

# Deskriptivna statistika – Zadaci

## Osnove statistike

Aleksandar Tomašević

April 2020.

## Sadržaj

<b>1</b>	<b>Uvod</b>	<b>1</b>
<b>2</b>	<b>Primer</b>	<b>2</b>
2.1	Pitanje, modaliteti, merna skala . . . . .	2
2.2	Distribucija frekvencija . . . . .	2
2.3	Mere centralne tendencije . . . . .	2
2.4	Boksplot . . . . .	5
2.5	Mere oblika rasporeda . . . . .	5
2.6	Razlike između grupa . . . . .	5

## 1 Uvod

U ovom dokumentu možete pronaći primere zadataka, odnosno pitanja koja se odnose na interpretaciju rezultata deskriptivne statistike dobijenih u SPSS-u.

Neki od osnovnih elemenata koje ovakav opis **jedne varijable** sadrži su:

- Formulacija pitanja u upitniku, ponuđeni odgovori (modaliteti), merna skala
- Distribucija frekvencija: histogram i/ili tabela frekvencija
- Mere centralne tendencije: modus, medijana, aritmetička sredina + mera varijabilnosti: standardna devijacija
- Boksplot
- Mere oblika rasporeda: asimetrija i kurtosis
- Poređenje neke od prethodnih mera između dve grupe (npr. muškarci i žene)

Od vas se na testu može tražiti da:

- Interpretirate neku meru deskriptivne statistike (sa gornjeg spiska)
- Izaberete tačan opis od ponuđenih opcija (npr. varijabla je a) asimetrična u levo b) simetrična c) asimetrična u desno)
- Interpretirate histogram ili boksplot (npr. uočiti modus na histogramu, prokomentarisati a/simetriju na osnovu histograma itd.)

Da biste mogli da odgovorite na ovakva ili slična pitanja na testu, neophodno je da znate kako u SPSS-u uraditi pomenute analize i kako se rezultati pravilno interpretiraju. U

nastavku ovog dokumenta biće prikazani rezultati i interpretacija za jednu varijablu, a vi možete vežbati na drugim varijablama.

## 2 Primer

U ovom primeru biće analizirana varijabla **povpp** odnosno *Poverenje u političke partije*.

### 2.1 Pitanje, modaliteti, merna skala

Pitanje se odnosi na poverenje građana u političke partije. U pitanju je skala odnosa (*scale*) koja sadrži 11 modaliteta, odnosno ponuđenih odgovora na skali od 0 (Nemam uopšte poverenja) do 10 (Potpuno poverenje).

8	povpolic	Numeric	2	0	Poverenje u policiju	{0, Nemam ...	77, 88, 99	9	Right	Scale	Input
9	povpolit	Numeric	2	0	Poverenje u politicare	{0, Nemam ...	77, 88, 99	9	Right	Scale	Input
10	povpp	Numeric	2	0	Poverenje u političke partije	{0, Nemam ...	77, 88, 99	9	Right	Scale	Input
11	povep	Numeric	2	0	Poverenje u Evropski parlament						
12	povun	Numeric	2	0	Poverenje u UN						
13	glasali	Numeric	1	0	Da li ste glasali na prethodnim parlamentarnim izborima						
14	glasalipartija	Numeric	2	0	Za koju partiju ste glasali na prethodnim izborima?						
15	velikapartija	Numeric	8	2	Za koju od većih partija ste glasali						
16	protesti	Numeric	1	0	Da li ste učestvovali na protestima u poslednjih 12 meseci						
17	polinternet	Numeric	1	0	Da li ste delili politički sadržaj na internetu u poslednjih 12 meseci						
18	partija	Numeric	1	0	Da li vam je neka od političkih partija bliska?						
19	partijabliska	Numeric	2	0	Koja partija vam je bliska?						
20	kolikobliska	Numeric	1	0	Koliko ste bliski toj partiji?						
21	glaseu	Numeric	2	0	Da li biste glasali za uclanjenje Srbije u EU						
22	eu	Numeric	8	2	Podrska pristupanju EU						
23	idskala	Numeric	2	0	Odredjenje na skali levica-desnica						
24	zadzivot	Numeric	2	0	Koliko ste zadovoljni vasim zivotom u celini	{0, Ekstrem...	77, 88, 99	9	Right	Scale	Input
25	zadeko	Numeric	2	0	Koliko ste zadovoljni stanjem nacionalne ekonomije u celini	{0, Ekstrem...	77, 88, 99	8	Right	Scale	Input

Slika 1: Variable view - Podaci o varijabli povpp

### 2.2 Distribucija frekvencija

Analyze ⇒ Descriptive Statistics ⇒ Frequencies.

Distribuciju frekvencija možemo prikazati tabelarno, kao na slici ispod.

U tabeli vidimo da je ukupno 1981 ispitanik dao odgovor na ovo pitanje i da je najčešći odgovor “Nemam uopšte poverenja”. Takav odgovor je dao 721 ispitanik ili 36.4% od svih ispitanika koji su odgovorili na ovo pitanje (*valid percent*).

Grafički, ovu distribuciju frekvencija prikazujemo na histogramu. Možete koristiti opciju Graphs ⇒ Chart Builder ili meni Analyze ⇒ Descriptive Statistics ⇒ Explore da biste dobili histogram.

Na histogramu vidimo ponovo da je najveći broj ispitanika dao odgovor 0, a najmanje odgovora nalazimo kod vrednosti 6,8 i 9.

### 2.3 Mere centralne tendencije

Mere centralne tendencije možete dobiti na različite načine u meniju Analyze ⇒ Descriptive Statistics ⇒ Frequencies, Descriptives ili Explore.

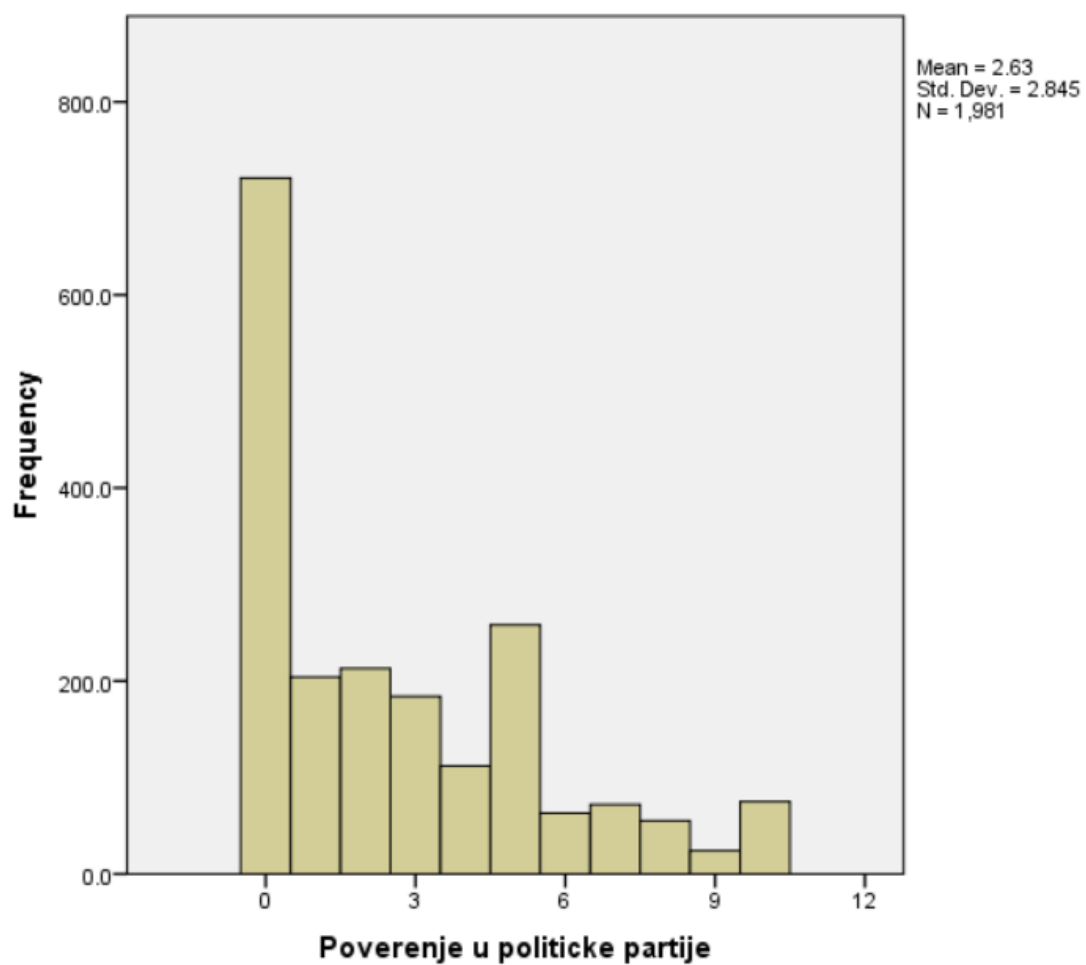
Na osnovu ovih vrednosti zaključujemo sledeće.

- Prosečan odgovor na postavljeno pitanje ima vrednost: 2.63

**Poverenje u politicke partije**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 Nemam uopste poverenja	721	35.2	36.4	36.4
	1 1	204	10.0	10.3	46.7
	2 2	213	10.4	10.8	57.4
	3 3	184	9.0	9.3	66.7
	4 4	112	5.5	5.7	72.4
	5 5	258	12.6	13.0	85.4
	6 6	63	3.1	3.2	88.6
	7 7	72	3.5	3.6	92.2
	8 8	55	2.7	2.8	95.0
	9 9	24	1.2	1.2	96.2
	10 Potpuno poverenje	75	3.7	3.8	100.0
	Total	1981	96.8	100.0	
Missing	77 Odbija	15	.7		
	88 Ne zna	50	2.4		
	Total	65	3.2		
Total		2046	100.0		

Slika 2: Tabela frekvencija



Slika 3: Histogram

Statistics		
Poverenje u politicke partije		
N	Valid	1981
	Missing	65
Mean		2.63
Median		2.00
Mode		0
Std. Deviation		2.845
Percentiles	25	.00
	50	2.00
	75	5.00

Slika 4: Deskriptivna statistika

- Najčešći odgovor, modus ili dominantna vrednost je vrednost 0
- 25% ispitanika koji najmanje veruju političkim partijama dali su odgovor 0
- 50% ispitanika je dalo odgovor manji ili jednak 2, odnosno odgovore 0, 1 ili 2
- 25% ispitanika koji najviše veruju političkim partijama dali su odgovor jednak ili veći od 5 (dakle: 5, 6, 7, 8, 9 ili 10)
- 50% središnjih ili tipičnih odgovora (između prvog i trećeg kvartila) kreću se između 0 i 5. Drugim rečima, odgovori veći od 5 su netipični, tj. nisu tipični građani koji imaju poverenja u političke partije.

Standardna devijacija iznosi 2.845.

Često zapisujemo zajedno vrednosti aritmetičke sredine (proseka) zajedno sa standardnom devijacijom.

Npr:  $2.63 \pm 2.845$  Prosek je 2.63 a tipični odgovori se nalaze u rasponu  $\pm 2.84$  u odnosu na prosek. Ako oduzmemo i saberemo ovaj broj od proseka dobićemo podataka da se tipične ili prosečne vrednosti nalaze u rasponu od 0 ( $2.63 - 2.845 = -0.2$  ali pošto je najmanja vrednost na skali 0, počinjemo od tog broja) do 5.47. Ovaj zaključak je nešto precizniji od onoga do kojeg smo došli analizom kvartila.

## 2.4 Boksplot

Boksplot možemo dobiti u istim menijima kao i histogram.

Na ovom boksplotu vidimo da je najveći deo vrednosti između 0 i 5, kao i da ne postoje ekstremne vrednosti unutar ovog skupa podataka.

## 2.5 Mere oblika rasporeda

Ove mere možemo dobiti u meniju Explore ili Descriptives.

Koeficijent asimetrije iznosi 0.942 što znači da je reč o blagoj formi asimetrije u desno. Na histogramu vidimo da je najveći deo vrednosti skocentrisan je levo od aritmetičke sredine (dominira vrednost 0) dok imamo manji broj netipičnijih vrednosti sa desne strane aritmetičke sredine (vrednosti 6, 7, 8, 9 i 10).

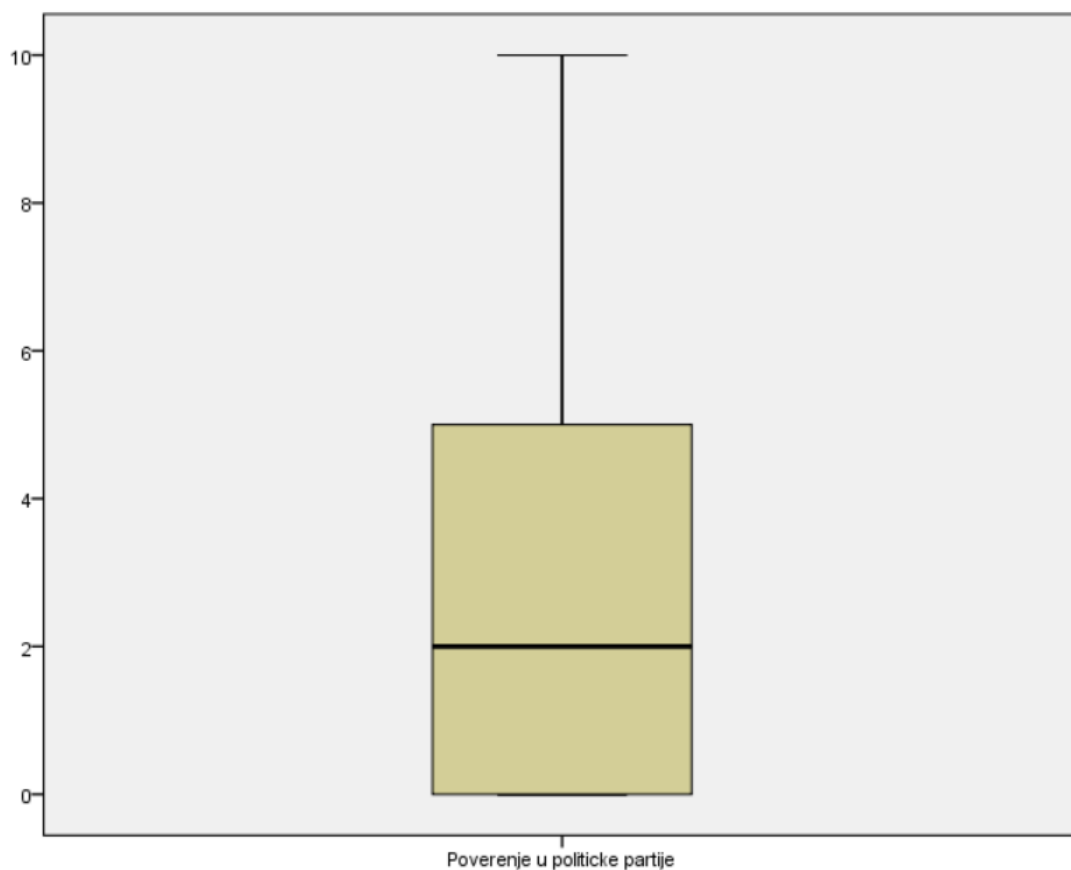
Koeficijent spšljoštenosti ili kurtosis iznosi -0.022 što znači da su ovi podaci približno normalno spljošteni.

## 2.6 Razlike između grupa

Poređićemo razlike između grupa u odnosu na varijablu **partija** odnosno *Da li vam je neka od političkih partija bliska?*. Na osnovu te varijable formiraju se dve grupe: 1) ispitanici koji su odgovorili “Da”, tj. oni kojima je neka partija bliska, 2) ispitanici koji su odgovorili “Ne”, odnosno oni kojima nijedna partija nije naročito bliska.

Razlike sumiramo kroz meni *Explore*.

Vidimo da postoje značajne razlike u prosečnim odgovorima. Prosek za grupu “Da” iznosi 3.96, dok za grupu “Ne” iznosi 1.9, što znači da oni kojima je neka partija bliska u proseku imaju veće poverenje u političke partije u odnosu na one kojima nijedna partija nije bliska.



Slika 5: Boksplo

Descriptives			Statistic	Std. Error
Poverenje u politicke partije	Mean		2.63	.064
	95% Confidence Interval for Mean	Lower Bound	2.50	
		Upper Bound	2.75	
	5% Trimmed Mean		2.38	
	Median		2.00	
	Variance		8.092	
	Std. Deviation		2.845	
	Minimum		0	
	Maximum		10	
	Range		10	
	Interquartile Range		5	
	Skewness		.942	.055
	Kurtosis		-.022	.110

Slika 6: Mere oblika rasporeda

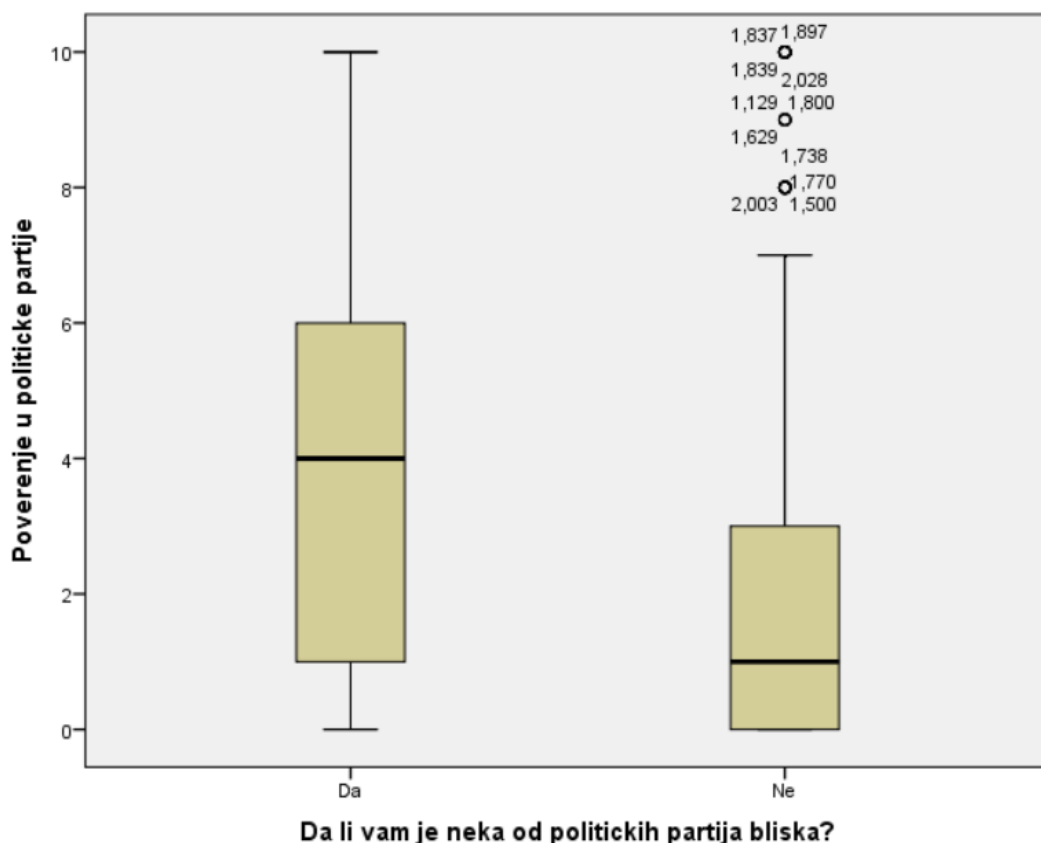
### Descriptives

Da li vam je neka od politickih partija bliska?				Statistic	Std. Error
Poverenje u politicke partije	1 Da	Mean		3.96	.120
		95% Confidence Interval for Mean	Lower Bound	3.73	
			Upper Bound	4.20	
		5% Trimmed Mean		3.85	
		Median		4.00	
		Variance		9.794	
		Std. Deviation		3.130	
		Minimum		0	
		Maximum		10	
		Range		10	
		Interquartile Range		5	
		Skewness		.355	.093
		Kurtosis		-.893	.187
	2 Ne	Mean		1.90	.069
		95% Confidence Interval for Mean	Lower Bound	1.76	
			Upper Bound	2.03	
		5% Trimmed Mean		1.64	
		Median		1.00	
		Variance		5.756	
		Std. Deviation		2.399	
		Minimum		0	
		Maximum		10	
		Range		10	
		Interquartile Range		3	
		Skewness		1.334	.070
		Kurtosis		1.224	.140

Slika 7: Razlike između grupa

Te razlike su izražene i u pogledu medijane. Medijalni odgovor prve grupe je 4, dok je medijalni odgovor druge grupe 1. To znači da je 50% pripadnika druge grupe na postavljeno pitanje dalo odgovor 0 ili 1.

Standardna devijacija je veća u prvoj grupi, što znači da je veća raznolikost ili varijabilitet odgovora u toj grupi, što se može videti na boksplotovima.



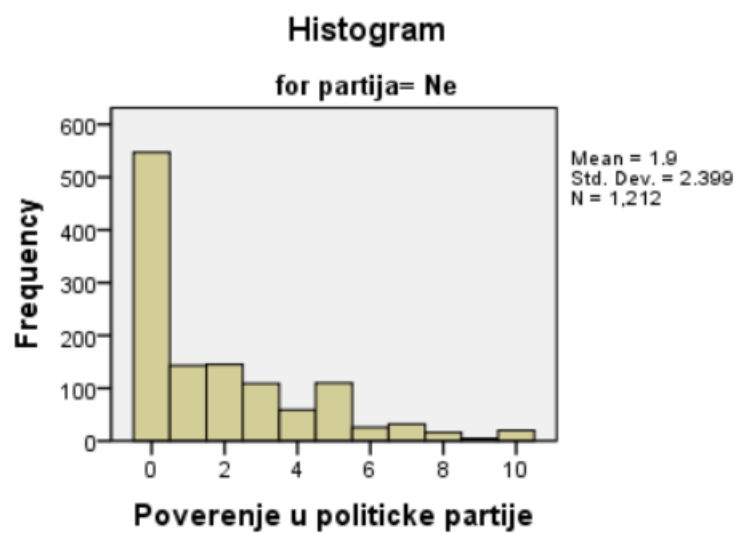
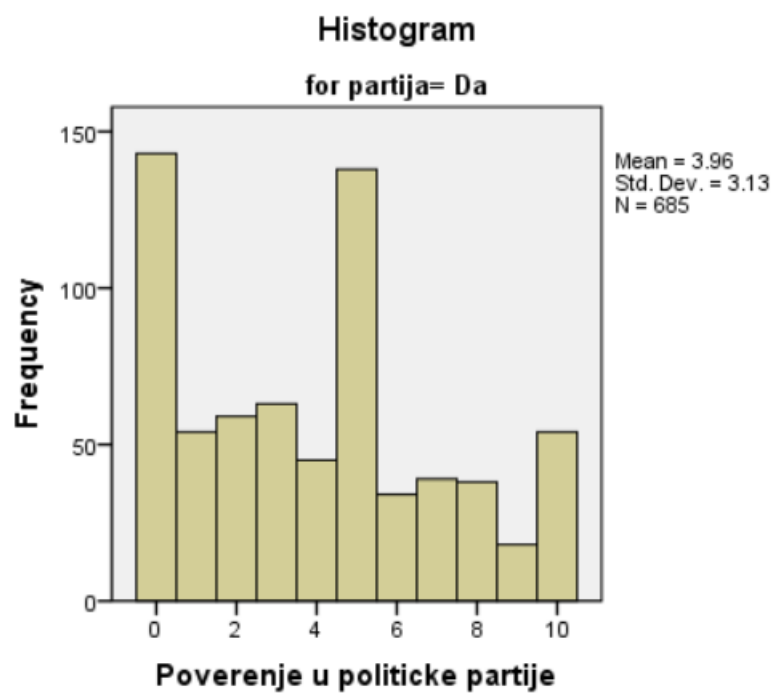
Slika 8: Boksplotovi za dve grupe

Prvi boksplot je izduženiji od drugog, što pokazuje da je varijabilitet veći u prvoj grupi. Takođe, primećujemo da u drugoj grupi postoje ekstremne vrednosti, a to su svi oni ispitanici koji su dali odgovore 8, 9 ili 10. Dakle, kod onih kojima nijedna partija nije bliska, ekstremno retki su ispitanici koji imaju visoko poverenje u političke partije.

Kada su u pitanju mere oblika rasporeda, prva grupa je blago asimetrična u desno, dok u drugoj grupi postoji izražena asimetrija u istu stranu (koeficijent asimetrije je 1.334). To se može uočiti i na histogramima, gde prvi histogram deluje relativno simetrično, dok je drugi jasno iskrivljen u desno.

Kada je kurtozis u pitanju, distribucija odgovora prve grupe je blago spljoštena, dok je u slučaju druge grupe izraženo izdužena (kurtozis je 1.334).





Slika 9: Histogrami dve grupe