

Zadaci za vežbe: Ocenjivanje parametara

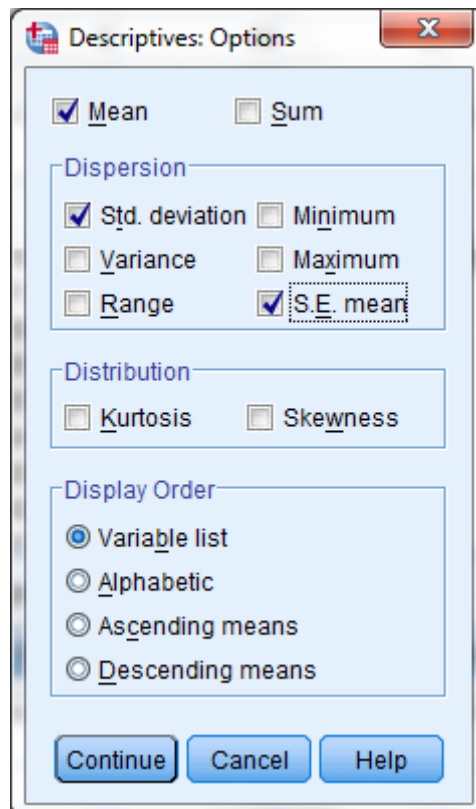
Rešenja zadataka uradila Jelica Miložić, student psihologije, PS160001

Zadatak 1.

U fajlu **gzptbu.sav** nalaze se, između ostalog, podaci o emocionalnoj uravnoteženosti (varijabla **EU**) i društvenosti (varijabla **DRUS**) slučajnog uzorka studenata.

- Izračunajte ocenu aritmetičke sredine i standardne devijacije na varijabli emocionalna uravnoteženost (**EU**) za populaciju iz koje je ovaj uzorak;
- Izračunajte standardnu grešku za aritmetičku sredinu na varijabli emocionalna uravnoteženost (**EU**);

REŠENJE: Računanje ocene aritmetičke sredine i standardne devijacije, kao i ocenjene standardne greške za aritmetičku sredinu na varijabli emocionalna uravnoteženost vrši se odabirom menija **Analyze-> Descriptive Statistics -> Descriptives**. U okvir **Variable(s)**: ubaci se varijabla emocionalna uravnoteženost i potom u **Options** uključe se opcije **Mean** (za računanje aritmetičke sredine), **Std. deviation** (za računanje standardne devijacije) i **S.E. mean** (za računanje standardne greške za aritmetičku sredinu). Definisane komande završava se klikom na **Continue** pa na **OK**.



U ispisu se dobija sledeća tabela:

	N	Mean		Std. Deviation
	Statistic	Statistic	Std. Error	Statistic
EMOCIONALNA URAVNOTEZENOST	960	14.71	.188	5.817
Valid N (listwise)	960			

Dakle, ocena aritmetičke sredine populacije iz koje je uzorak iznosi 14.71, ocena standardne devijacije je 5.82, a ocenjena standardna greška je 0.19.

Zadatak 2.

Fajl je isti kao za zadatak 1.

- Napravite 95% i 99% interval poverenja za aritmetičku sredinu varijable emocionalna uravnoteženost (**EU**)

REŠENJE: Računanje intervala poverenja za varijablu emocionalna uravnoteženost vrši se pomoću menija **Analyze-> Descriptive Statistics -> Explore**. U okvir **Dependent List** ubaci se varijabla emocionalna uravnoteženost, a posle klika na opciju **Statistics** u polju posle **Confidence Interval for Mean**: treba samo proveriti da li piše 95, što je inače automatski (tj. po difoltu) već uključeno. U ispisu se dobija sledeća tabela:

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
EMOCIONALNA URAVNOTEZENOST	960	100.0%	0	0.0%	960	100.0%

		Statistic	Std. Error
EMOCIONALNA URAVNOTEZENOST	Mean	14.71	.188
	95% Confidence Interval for Mean	Lower Bound	14.35
		Upper Bound	15.08
	5% Trimmed Mean	14.74	
	Median	15.00	
	Variance	33.833	
	Std. Deviation	5.817	
	Minimum	0	
	Maximum	29	
	Range	29	
	Interquartile Range	8	
Skewness	-.108	.079	
Kurtosis	-.543	.158	

Raspon 95% intervala poverenja u ovom slučaju ide od 14.35 (**Lower Bound** – donja granica) do 15.08 (**Upper Bound** – gornja granica).

Dakle, sa 95% sigurnosti (i 5% rizika greške) zaključujemo da je parametar obuhvaćen ovim intervalom. Stepenn sigurnosti u zaključak proističe iz statističke teorije koja stoji u osnovi ovog postupka: ako bismo na uzorcima iste veličine napravili na isti način beskonačan broj 95% intervala poverenja, onda bi 95% tako konstruisanih intervala obuhvatilo parametar, a 5% intervala ne bi obuhvatilo parametar.

Definisanje procedure za 99% interval poverenja je isto kao i za 95% interval poverenja, samo što se u opciji **Descriptives** u polju iza **Confidence Interval for Mean**: umesto 95, upiše 99. U ispisu u sledećoj tabeli vidimo da je donja granica intervala 14.23, a gornja 15.20.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
EMOCIONALNA URAVNOTEZENOST	960	100.0%	0	0.0%	960	100.0%

Descriptives

		Statistic	Std. Error
EMOCIONALNA URAVNOTEZENOST	Mean	14.71	.188
	99% Confidence Interval for Mean	Lower Bound Upper Bound	14.23 15.20
	5% Trimmed Mean	14.74	
	Median	15.00	
	Variance	33.833	
	Std. Deviation	5.817	
	Minimum	0	
	Maximum	29	
	Range	29	
	Interquartile Range	8	
	Skewness	-.108	.079
	Kurtosis	-.543	.158

Sa 99% sigurnosti (i 1% rizika greške) zaključujemo da je parametar obuhvaćen ovim intervalom.

Možemo uočiti da je 99% interval poverenja nešto širi od 95% intervala poverenja. To je „cena“ koju plaćamo za veću uverenost u zaključak da je parametar obuhvaćen intervalom.

Napomena: U realnim istraživanjima ne pravimo i 95% i 99% interval poverenja. Koji ćemo interval poverenja (95% ili 99% ili neki drugi) praviti u našem konkretnom istraživanju odlučujemo pri planiranju istraživanja, tj. pre izvođenja istraživanja.

Zadatak 3.

Fajl isti kao za zadatak 1.

- Izračunajte ocenu aritmetičke sredine i standardne devijacije za varijablu emocionalna uravnoteženost (**EU**) ali odvojeno za muški i ženski deo populacije iz koje je ovaj uzorak (na varijabli **pol** muškarci su označeni cifrom 1 a žene cifrom 2). Utvrdite u koju aritmetičku sredinu biste imali više poverenja kao ocenu parametra centralne tendencije: onu ocenu koju ste napravili za subpopulaciju muškaraca ili onu koju ste napravili za subpopulaciju žena?

REŠENJE: Za podelu uzorka na muški i ženski pol koristimo komandu **Data -> Split File**. U komandi **Split File** izaberemo ili **Compare groups** (ako želimo da u ispisu pri prikazivanju rezultata analiza bude jedna tabela podeljena na delove za muškarce i žene) ili **Organize output by groups** (ako želimo dve fizički razdvojene tabele za muškarce i žene). U prozorčić **Groups Based on:** ubacimo varijablu pol. Kao rezultat ove komande pojavljuje se upozorenje **Split by pol** u donjem desnom uglu prozora za podatke. Prema tome, dokle god se ova opcija ne deaktivira (izborom **Data -> Split File** pa uključivanjem **Analyze all cases, do not create groups**) sve statističke analize i grafici biće rađeni odvojeno za muškarce i žene. Mi smo uključili opciju **Compare groups**.

Statističku proceduru potom definišemo isto kao i u zadatku 1: **Analyze-> Descriptive Statistics -> Descriptives**. U okvir **Variable(s):** ubaci se varijabla emocionalna uravnoteženost (**EU**), a u **Options** se uključe opcije **Mean**, **Std. deviation** i **S.E. mean**.

U ispisu dobijamo sledeću tabelu:

Descriptive Statistics					
		N	Mean	Std. Deviation	
Pol		Statistic	Statistic	Std. Error	Statistic
MUSKARCI	EMOCIONALNA	480	15.84	.258	5.662
	URAVNOTEZENOST				
	Valid N (listwise)	480			
ZENE	EMOCIONALNA	480	13.59	.263	5.758
	URAVNOTEZENOST				
	Valid N (listwise)	480			

Ocenjena standardna greška za aritmetičku sredinu iznosi 0.258 za muškarce, a 0.263 za žene. Ocenjena standardna greška, pošto su uzorci iste veličine, u ovom slučaju zavisi od standardne devijacije koja je neznatno veća u uzorku žena.

Zaključak: više poverenja bismo imali u ocenu aritmetičke sredine za subpopulaciju muškaraca jer ova aritmetička sredina ima nešto manju standardnu grešku. Ipak, razlika u veličini ocenjenih standardnih grešaka je zaista mala, te ni razlika u stepenu poverenja u ove ocene ne bi bila velika.

*Napomena: Mogli bismo isti cilj postići (ali bez korišćenja komande **Split file**) komandom **Analyze-> Descriptive Statistics -> Explore**. U okvir **Dependent List**: ubacili bismo varijablu emocionalna uravnoteženost (**EU**) a u okvir **Factor list**: varijablu pol.*

Zadatak 4.

Bilo koji fajl sa podacima (konkretni podaci za ovaj zadatak nisu bitni) ili upisati bilo kakav podatak (npr. cifru 1) u praznom prozoru za podatke. Ovo je neophodno kako bi mogla da bude korišćena komanda COMPUTE i kako bi SPSS imao prostor za unosenje rezultata.

Korišćenjem komande **Transform/Compute** i funkcija **IDF.NORMAL(prob, mean, stddev)** za normalnu funkciju gustine, a **IDF.T(p, df)** za Studentovu t-funkciju gustine napravite sledeće:

- Ako slučajna varijabla Z ima standardizovanu normalnu raspodelu izračunajte vrednost z1 za koju važi sledeće: verovatnoća da slučajna varijabla Z uzme vrednost jednaku z1 ili veću od z1 jednaka je 0.025;

REŠENJE: Funkcija **IDF.NORMAL(prob, mean, stddev)** daje u opštem slučaju kvantil, tj. vrednost z1 za koju važi sledeće: verovatnoća da slučajna varijabla koja ima standardizovanu normalnu raspodelu uzme neku vrednost jednaku z1 ili manju od z1 jednaka je verovatnoći **prob**, tj. argumentu **prob** iz funkcije **IDF.NORMAL(prob, mean, stddev)**. Traži se, dakle, vrednost kvantila 0.975 (tj. 1-0.025) za standardizovanu normalnu raspodelu. U meniju **Transform -> Compute Variable**, u polju **Target Variable**: upiše se z1, a u okvir **Numeric expression**: unese se funkcija **IDF.NORMAL(prob, mean, stddev)** i definiše kao **IDF.NORMAL(1-0.025, 0, 1)**. Kao rezultat u "varijabli" z1 dobija se 1.96. Znači, verovatnoća da varijabla koja ima standardizovanu normalnu raspodelu na slučaj uzme neku vrednost jednaku ili veću od te tačke (1.96) iznosi 0.025.

- Ako slučajna varijabla t (t statistik) ima Studentovu ili T raspodelu sa 20 stepeni slobode izračunajte vrednost t1 za koju važi sledeće: verovatnoća da slučajna varijabla t uzme vrednost jednaku t1 ili veću od t1 jednaka je 0.025;

REŠENJE: Funkcija **IDF.T(p, df)** daje u opštem slučaju kvantil, tj. vrednost t1 za koju važi sledeće: verovatnoća da slučajna varijabla koja ima Studentovu raspodelu sa parametrom **df** uzme neku vrednost manju od t1 jednaka je verovatnoći **p**, tj. argumentu **p** iz funkcije **IDF.T(p, df)**. Traži se, dakle, vrednost kvantila 0.975 za Studentovu ili T raspodelu. Pomoću funkcije **IDF.T(1-0.025, 20)** u meniju **Compute** (naziv „variable“ t1) dobija se 2.09.

- Ako slučajna varijabla t (t statistik) ima Studentovu ili T raspodelu sa 50 stepeni slobode izračunajte vrednost t_2 za koju važi sledeće: verovatnoća da slučajna varijabla t uzme vrednost jednaku t_2 ili veću od t_2 jednaka je 0.025;

REŠENJA: Pomoću funkcije **IDF.T(1-0.025, 50)** u meniju **Compute** (naziv „varijable“ t_2) dobija se 2.01.

- Ako slučajna varijabla t (t statistik) ima Studentovu ili T raspodelu sa 100 stepeni slobode izračunajte vrednost t_3 za koju važi sledeće: verovatnoća da slučajna varijabla t uzme vrednost jednaku t_3 ili veću od t_3 jednaka je 0.025;

REŠENJA: Pomoću funkcije **IDF.T(1-0.025, 100)** u meniju **Compute** (naziv „varijable“ t_3) dobija se 1.98.

- Ako slučajna varijabla t (t statistik) ima Studentovu ili T raspodelu sa 1000 stepeni slobode izračunajte vrednost t_4 za koju važi sledeće: verovatnoća da slučajna varijabla t uzme vrednost jednaku t_4 ili veću od t_4 jednaka je 0.025;

REŠENJA: Pomoću funkcije **IDF.T(1-0.025, 1000)** u meniju **Compute** (naziv „varijable“ t_4) dobija se 1.96.

- Uporedite vrednost z_1 sa vrednostima t_1 , t_2 , t_3 i t_4 . Šta na osnovu toga možete da zaključite? Postoji li statistička teorema na osnovu koje se ovi rezultati mogu predvideti?

REŠENJE:

z_1	t_1	t_2	t_3	t_4
1.96	2.09	2.01	1.98	1.96

Kao što vidimo iz vrednosti koje su upisane u kolonama t_1 , t_2 , t_3 i t_4 , kvantil 0.975 za Studentovu raspodelu sa povećanjem broja stepeni slobode približava se vrednosti 1.96, tj. vrednosti kvantila 0.975 za standardizovanu normalnu raspodelu.

Ovi rezultati mogu se predvideti na osnovu sledeće statističke teoreme: **Sa povećanjem broja stepeni slobode Studentova raspodela teži standardizovanoj normalnoj.**

Napomena: Odlaskom na adresu <http://rpsychologist.com/d3/tdist/> možete posmatrati kako se u zavisnosti od broja stepeni slobode menja oblik t -funkcije gustine i kako se t -distribucija približava standardizovanoj normalnoj.