ako je uzeta vrednost veća od 0,05 nulta hipoteza se prihvata, dok se u suprotnom odbacuje. U našem primeru je Sig. = 0,963 > 0,05 pa se uzima da je

Sig. (2-tailed) = 0,285. Kako je ta vrednost veća od 0,05 nulta hipoteza se prihvata.

## Izračunavanje nove varijable

### Transform – Compute

U levi prozor upisuje se ime nove varijable a u desni izraz pomoću koga se ona izračunava (na primer, SISTOLN – SISTOL2). Kliknuti **OK**.

## Upareni t test (dva merenja)

Prvo se proveriti da li razlika varijabli ima normalnu raspodelu. Ako ima, onda

### Analyze – Compare Means – Paired-Samples T Test

Varijable koje predstavljaju rezultate dva merenja nad istim objektima uneti u prozor **Paired Variables** i kliknuti **O.K**.

### Primer.

#### Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 SISTOLNI	<b>148,7879</b>	<b>33</b>	<b>20,1568</b>	3,5089
SISTOL2	<b>134,3939</b>	<b>33</b>	<b>15,7994</b>	2,7503

#### Paired Samples Test

	Paired Difference s			t	df	Sig. (2- tailed)
	Mean	Std. Deviation	Std. Error Mean			
Pair 1 SISTOLNI - SISTOL2	14,3939	15,7499	2,7417	5,250	32	<b>,000</b>

U prvoj tabeli se daju aritmetička sredina i standardna devijacija prve i druge varijable (uzorka). Ako je u drugoj tabeli Sig. (2-tailed) > 0,05 prihvata se hipoteza da su razlike između varijabli u proseku jednake nuli. U suprotnom se ta hipoteza odbacuje, odnosno zaključuje se da su razlike između ovih varijabli statistički značajne. U našem primeru je Sig. = 0,000 < 0,05 pa se zaključuje da su razlike između varijabli statistički značajne. Broj 0,000 je zaokružen, a zapravo nije jednak 0. U ovom slučaju se piše  $p < 0,0005$ .

## Analiza varijansi (ANOVA) $H_0(\mu_1 = \mu_2 = \dots = \mu_r)$

Prvo treba ispitati da li varijabla ima normalnu raspodelu u svim populacijama koristeći

Analyze - Descriptive – Explore – Plots – Normality plots

Ako ima, onda

Analyze - Compare Means – One-way-ANOVA.

Neprekidnu varijablu prebaciti u prozor **Dependent List**. Varijablu koja određuje populacije prebaciti u prozor **Factor**. Korisno je da se posle ovoga klikne na **Options** i štiklira **Descriptive**. Kliknuti na **OK**.

Ako je Sig. > 0,05, onda se prihvata nulta hipoteza o jednakosti srednjih vrednosti (aritmetičkih sredina). U suprotnom, hipoteza se odbacuje. Da bi se ispitalo između kojih populacija je razlika značajna, ispituje se najpre homogenost

Analyze - Compare Means - One-way-ANOVA – Options - Homogeneity

a zatim

Analyze - Compare Means – One-way-ANOVA - Post Hoc

Ako je rezultat ispitivanja homogenosti Sig. < 0,05 onda se koristiti neka od donjih metoda (na primer, Dunnett's T3). U suprotnom koristi se neka od gornjih metoda (na primer, Bonferroni).

### Primer.

Descriptives  
FRAK2

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1,00	29	<b>34,4310</b>	<b>14,9812</b>	2,7819	28,7325	40,1296	8,90	63,60
2,00	22	<b>26,1364</b>	<b>13,0087</b>	2,7735	20,3686	31,9041	6,20	59,60
3,00	24	<b>19,1250</b>	<b>9,7706</b>	1,9944	14,9993	23,2507	4,60	41,30
4,00	25	<b>20,0240</b>	<b>9,7536</b>	1,9507	15,9979	24,0501	6,00	41,00
Total	100	25,3310	13,6186	1,3619	22,6288	28,0332	4,60	63,60

ANOVA  
FRAK2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4044,230	3	1348,077	9,039	<b>,000</b>
Within Groups	14316,824	96	149,134		
Total	18361,054	99			

Kako je rezultat Sig. < 0,05 odbacuje se hipoteza o jednakosti srednjih vrednosti.

Test of Homogeneity of Variances  
FRAK2

Levene Statistic	df1	df2	Sig.
2,962	3	96	<b>,036</b>

Sig. = 0,036 < 0,05 ukazuje da treba primeniti neku od donjih metoda.

Primenom testa Dunnett's T3 dobija se

Multiple Comparisons  
Dependent Variable: FRAK2

**Dunnett T3**

		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I) GRUP	(J) GRUP				Lower Bound	Upper Bound
1,00	2,00	8,2947	3,4527	,212	-2,4608	19,0502
	3,00	15,3060	3,3699	<b>,000</b>	5,9418	24,6702
	4,00	14,4070	3,3328	<b>,001</b>	5,1119	23,7021
2,00	1,00	-8,2947	3,4527	,212	-19,0502	2,4608
	3,00	7,0114	3,6045	,243	-2,4288	16,4515
	4,00	6,1124	3,5699	,378	-3,2579	15,4826
3,00	1,00	-15,3060	3,3699	,000	-24,6702	-5,9418
	2,00	-7,0114	3,6045	,243	-16,4515	2,4288
	4,00	-,8990	3,4899	1,000	-8,5438	6,7458
4,00	1,00	-14,4070	3,3328	,001	-23,7021	-5,1119
	2,00	-6,1124	3,5699	,378	-15,4826	3,2579
	3,00	,8990	3,4899	1,000	-6,7458	8,5438

\* The mean difference is significant at the .05 level.

Razlika između populacije 1 i populacije 3 je statistički značajna (Sig. = 0,000).  
Razlika između populacije 1 i populacije 4 je statistički značajna (Sig. = 0,001).  
Ostale razlike nisu statistički značajne.

## Mann-Whitney-ev test

Prvo se proveriti da li varijabla ima normalnu raspodelu u obe populacije, pa ako **nema** onda

Analyze - Nonparametric tests – 2 Independent Samples Test

Neprekidnu varijablu prebaciti u prozor **Test Variable List**. Varijablu koja određuje populacije prebaciti u prozor **Grouping Variable**. Kliknuti na **Define Groups** a zatim upisati brojeve populacija (na primer, 1 i 2). Kliknuti **Continue** i **OK**.

Ako je u drugoj tabeli Exact Sig. [ $2 \times (1\text{-tailed Sig.})$ ]  $> 0,05$  hipoteza o jednakosti medijana se prihvata. U suprotnom se odbacuje.

### Primer.

Ranks	GRAD	N	Mean Rank	Sum of Ranks
HOLEST	1,00	6	5,58	33,50
	2,00	9	9,61	86,50
	Total	15		

Test Statistics	
	HOLEST
Mann-Whitney U	12,500
Wilcoxon W	33,500
Z	-1,712
Asymp. Sig. (2-tailed)	,087
Exact Sig. [ $2 \times (1\text{-tailed Sig.})$ ]	<b>,088</b>

a Not corrected for ties.

b Grouping Variable: GRAD

Kako je Exact Sig. [ $2 \times (1\text{-tailed Sig.})$ ] = 0,088  $> 0,05$  hipoteza o jednakosti medijana promenljive HOLEST se prihvata.

## Wilcoxon-ov test

Prvo se proverava da li **razlika** varijabli ima normalnu raspodelu. Ako **nema**, onda

### Analyze - Nonparametric tests – 2 Related Test

Varijable koje predstavljaju rezultate dva merenja nad istim objektima uneti u prozor **Test Pair(s) List** i kliknuti **O.K.**

Ako je u drugoj tabeli Asymp. Sig. (2-tailed Sig) > 0,05 prihvata se hipoteza da su razlike između varijabli jednake nuli. U suprotnom se odbacuje.

### Primer.

Ranks		N	Mean Rank	Sum of Ranks
HOLEST2 - HOLESTER	Negative Ranks	18	14,08	253,50
	Positive Ranks	6	7,75	46,50
	Ties	9		
	Total	33		
a HOLEST2 < HOLESTER				
b HOLEST2 > HOLESTER				
c HOLESTER = HOLEST2				

Test Statistics	
HOLEST2 - HOLESTER	
Z	-2,960
Asymp. Sig. (2-tailed)	<b>,003</b>

- a Based on positive ranks.
- b Wilcoxon Signed Ranks Test

Kako je Asimpt. Sig. (2-tailed Sig) = 0,003 < 0,05 odbacuje se hipoteza da su razlike između varijabli HOLEST i HOLEST2 jednake nuli, odnosno razlike između ovih varijabli su statistički značajne.



## Kruskal-Wallis-ov test

Prvo treba ispitati da li varijabla ima normalnu raspodelu u svim populacijama koristeći

Analyze - Descriptive – Explore – Plots – Normality plots

Ako **nema**, onda

Analyze - Nonparametric tests – K indenpedent-Samples Test

Neprekidnu varijablu prebaciti u prozor **Test Variable List**. Varijablu koja određuje populacije prebaciti u prozor **Grouping Variable**, kliknuti na **Define Range**, uneti najmanju i najveću oznaku grupe (na primer, 1 i 4) i kliknuti **Continue** i **O.K.**

Ako je u drugoj tabeli Asymp. Sig. > 0,05 prihvata se hipoteza o jednakosti medijana. U suprotnom se odbacuje.

### Primer.

Ranks			
	GRAD	N	Mean Rank
HOLEST	1,00	6	11,17
	2,00	9	19,00
	3,00	15	15,13
	Total	30	

Test Statistics	
	HOLEST
Chi-Square	2,907
df	2
Asymp. Sig.	<b>,234</b>

a Kruskal Wallis Test

b Grouping Variable: GRAD

Kako je Asimp. Sig. = 0,234 > 0,05 prihvata se hipoteza o jednakosti medijana varijable HOLEST između populacija.

## Friedman-ov test

### Analyze - Nonparametric tests – K Related Samples Test

Varijable koje predstavljaju rezultate više od dva merenja nad istim objektima uneti u prozor **Test Variables** i kliknuti **O.K.**

Ako je u drugoj tabeli Asymp. Sig. > 0,05 prihvata se hipoteza da su razlike između varijabli (merenja) jednake nuli. U suprotnom se odbacuje.

### Primer.

Ranks	
	Mean Rank
HORM1	2,32
HORM2	1,78
HORM3	1,90

Test Statistics	
N	106
Chi-Square	18,509
df	2
Asymp. Sig.	,000

a. Friedman Test

Kako je Asimp. Sig. = 0,000 < 0,05 odbacuje se hipoteza da su razlike između varijabli (merenja) HORM1, HORM2, HORM3 jednake nuli, odnosno razlike između njih su statistički značajne.

## Tabele kontingencije (Hi-kvadrat test)

### Analyze – Descriptive Statistics – Crosstabs

Jednu kategorijsku varijablu uneti u prozor **Row(s)** a drugu u prozor **Column(s)**, zatim kliknuti **Statistics**, štiklirati **Chi-square**, kliknuti na **Continue**, kliknuti na **Cells**, štiklirati **Row**, kliknuti na **Continue** i kliknuti **O.K.**

Ako je u drugoj tabeli Asymp. Sig. (2-sided), u vrsti Pearson Chi-Square, veća od 0,05 prihvata se hipoteza o nezavisnosti varijabli. U suprotnom se odbacuje.

Ako je prva tabela oblika 2 x 2 i sve frekvencije u tabeli su veće ili jednake 5, a Asymp. Sig. (2-sided) u vrsti Continuity Correction veća od 0,05 prihvata se hipoteza o nezavisnosti varijabli. U suprotnom se odbacuje.

Ako je prva tabela oblika 2 x 2 i neka od frekvencija u tabeli je manja od 5, a Exact Sig. (2-sided) u vrsti Fisher's Exact Test veća od 0,05 prihvata se hipoteza o nezavisnosti varijabli. U suprotnom se odbacuje.

## Primer.

OBRAZOVA \* AHOL01 Crosstabulation

			AB		Total
			,00	1,00	
OBRAZOV	sss	Count	26	26	52
		% within OBRAZOV A	50,0%	50,0%	100,0%
	všs	Count	33	30	63
		% within OBRAZOV A	52,4%	47,6%	100,0%
	vss	Count	85	79	164
		% within OBRAZOV A	51,8%	48,2%	100,0%
	mr_sci	Count	38	38	76
		% within OBRAZOV A	50,0%	50,0%	100,0%
	dr_sci	Count	15	30	45
		% within OBRAZOV A	33,3%	66,7%	100,0%
Total		Count	197	203	400
		% within OBRAZOV A	49,3%	50,8%	100,0%

## Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,274	4	<b>,260</b>
Likelihood Ratio	5,369	4	,251
Linear-by-Linear Association	2,064	1	,151
N of Valid Cases	400		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 22,16.

Kako je u drugoj tabeli Asymp. Sig. (2-sided) = 0,260 > 0,05 prihvata se hipoteza o nezavisnosti varijabli OBRAZOV i AB.

## Primer.

### Crosstab

			DUVAN		Total
POL	muski	Count	,00	1,00	
		% within POL	192	208	400
	zenski	Count	48,0%	<b>52,0%</b>	100,0%
		% within POL	229	171	400
Total		Count	57,3%	<b>42,8%</b>	100,0%
		% within POL	421	379	800
			52,6%	47,4%	100,0%

### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6,864	1	,009		
Continuity Correction	6,498	1	<b>,011</b>		
Likelihood Ratio	6,874	1	,009		
Fisher's Exact Test				,011	,005
Linear-by-Linear Association	6,855	1	,009		
N of Valid Cases	800				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 189,50.

Prva tabela je oblika 2 x 2 a sve frekvencije (192, 208, 229, 171) su veće od 5, pa se posmatra Asymp. Sig. (2-sided) u vrsti Continuity Correction. Kako je Asymp. Sig. (2-sided) = 0,011 < 0,05 odbacuje se hipoteza o nezavisnosti varijabli POL i DUVAN. Iz prve tabele se vidi da je vrednost 1 varijable DUVAN kod muškog pola zastupljena sa 52,0 % a kod ženskog sa 42,8 %.

## Linearna regresija

### Analyze – Regression – Linear

U prozor **Dependent** uneti zavisnu varijablu a u prozor **Independent(s)** uneti nezavisnu varijablu, odnosno nezavisne varijable. Kliknuti **O.K.** Ako želimo da dobijemo modele (kada ima više nezavisnih varijabli), kliknuti **Method** i **Backward**.

Program daje koeficijente regresione prave i testira ih.

Ako je Sig. u Tabeli Coefficients manje od 0,05 to znači da nezavisna varijabla, koja se nalazi u toj vrsti, značajno utiče na zavisnu varijablu.

### Primer.

#### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,410	<b>,168</b>	,161	11,0077

a Predictors: (Constant), HORM1

#### ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	2744,640	1	2744,640	22,651	<b>,000</b>
Residual	13570,886	112	121,169		
Total	16315,526	113			

a Predictors: (Constant), HORM1  
b Dependent Variable: FRAK1

#### Coefficients

Model	Unstandardized Coefficient B	Std. Error	Standardized Coefficient Beta	t	Sig.
1 (Constant)	<b>28,379</b>	1,421		19,969	<b>,000</b>
HORM1	<b>-,123</b>	,026	-,410	-4,759	<b>,000</b>

a Dependent Variable: FRAK1

Nezavisna varijabla je HORM1, a zavisna FRAK1. Jednačina regresione prave je

$$\text{FRAK1} = 28,379 - 0,123 \cdot \text{HORM1}.$$

Kako je Sig. = 0,000 < 0,05 varijabla HORM1 značajno utiče na varijablu FRAK1.

## Primer.

### Coefficients

Model		Unstandar dized Coefficient s	Std. Error	Standardiz ed Coefficient s	t	Sig.
1	(Constant)	75,102	8,464		8,874	,000
	ANTICIP	-,208	,141	-,161	-1,482	,141
	FEAR	-,219	,120	-,190	-1,833	,070
	SHYNESS	-,234	,115	-,193	-2,033	,044
	SENTIMEN	4,310E-02	,100	,039	,429	<b>,669</b>
2	(Constant)	76,085	8,117		9,373	,000
	ANTICIP	-,198	,138	-,152	-1,433	<b>,155</b>
	FEAR	-,212	,118	-,184	-1,796	,075
	SHYNESS	-,238	,114	-,196	-2,080	,040
3	(Constant)	<b>69,556</b>	6,749		10,306	<b>,000</b>
	FEAR	<b>-,292</b>	,105	-,253	-2,794	<b>,006</b>
	SHYNESS	<b>-,284</b>	,110	-,235	-2,586	<b>,011</b>

a Dependent Variable: EXSTRV

Iz Modela 1 se vidi da varijabla SENTIMEN najmanje utiče na varijablu EXSTRV (ima najveće Sig. = 0,669) pa je izostavljena u Modelu 2. Iz Modela 2 se vidi da varijabla ANTICIP najmanje utiče na varijablu EXSTRV (ima najveće Sig. = 0,155) pa je izostavljena u Modelu 3. Iz Modela 3 se vidi da promenljive FEAR (Sig. = 0,006) i SHYNESS (Sig. = 0,011) značajno utiču na zavisnu promenljivu. Jednačina regresione ravni je

$$\text{EXSTRV} = 69,556 - 0,292 \cdot \text{FEAR} - 0,284 \cdot \text{SHYNESS}.$$

## Binarna logistička regresija

### Analyze – Regression – Binary Logistic

U prozor **Dependent** uneti binarnu varijablu a u prozor **Covariate(s)** uneti nezavisnu varijablu, odnosno nezavisne varijable. Kliknuti **O.K.**

Za dobijanje modela kliknuti **Method** i **Backward:Wald**. Da bi se dobio interval poverenja za  $\exp(B)$  treba kliknuti na **Options** i štiklirati **CI for  $\exp(B)$** , zatim kliknuti **Continue** i **O.K.**

Nezavisne (prediktorske) varijable kojima odgovara broj Sig. manji od 0,05 značajno utiču na zavisnu, binarnu varijablu. Ako je Exp(B) veće od 1 to znači da ako se ta prediktorska varijabla poveća, rizik da binarna varijabla dobije vrednost 1 se povećava. Ako je Exp(B) manje od 1 to znači da ako se ta prediktorska varijabla poveća, rizik da binarna varijabla dobije vrednost 1 se smanjuje.

## Primer.

Ovde se radi o binarnoj varijabli INF kod koje broj 1 znači da je odgovarajuća osoba imala infarkt, a broj 0 da nije.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95,0% C.I. for EXP(B)	
								Lower	Upper
Step 1	EXSTRAV	,002	,012	,021	1	,885	1,002	,979	1,025
	DISORDER	-,032	,014	5,115	1	<b>,024</b>	<b>,968</b>	<b>,942</b>	<b>,996</b>
	HARM	-,042	,021	4,108	1	<b>,043</b>	<b>,959</b>	<b>,921</b>	<b>,999</b>
	ANTICIP	,022	,019	1,349	1	,245	1,022	,985	1,060
	FEAR	,005	,012	,172	1	,679	1,005	,981	1,030
	NOVELITY	,048	,019	6,640	1	<b>,010</b>	<b>1,049</b>	<b>1,011</b>	<b>1,088</b>
	SENTIMEN	-,010	,008	1,384	1	,239	,990	,974	1,007
	Constant	,348	,947	,135	1	,713	1,416		

a Variable(s) entered on step 1: EXSTRAV, DISORDER, HARM, ANTICIP, FEAR, NOVELITY, SENTIMEN.

Na binarnu varijablu INF (na pojavu infarkta) značajno utiču varijable DISORDER (Sig. = 0,024), HARM (Sig. = 0,043) i NOVELITY (Sig. = 0,010).

Odds ratio (Exp(B)) za varijablu DISORDER je 0,968, što znači da ako se varijabla DISORDER poveća za 1, rizik da binarna varijabla dobije vrednost 1 (da odgovarajuća osoba ima infarkt) se smanji za  $1 - 0,968 = 0,032$  t.j za 3,2 %.

Odds ratio (Exp(B)) za varijablu HARM je 0,959, što znači da ako se varijabla HARM poveća za 1, rizik da binarna varijabla dobije vrednost 1 (da odgovarajuća osoba ima infarkt) se smanji za  $1 - 0,959 = 0,041$  t.j za 4,1 %.

Odds ratio (Exp(B)) za varijablu NOVELITY je 1,049, što znači da ako se varijabla NOVELITY poveća za 1, rizik da binarna varijabla dobije vrednost 1 (da odgovarajuća osoba ima infarkt) se poveća za  $1,049 - 1 = 0,049$  t.j za 4,9 %.

## Linearna korelacija

### Analyze – Correlate – Bivariate

Dve ili više varijabli uneti u prozor **Variables** a zatim kliknuti **O.K.** Ako se štiklira **Spearman** onda se dobija rezultat Spearman-ovog testa.

Program daje Pearson-ov koeficijent korelacije (Pearson Correlation) i statističku značajnost Sig. (2-tailed). Pearson-ov koeficijent pokazuje jačinu veze između varijabli a Sig. (2-tailed) sa koliko poverenja treba posmatrati dobijene rezultate. Ako je Sig. (2-tailed) < 0,05 radi se o značajnoj korelaciji.

### Primer.

#### Correlations

		FRAK1	HORM1
FRAK1	Pearson Correlation	1,000	-,410**
	Sig. (2-tailed)	,	,000
	N	114	114
HORM1	Pearson Correlation	-,410	1,000
	Sig. (2-tailed)	,000	,
	N	114	122

\*\* Correlation is significant at the 0.01 level (2-tailed).

U ovom primeru Pearson-ov koeficijent korelacije je – 0,410 što govori da je korelacija između varijabli srednje jačine. Znak minus pokazuje da ako jedna varijabla raste druga opada. Pošto je Sig. (2-tailed) = 0,000 < 0,05 korelacija je značajna.

## ROC krive

### Graphs – ROC curve

Neprekidnu promenljivu uneti u prozor **Test Variable**, promenljivu koja definiše ishod (živ – mrtav, zdrav – bolestan) uneti u prozor **State Variable**, u prozor **Value of State Variable** uneti onu vrednost koja određuje ishod (na primer, 1) a zatim štiklirati **With diagonal reference line**, **Standard error and confidence interval** i **Coordinate points of the ROC curve**.



Primer. Ovde se radi o binarnoj varijabli INFARKT kod koje broj 1 znači da je odgovarajuća osoba imala infarkt, a broj 0 da nije. Numerička promenljiva je HOLESTER.

#### Area Under the Curve

Test Result Variable(s): HOLESTER

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
<b>,803</b>	,080	<b>,003</b>	,646	,961

The test result variable(s): HOLESTER has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a Under the nonparametric assumption

b Null hypothesis: true area = 0.5

Dobijeno je area = 0,803 a Sig. = 0,003 pa promenljiva Holester može da bude marker za binarnu promenljivu.

#### Coordinates of the Curve

Test Result Variable(s): HOLESTER

Positive if Greater Than or Equal To	Sensitivity	1 - Specificity
2,5000	1,000	1,000
3,6500	1,000	,938
3,9000	1,000	,813
4,1000	,941	,688
4,5000	,941	,625
4,8500	,941	,500
4,9500	,882	,438
5,3000	,882	,313
<b>6,3000</b>	<b>,882</b>	<b>,250</b>
7,4000	,824	,250
7,9000	,765	,250
8,2000	,353	,125
8,7000	,294	,125
9,1500	,118	,000
9,6500	,059	,000
11,0000	,000	,000

The test result variable(s): HOLESTER has at least one tie between the positive actual state group and the negative actual state group.

a The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

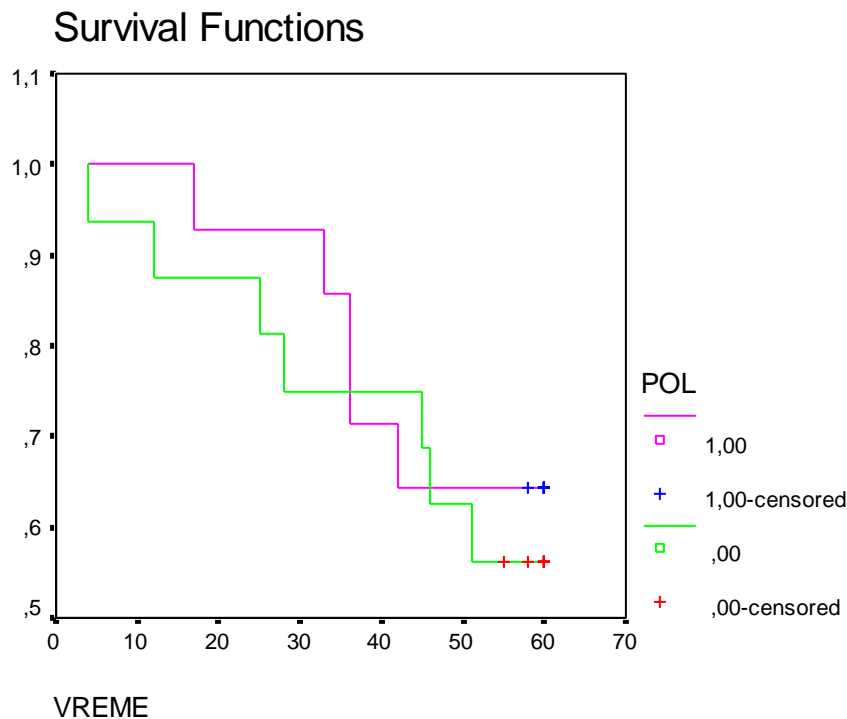
Proizvod senzitivnosti i specifičnosti je maksimalan ako je vrednost promenljive Holester jednaka 6,3 (senzitivnost je 0,882 a specifičnost  $1 - 0,250 = 0,750$ ).

## Kaplan-Meier-ove krive

### Analyze – Survival – Kaplan-Meier

U prozor **Time** uneti vreme, u prozor **Status** uneti promenljivu koja određuje ishod (živ – mrtav), kliknuti na **Define Event**, u prozor **Single Value** uneti vrednost koja određuje ishod (na primer, 1) i kliknuti **Continue**. Zatim kliknuti na **Options**, štiklirati **Survival** i kliknuti **Continue**. Promenljivu koja određuje grupe uneti u prozor **Factor**. Kliknuti na **Compare Factor**, štiklirati **Log rank**, kliknuti **Continue** a zatim **O.K.**

Primer. U ovom primeru se ispituje da li je razlika u preživljavanju (VREME, STATUS) između polova statistički značajna.



	Statistic	df	Significance
Log Rank	,19	1	,6600

Kako je Significance = 0,6600  $\geq$  0,05 razlika u preživljavanju između polova nije statistički značajna.

## Cox-ova regresija

### Analyze – Survival – Cox Regression

U prozor **Time** uneti vreme, u prozor **Status** uneti promenljivu koja određuje ishod (živ – mrtav), kliknuti na **Define Event**, u prozor **Single Value** uneti vrednost koja određuje ishod (na primer, 1) i kliknuti **Continue**. U prozor **Covariates** uneti promenljive čiji se uticaj na preživljavanje ispituje. Da bi se dobio interval poverenja za  $\exp(B)$  teba kliknuti na **Options** i štiklirati **CI for  $\exp(B)$** , zatim kliknuti **Continue** i **O.K.**

Za dobijanje modela kliknuti **Method** i **Backward:Wald** a zatim kliknuti **O.K.**

**Primer.** U ovom primeru se ispituje uticaj promenljivih GODINE, POL, SISTOL, HOLEST i PUSENJE na preživljavanje (VREME, STATUS).

#### Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95% CI for Exp(B)	
							Lower	Upper
GODINE	-,059	,033	3,266	1	,071	,943	,884	1,005
POL	2,338	1,052	4,942	1	<b>,026</b>	10,356	1,319	81,330
SISTOL	,284	,092	9,620	1	<b>,002</b>	1,328	1,110	1,590
HOLEST	-,015	,024	,414	1	,520	,985	,941	1,031
PUSENJE	,589	1,105	,285	1	,594	1,803	,207	15,718

Na preživljavanje utiču one promenljive za koje je  $\text{Sig.} < 0,05$  t.j. promenljive POL i SISTOLNI.