

Testiranje hipoteza

Logika testiranja hipoteza

- ▶ Istraživači gotovo uvijek koriste uzorak da odgovore na pitanja o populaciji
- ▶ **Definicija:** *testiranje hipoteza je metod korišćenja uzorka da bi ocijenili teorijske pretpostavke o populaciji*
- ▶ **Statistika vs. Parametar**
- ▶ **Četiri koraka** za testiranje hipoteza:
 1. *Formulisanje hipoteze o populaciji*
 2. *Odrediti kriterijum za donošenje odluke*
 3. *Prikupljanje podataka i izračunavanje statistike uzorka*
 4. *Donošenje odluke*

Korak 1: Formulisanje hipoteze

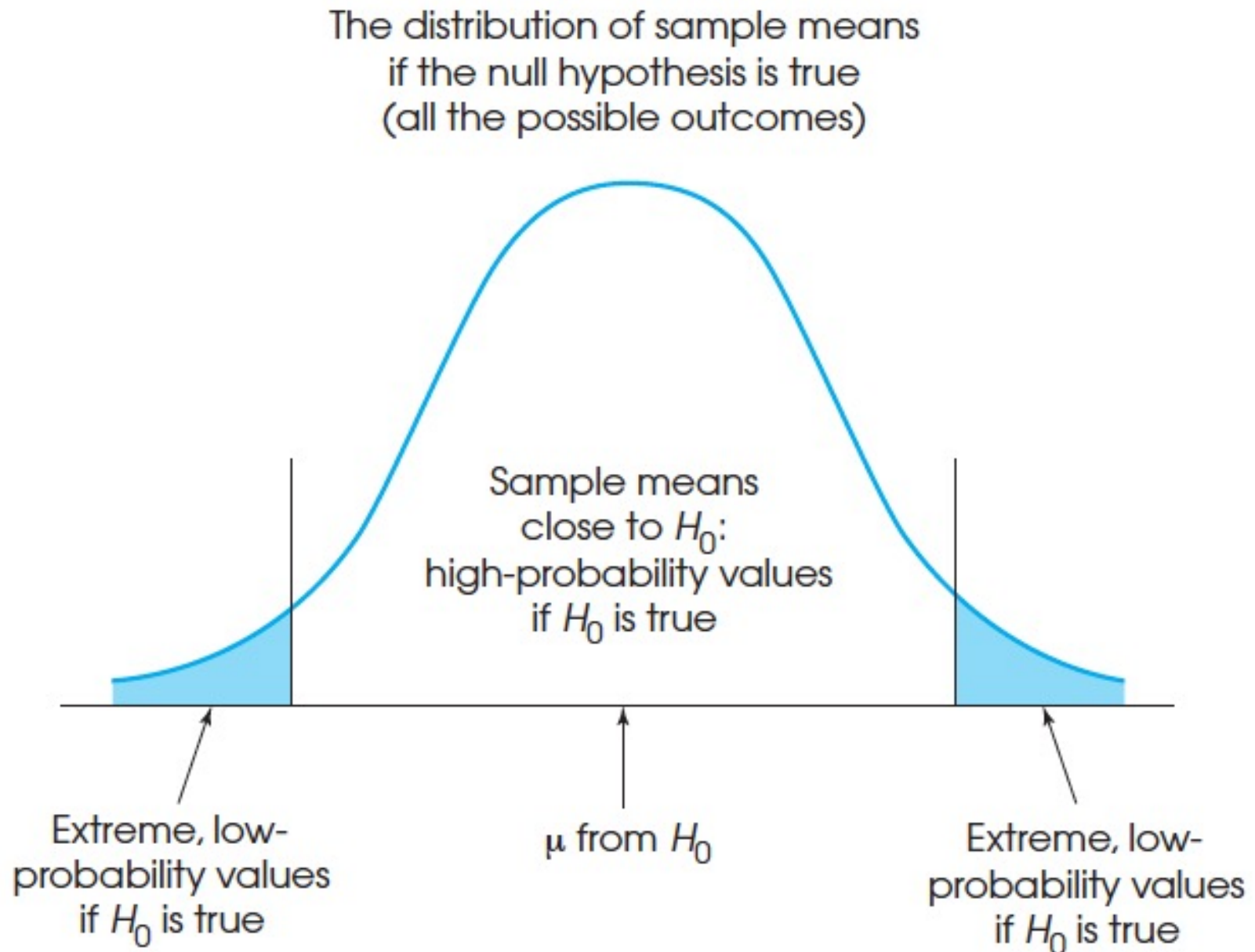
- ▶ Tiče se nepoznate karakteristike populacije (parametra)
- ▶ Formuliramo dvije hipoteze: **nultu** i **alternativnu**
- ▶ **Nulta hipoteza (H_0):** U populaciji nema promjena, razlika ili odnosa. Pretpostavlja da nezavisna varijabla (tretman) **nema** uticaj na zavisnu varijablu u populaciji.
- ▶ **Alternativna hipoteza (H_1):** U populaciji ima promjena, razlika, ili odnos između dva fenomena.
- ▶ **Primjer: (H_0)** – “Učestalo konzumiranje političkih medija nema efekta na nivo političko znanje
(H_1) - “Učestalo konzumiranje političkih medija utiče na nivo političkog znanja ”
- ▶ Zašto odbacujemo nultu a ne prihvatamo alternativnu?

Korak 2: Određivanje kriterijuma

- ▶ Podaci iz uzorka nam sugerišu da li imamo dovoljno dokaza da **opovrgnemo** nultu hipotezu ili ne
- ▶ Da bi donijeli tu odluku potrebna nam je **referentna tačka** – kritična vrijednost
- ▶ Ukoliko je vrijednost u uzorku **tipična** onda je vjerovatnije da nećemo opovrgnuti nultu hipotezu (i obrnuto).
- ▶ **Nivo statističke značajnosti**
 - Označava granice koje razdvajaju uzorke sa "malom" i "velikom" vjerovatnoćom
 - Arbitrarno određen nivo pouzdanosti
 - Osim statističke, zanima nas i substantivna značajnost

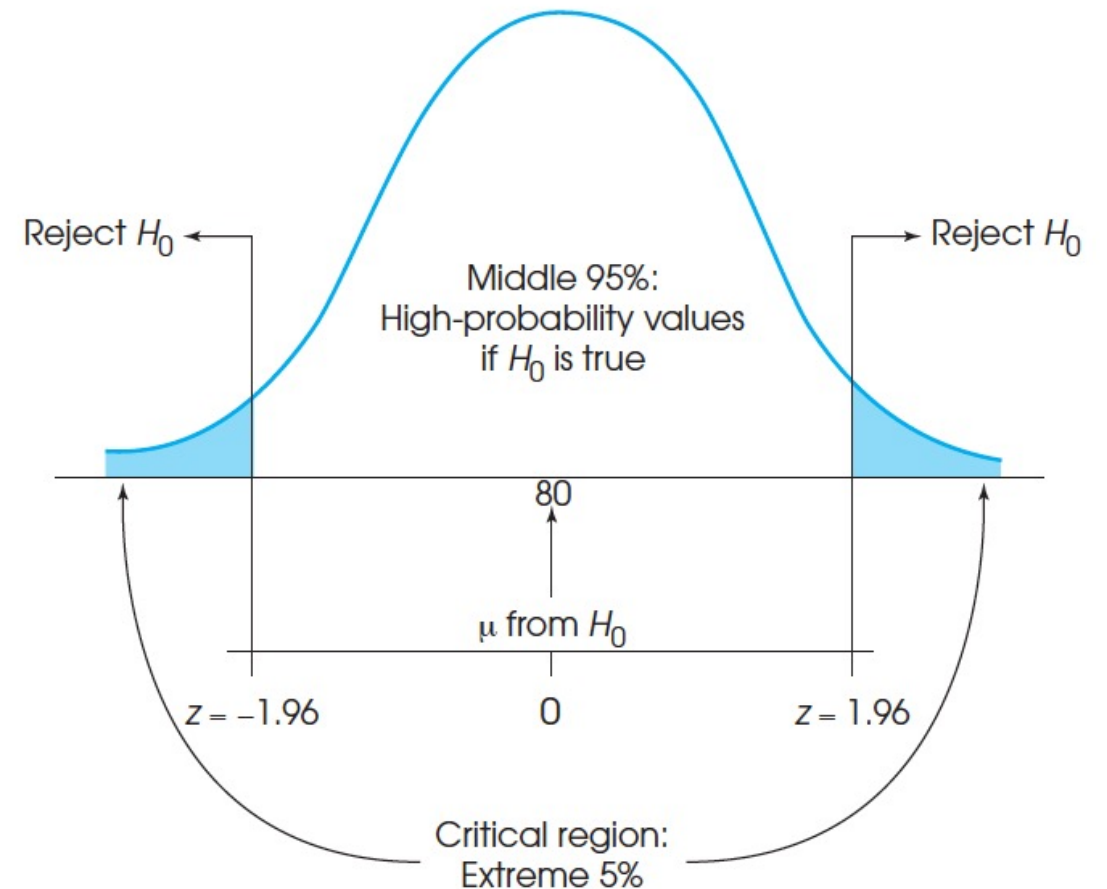
FIGURE 8.3

The set of potential samples is divided into those that are likely to be obtained and those that are very unlikely to be obtained if the null hypothesis is true.



Korak 2: Određivanje kriterijuma

- ▶ **Nivo alfa (α):** označava vjerovatnoću sa kojom obilježavamo “malo vjerovatne” uzoračke karakteristike
- ▶ **Kritične oblasti (regije):** sastoje se od ekstremnih vrijednosti uzoraka koje su potrebne da bi odbacili pretpostavku da je nulta hipoteza uvijek ispravna



Korak 3: Prikupljanje podataka i karakteristike uzorka

- ▶ Za računanje statistike uzorka potrebne su nam sljedeće informacije:
 - ❑ Aritmetička sredina uzorka
 - ❑ Pretpostavljena vrijednost aritmetičke sredine u populaciji
 - ❑ Veličina uzorka
 - ❑ Prosječno odstupanje pojedinačnih opservacija od prosjeka u uzorku

Korak 4: Donošenje odluke

- ▶ Donosimo odluku o tome da li imamo dovoljno dokaza da odbacimo nultu hipotezu. Postoje dvije mogućnosti:
 - a. Podaci iz uzorka se nalaze u kritičnom regionu**
 - a. Uzorak koji je malo vjerovatno dobiti ukoliko je nulta hipoteza tačna
 - b. Odlučujemo da odbacimo nultu hipotezu!
 - c. Interpretacija uključuje i nivo pouzdanosti

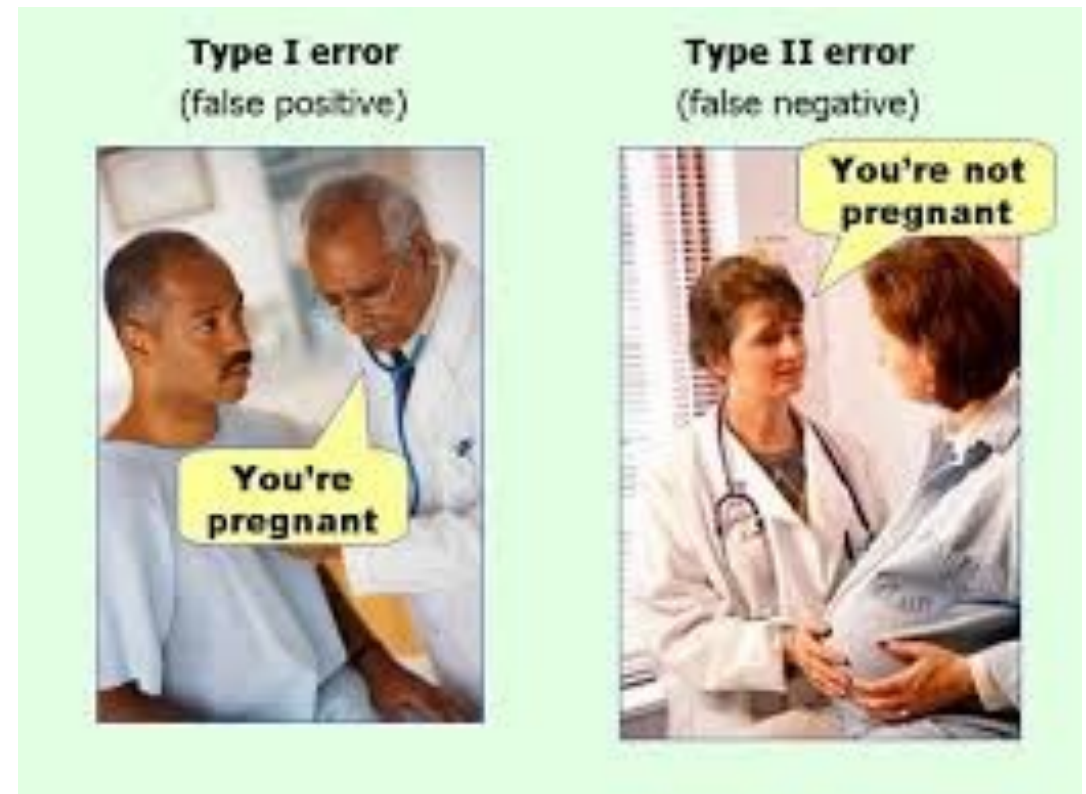
 - b. Podaci iz uzorka nijesu u kritičnom regionu**
 - a. Uzorak je prilično blizu pretpostavljenoj vrijednosti u nultoj hipotezi
 - b. Odlučujemo da nemamo dovoljno dokaz da odbacimo nultu hipotezu

Nesigurnost i greške

- ▶ Generalizacije **uvijek** dolaze uz određenu nesigurnost (uzorci nijesu savršeni)
- ▶ Što je vrijednost u uzorku **ekstremnija**, to je vjerovatnoća greške manja
- ▶ Vrste grešaka: **Greška I tipa** i **Greška II tipa**
- ▶ **Greška I tipa:** se dešava kada istraživač odbaci nultu hipotezu koja je tačna. Istraživač zaključuje da postoji efekat, kad u suštini taj efekat ne postoji.
 - ▶ Vjerovatnoća da smo napravili grešku prvog tipa označava se sa “**p**”
 - ▶ Najčešće pominjana vrijednost “p” je $< 0,05$ (5%)

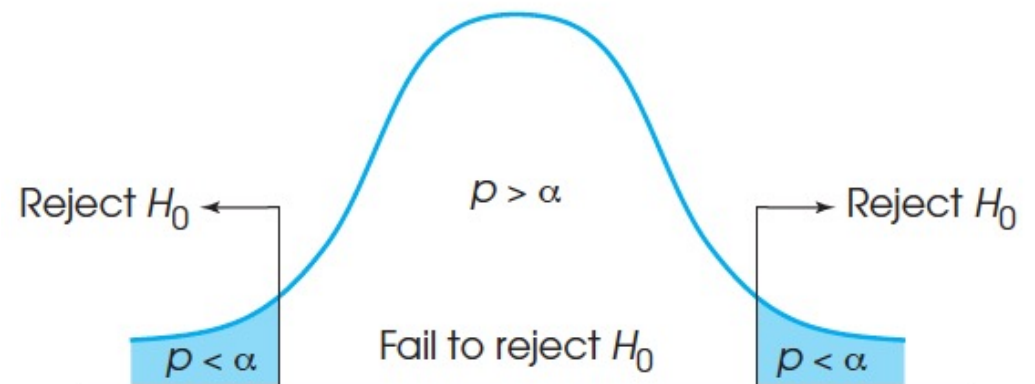
Nesigurnost i greške

- ▶ **Greška II tipa:** se dešava kada istraživač ne odbaci netačnu nultu hipotezu. Suštinski, nijesmo uspjeli otkriti postojeći efekat
- ▶ Manje je “opasna” od Greške I tipa
- ▶ Znači da nemamo dovoljno dokaza za potvrdimo očeivanja istraživača
- ▶ Ne može se izraziti jednim brojem kao u slučaju Greške I tipa



Odabir nivoa značajnosti

- ▶ U potpunosti **arbitrarno**, ali postoje **konvencije** (najčešće 5%)
- ▶ **Razlikuje** se zavisno od naučnog polja (teme)
- ▶ Određuje se **prije** nego je istraživanje sprovedeno
- ▶ **Inflacija** niskih p-vrijednosti u časopisima
- ▶ **Efekat “fioke”** u društvenim naukama
- ▶ Pretpostavke?



Bivarijantna statistika

Odnos između dvije varijable

► Suštinski nas zanimaju tri informacije:

1. STATISTIČKA ZNAČAJNOST:

- Da li se vrijednost koja sumira odnos između dva fenomena nalazi u kritičnoj regiji
- Preciznije, sa kojom vjerovatnoćom možemo reći da smo izbjegli **Grešku I tipa**

2. SNAGA ODNOSA:

1. Nevažno ukoliko veza nije statistički značajna!
2. Ukoliko jeste, zanima nas supstantivna značajnost
3. Nema jasnog standarda za kategorizaciju snage odnosa

3. SMJER ODNOSA: pozitivna ili negaativan

Krostabulacije

- ▶ **Krostabulacije vs. Korelacija**

- ▶ U slučaju da imamo **kategoričke** podatke (*nominalne* i *ordinalne*) najčešće koristimo krostabulacije.

- ▶ Tehnika ustaljena kad je u pitanju analiza **javnog mnjenja** ili **ponašanja birača**

- ▶ **Tipična pitanja:**

Da li postoji veza između rase (nominalna) i podrške za smrtnu kaznu (ordinalna)?

Jesu li rod i partijska pripadnost povezani?

Da li je nivo obrazovanje osobe povezan sa podrškom za zaštitu životne sredine?

Da li su članove sindikata vjerovatniji glasači lijeve partije X od onih koji nijesu članovi sindikata?

Krostabulacije

- ▶ **Cilj:** Kreirati tabelu sa vrijednostima jedne varijable kao redovima i vrijednostima druge varijable kao kolonama
- ▶ Najčešće se nezavisna varijabla nalazi u **kolonama**, a zavisna u **redovima**
- ▶ Svaka “**ćelija**” u tabeli sadrži broj pojedinačnih opservacija koji se nalaze na **presjeku** dvije datog reda i kolone
- ▶ Interpretacija se najčešće odnosi na **procente** u kolonama (ili redovima)

Tabela kontingencije 5:
Glasanje za partije, prema starosnoj strukturi

		DPS	SDP	SNP	NOVA	PZP	Apstinenti	Ostali	Ukupno
18-34	N	151	16	56	18	28	106	40	415
	%	36.4%	3.9%	13.5%	4.3%	6.7%	25.5%	9.6%	100.0%
35-55	N	179	18	81	24	20	79	41	442
	%	40.5%	4.1%	18.3%	5.4%	4.5%	17.9%	9.3%	100.0%
55+	N	141	19	74	9	28	66	33	370
	%	38.1%	5.1%	20.0%	2.4%	7.6%	17.8%	8.9%	100.0%
Ukupno	N	471	53	211	51	76	251	114	1227
	%	38.4%	4.3%	17.2%	4.2%	6.2%	20.5%	9.3%	100.0%

Napomena: $X^2(12, N = 1227) = 22.881; p < .05$

Krostabulacija

- ▶ Statistički test za dvije kategoričke varijable zove se **Pirson Hi-kvadrat test**
- ▶ Vrijednost Hi-kvadrata je manje koristan i važan od **statističke značajnosti**
 - ▶ P-vrijednost nam sugerše da li možemo očekivati vezu između dvije varijable i u populaciji
 - ▶ Prihvaćeno je da očekujemo da $p < 0.05$ (5%)

Tabela 1. Krostabulacija rase listavova prema smrtnoj kazni

	<i>Bijelci</i>	<i>Crnci</i>
<i>U potpunosti se slažem (%)</i>	57.7	33.5
<i>Slažem se (%)</i>	18.3	18.3
<i>Protivim se (%)</i>	10.5	12.6
<i>U potpunosti se protivim (%)</i>	13.6	35.6
<i>N</i>	1320	191

Izvor: ANES 2000 . Pirson Hi-kvadrat : 67.559; p-vrijednost = 0.000



Primjeri krostabulacija

Tabela kontingencije 8: *Porodična istorija u odnosu na partijsko opredjeljenje u CG*¹³³

		Partizane	Četnike	Zelenaše	Nikoga nisu podržali	Ukupno
DPS	N	205	19	9	81	314
	%	65.3%	6.1%	2.9%	25.8%	100.0%
SDP	N	24	3	1	10	38
	%	63.2%	7.9%	2.6%	26.3%	100.0%
SNP	N	97	42	0	25	164
	%	59.1%	25.6%	.0%	15.2%	100.0%
PZP	N	24	5	0	12	41
	%	58.5%	12.2%	.0%	29.3%	100.0%
NOVA	N	27	27	0	10	64
	%	42.2%	42.2%	.0%	15.6%	100.0%
Apstinenti	N	74	20	4	61	159
	%	46.5%	12.6%	2.5%	38.4%	100.0%
Ostali	N	40	14	2	23	79
	%	50.6%	17.7%	2.5%	29.1%	100.0%
Ukupno	N	491	130	16	222	859
	%	57.2%	15.1%	1.9%	25.8%	100.0%

Napomena: $X^2(18, N = 1223) = 100.037; p < .01$



Tabela kontingencije 1:
Glasanje za partije, prema nacionalnosti

		DPS	SDP	SNP	NOVA	PZP	Apstinenti	Ostali	Ukupno
Crnogorska	N	331	29	24	27	12	112	24	559
	%	59.2%	5.2%	4.3%	4.8%	2.1%	20.0%	4.3%	100.0%
Srpska	N	22	6	185	23	63	96	35	430
	%	5.1%	1.4%	43.0%	5.3%	14.7%	22.3%	8.1%	100.0%
Albanska	N	29	1	1	0	0	13	22	66
	%	43.9%	1.5%	1.5%	.0%	.0%	19.7%	33.3%	100.0%
Bošnjačka	N	22	4	0	0	0	5	12	43
	%	51.2%	9.3%	.0%	.0%	.0%	11.6%	27.9%	100.0%
Muslimanska	N	61	11	0	1	0	16	10	99
	%	61.6%	11.1%	.0%	1.0%	.0%	16.2%	10.1%	100.0%
Hrvatska	N	2	1	0	0	0	2	8	13
	%	15.4%	7.7%	.0%	.0%	.0%	15.4%	61.5%	100.0%
Drugo	N	3	1	1	0	1	4	3	13
	%	23.1%	7.7%	7.7%	.0%	7.7%	30.8%	23.1%	100.0%
Ukupno	N	470	53	211	51	76	248	114	1223
	%	38.4%	4.3%	17.3%	4.2%	6.2%	20.3%	9.3%	100.0%

Napomena: $X^2(36, N = 1223) = 692.383; p < .01$



Tabela kontingencije 2:
Glasanje za partije, prema konfesionalnoj pripadnosti

		DPS	SDP	SNP	NOVA	PZP	Apstinenti	Ostali	Ukupno
Srpska pravoslavna crkva	N	137	10	200	39	76	139	40	641
	%	21.4%	1.6%	31.2%	6.1%	11.9%	21.7%	6.2%	100.0%
Crnogorska pravoslavna crkva	N	173	19	6	7	1	35	7	248
	%	69.8%	7.7%	2.4%	2.8%	.4%	14.1%	2.8%	100.0%
Islamska	N	115	17	1	2	0	36	38	209
	%	55.0%	8.1%	.5%	1.0%	.0%	17.2%	18.2%	100.0%
Katolička	N	25	4	0	1	0	12	21	63
	%	39.7%	6.3%	.0%	1.6%	.0%	19.0%	33.3%	100.0%
Neka druga	N	1	1	1	1	0	2	1	7
	%	14.3%	14.3%	14.3%	14.3%	.0%	28.6%	14.3%	100.0%
Bez vjerske pripadnosti	N	18	2	1	2	0	25	3	51
	%	35.3%	3.9%	2.0%	3.9%	.0%	49.0%	5.9%	100.0%
Ukupno	N	469	53	209	52	77	249	110	1219
	%	38.5%	4.3%	17.1%	4.3%	6.3%	20.4%	9.0%	100.0%

Napomena: $X^2(30, N = 1219) = 496.093; p < .01$



Tabela kontingencije 4:
Zainteresovanost za politiku, prema polu

		Veoma zainteresovan/a	Donekle zainteresovan/a	Ne baš zainteresovan/a	Nisam uopšte zainteresovan/a	Ukupno
Muški	N	102	315	158	92	667
	%	15.3%	47.2%	23.7%	13.8%	100.0%
Ženski	N	43	182	244	248	717
	%	6.0%	25.4%	34.0%	34.6%	100.0%

Napomena: $X^2(3, N = 1384) = 147.960; p < .01$

Tabela 2.1 Kosmopolitizam po zemljama – dummy varijabla

	ostali	kosmopolitizam
BiH	85,3%	14,7%
Hrvatska	81,1%	18,9%
Crna Gora	73,5%	26,5%
Srbija	75,4%	24,6%
Slovenija	80,2%	19,8%
Makedonija	77,2%	22,8%
total	78,8%	21,2%

Table 1. Age Distribution of Supporters (Percentage)

	SDP	SDA	HDZ	SNSD
18-30	63.33	31.82	66.67	57.14
31-46	23.33	27.77	22.22	25.71
47-62	6.67	36.36	5.56	17.14
> 62	6.67	4.55	5.56	0.00

Source: European Values Study, B&H

Table 4. Ideological Identification among Party Electorates

Party	Left	Center	Right
SDP	32.89%	42.11%	25.00%
SDA	23.08%	30.77%	46.15%
HDZ	5.66%	20.75%	73.58%
SNSD	17.86%	32.14%	50.00%

Table 3
Twitter Style by Party

	DEMOCRAT	REPUBLICAN	THIRD PARTY
Attack ^c	6.4%	5.6%	5.7%
Attack Other ^{abc}	3.5%	5.6%	7.5%
Media ^{bc}	11.3%	12.6%	14.3%
Campaign	25.2%	25.3%	17.7%
Issues ^{bc}	9.8%	11.1%	13.6%
Mobilization	7.8%	7.2%	5.6%
Personal ^{bc}	29.6%	27.3%	31.1%
User Interaction ^{bc}	15.8%	12.2%	20.1%
Obama ^{abc}	2.8%	5.5%	6.4%
Romney ^{bc}	2.4%	2.9%	4.3%

a= Significant difference between Democrats and Republicans, $p \leq .10$.

b= Significant difference between Third Party and Democrats, $p \leq .10$.

c= Significant difference between Third Party and Republicans, $p \leq .10$.

Table 4
Twitter Style by Gender

	WOMEN	MEN
Attack+	6.5%	4.4%
Attack Other	5.3%	4.8%
Media	11.5%	12.5%
Campaign*	24.6%	24.0%
Issues**	12.8%	10.2%
Mobilization**	9.1%	6.5%
Personal	26.5%	29.7%
User Interaction	15.2%	14.9%
Obama	4.2%	4.5%
Romney	2.4%	3.1%

+ $p \leq .10$, * $p \leq .05$, ** $p \leq .01$.

Table 5
Twitter Style by Incumbency

	INCUMBENT	CHALLENGER
Attack**	1.3%	8.0%
Attack Other	5.8%	4.6%
Media*	13.5%	11.7%
Campaign**	19.7%	26.2%
Issues	14.9%	9.7%
Mobilization**	5.3%	8.1%
Personal*	35.3%	26.0%
User Interaction**	10%	16.8%
Obama	5.4%	4.0%
Romney*	3.1%	2.8%

** $p \leq .01$; * $p \leq .05$.

Table 6
Twitter Style by Competitiveness

	COMPETITIVE	NON-COMPETITIVE
Attack *	9%	5.4%
Attack Other+	2.2%	5.4%
Media	8.7%	12.9%
Campaign	20.0%	24.8%
Issues	11.1%	10.8%
Mobilization*	9.8%	6.7%
Personal	28.4%	28.8%
User Interaction	13.2%	15.2%
Obama+	1.6%	5.0%
Romney*	0.7%	3.2%

+ $p \leq .10$, * $p \leq .05$, ** $p \leq .01$.



Krostabulacije u formi grafika



Online analiza

U ovom dijelu sajta, moguće je lako i bez poznavanja statistike i rada u statističkim programima istraživati MNES bazu, kreirati grafike i upoređivati podatke u odnosu na demografiju.

Prilikom ukrštanja, uvijek se pojave razlike, manje ili veće, u odgovorima između ukrštenih kategorija. Ukoliko je ukrštanje statistički značajno, tj. ukoliko u populaciji **postoje statistički značajne razlike** između kategorija koje se upoređuju, na nivou od $p < 0.05$, na grafiku će se pojaviti informacija o tome. Ukoliko ukrštanje nije statistički značajno, ne možemo tvrditi da takve razlike postoje u populaciji o čemu će, takođe, biti informacija u samom grafiku.

Ukrštanja (krostabulacije)

Izaberite pitanje

-- izaberi pitanje --



Demografija

-- izaberi pitanje --



 Prikaži grafik

Da li stojite dok svira državna himna? / Pol ispitanika

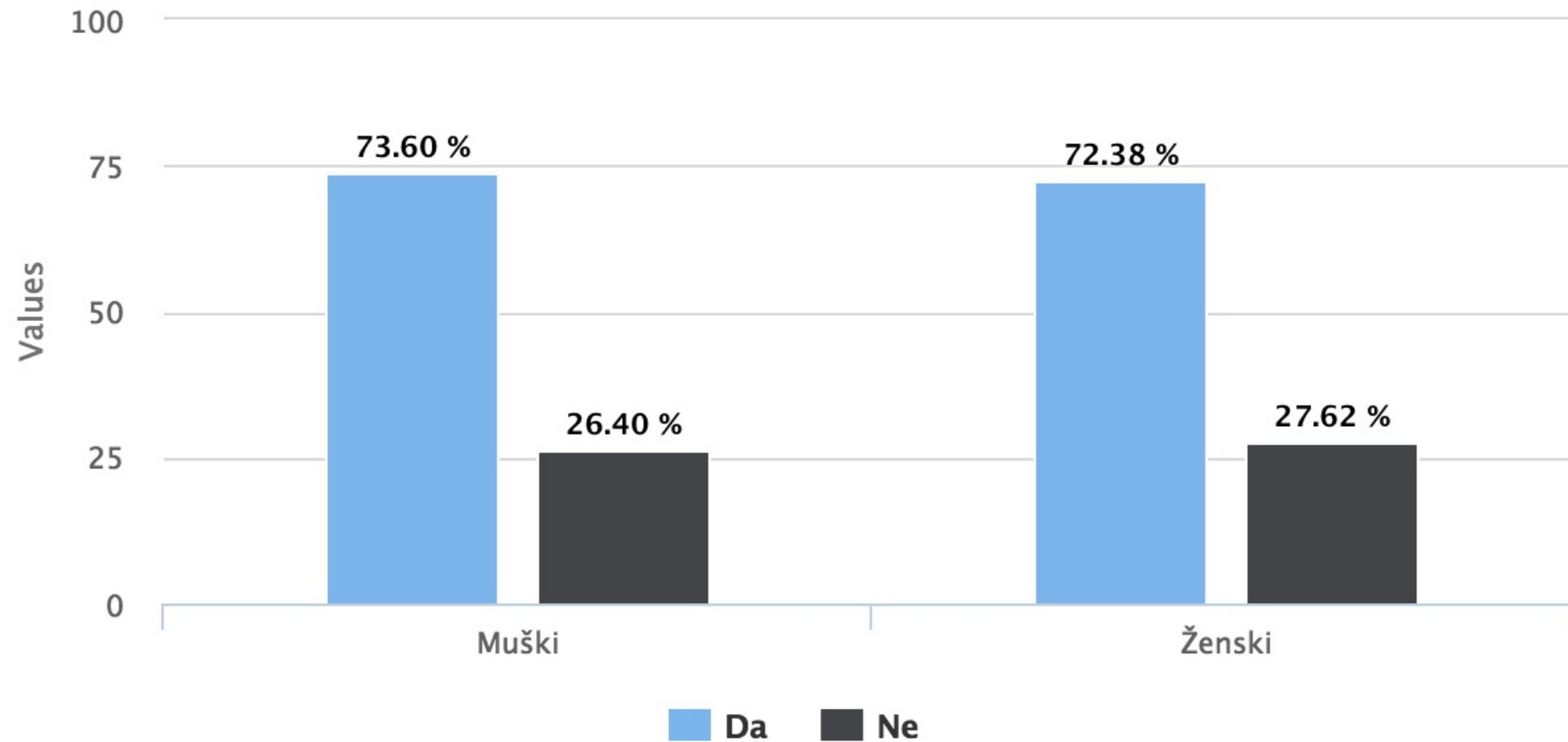


N (broj ispitanika) = 1144

$\chi^2(1) = 0.2123, p < 0.05$

$\chi^2(1) < 3.84$

Ukrštanje nije statistički značajno



Povjerenje u institucije – Političke partije / Pol ispitanika

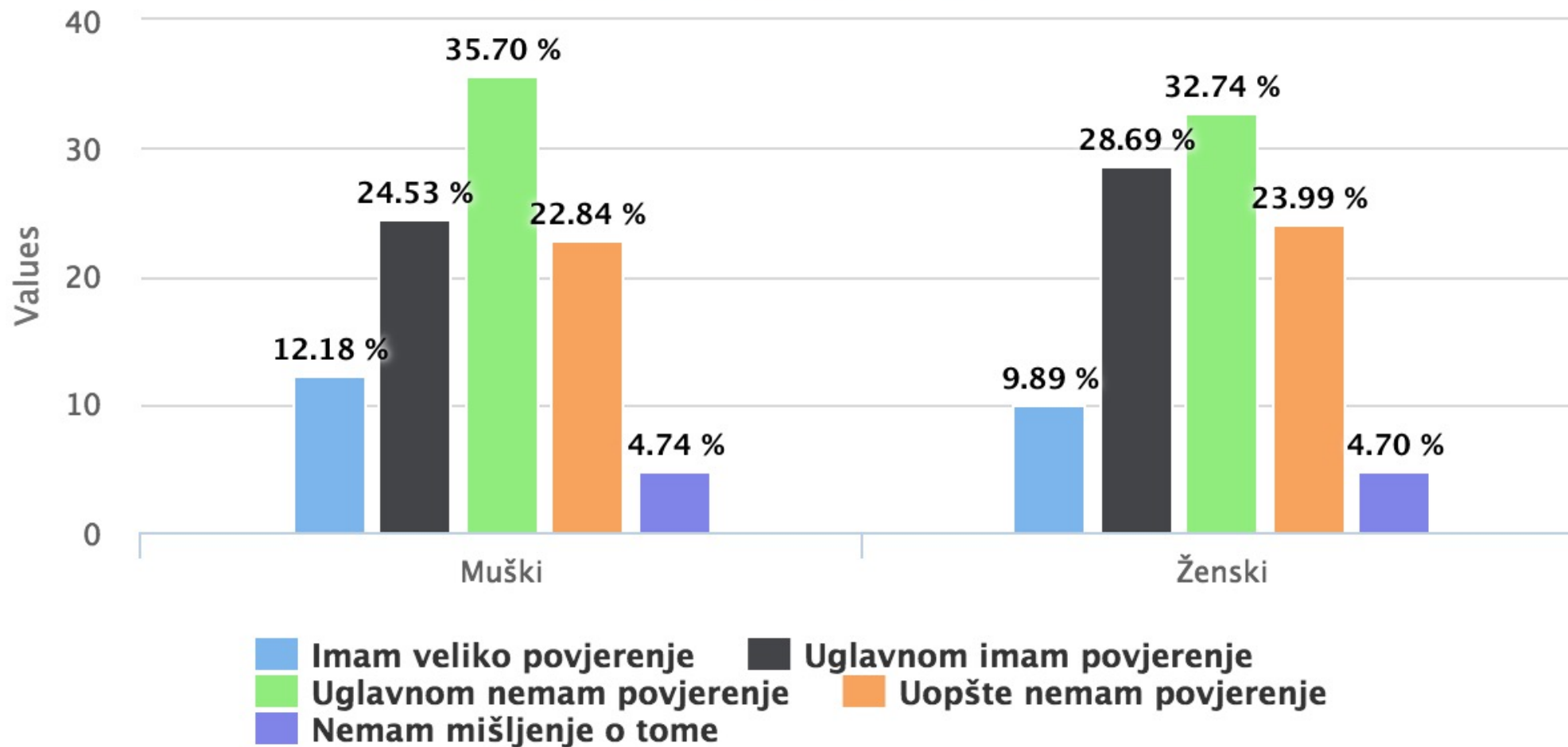


N (broj ispitanika) = 1209

$\chi^2(4) = 4.3472, p < 0.05$

$\chi^2(4) < 9.49$

Ukrštanje nije statistički značajno



Povjerenje u institucije – Vojska CG / Pol ispitanika

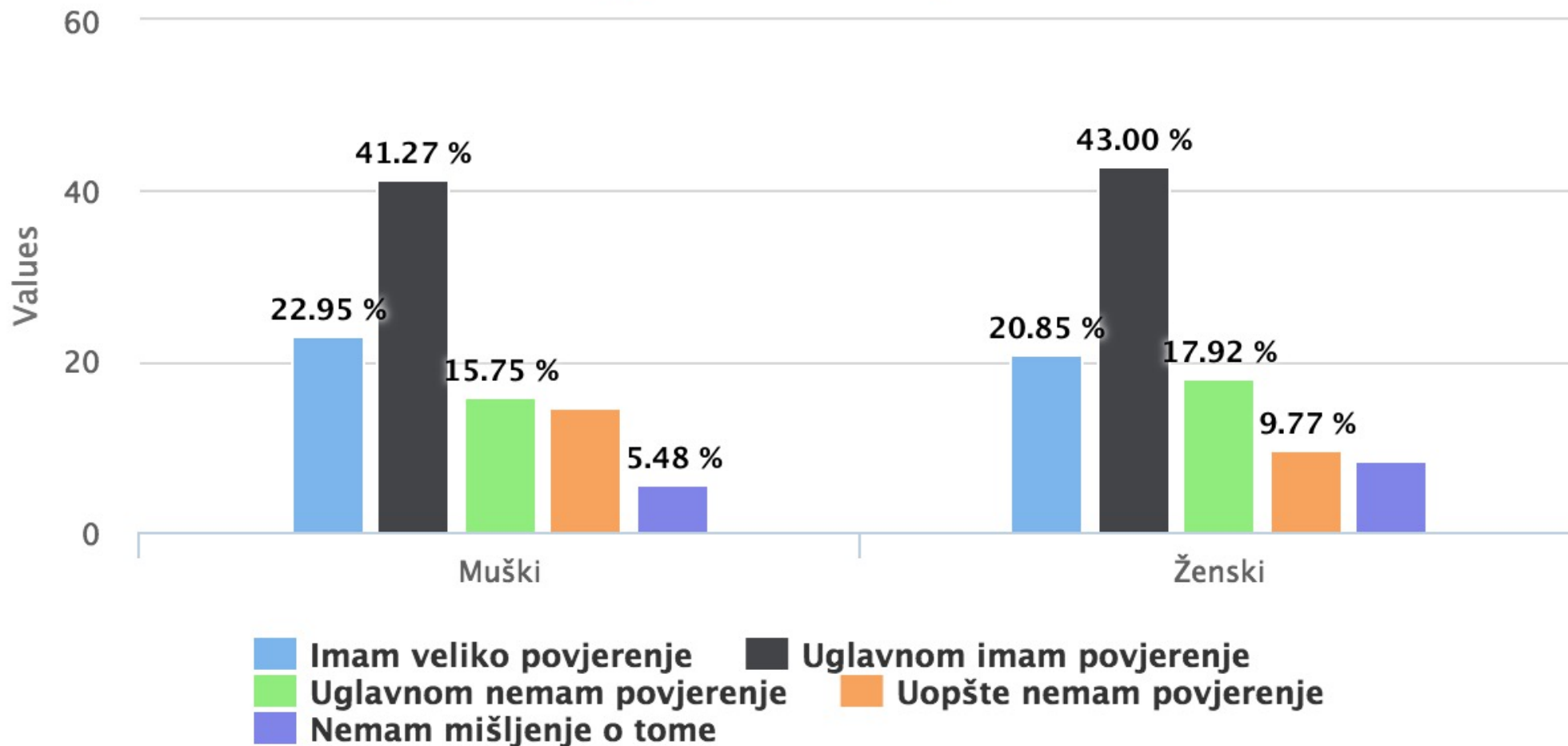


N (broj ispitanika) = 1197

$\chi^2(4) = 11.0974, p < 0.05$

$\chi^2(4) > 9.49$

Ukrštanje je statistički značajno



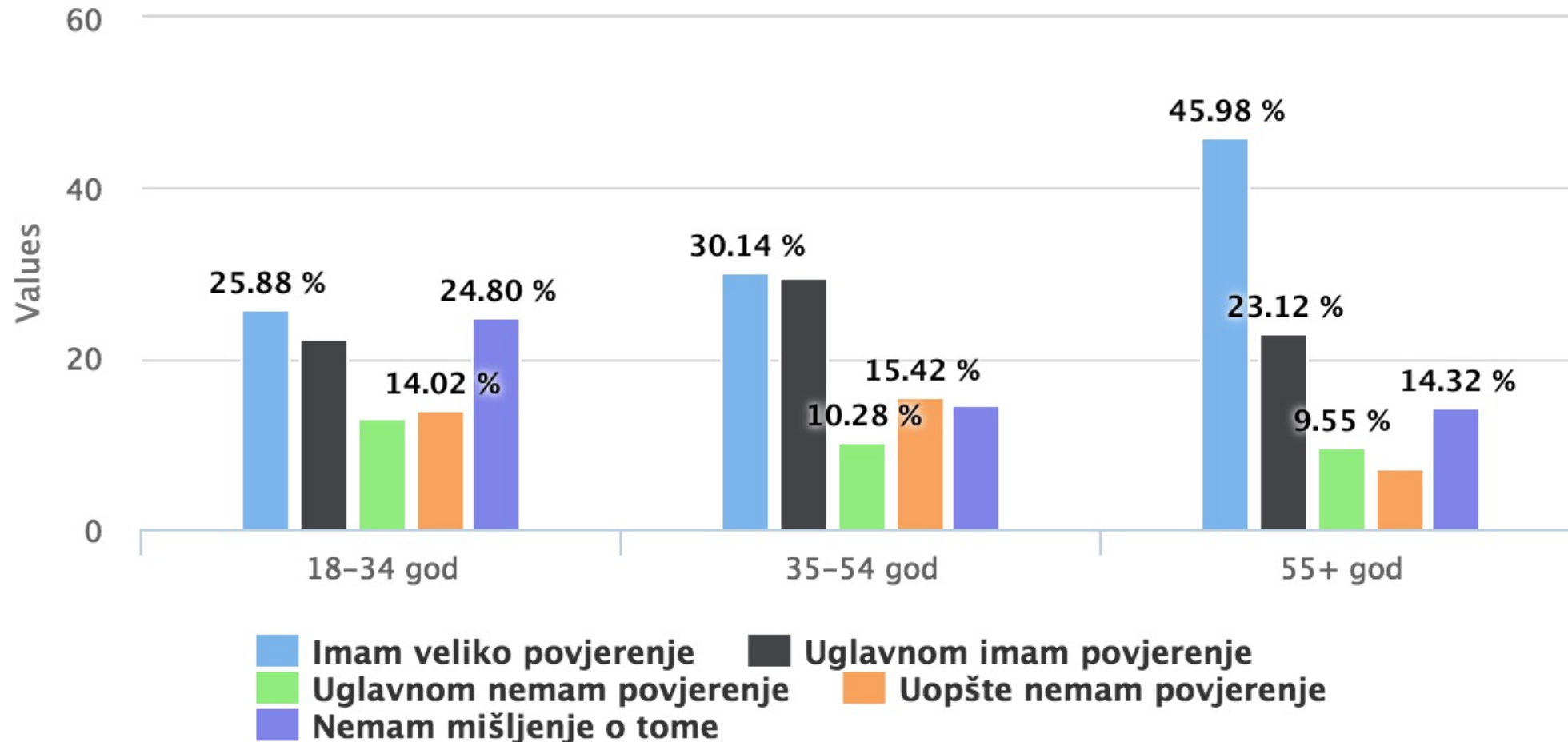
Povjerenje u institucije – Srpska pravoslavna crkva / Godine

N (broj ispitanika) = 1204

$\chi^2(8) = 61.4851, p < 0.05$

$\chi^2(8) > 15.51$

Ukrštanje je statistički značajno



Da li se slažete sa uvođenjem dodatnih slova u crnogorsku azbuku? / Obrazovanje

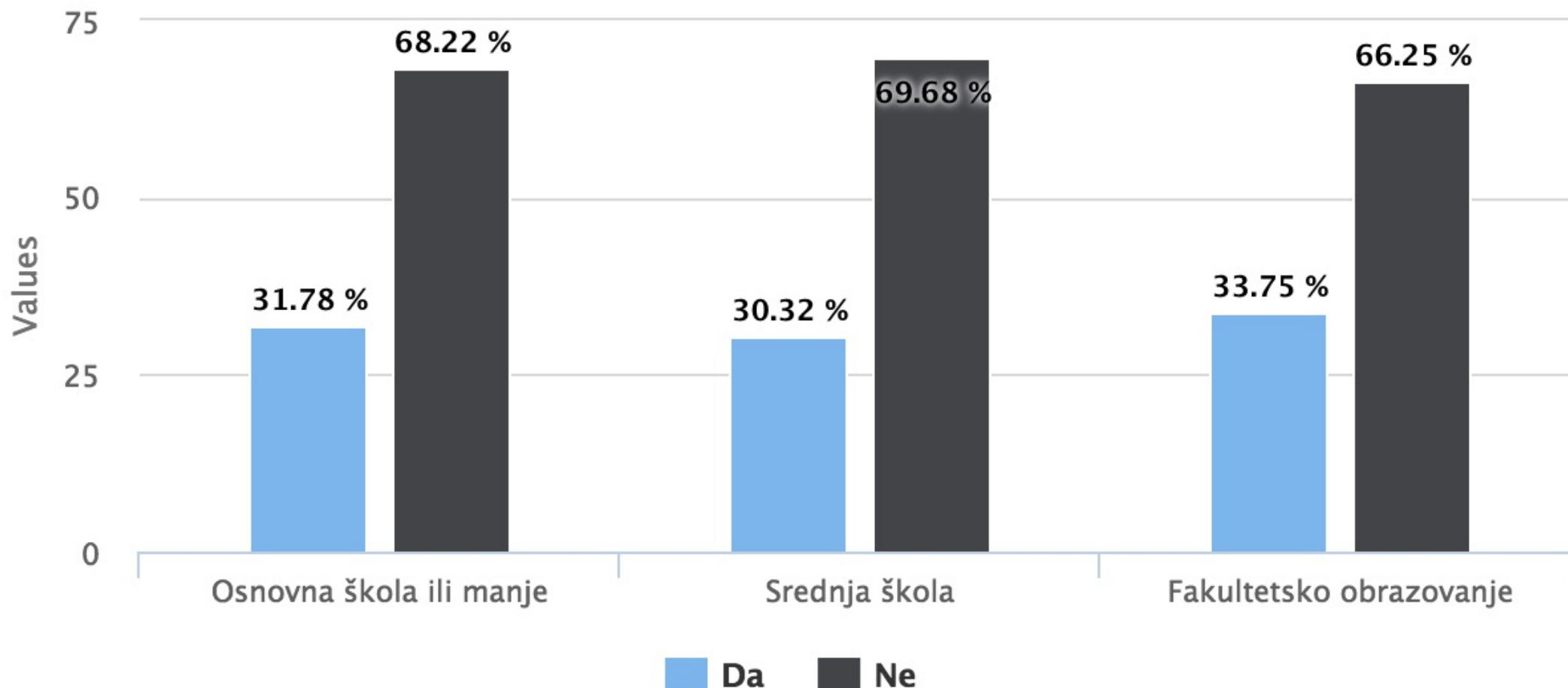


N (broj ispitanika) = 1133

$\chi^2(2) = 1.1086, p < 0.05$

$\chi^2(2) < 5.99$

Ukrštanje nije statistički značajno



Da li se slažete sa uvođenjem dodatnih slova u crnogorsku azbuku? / Godine

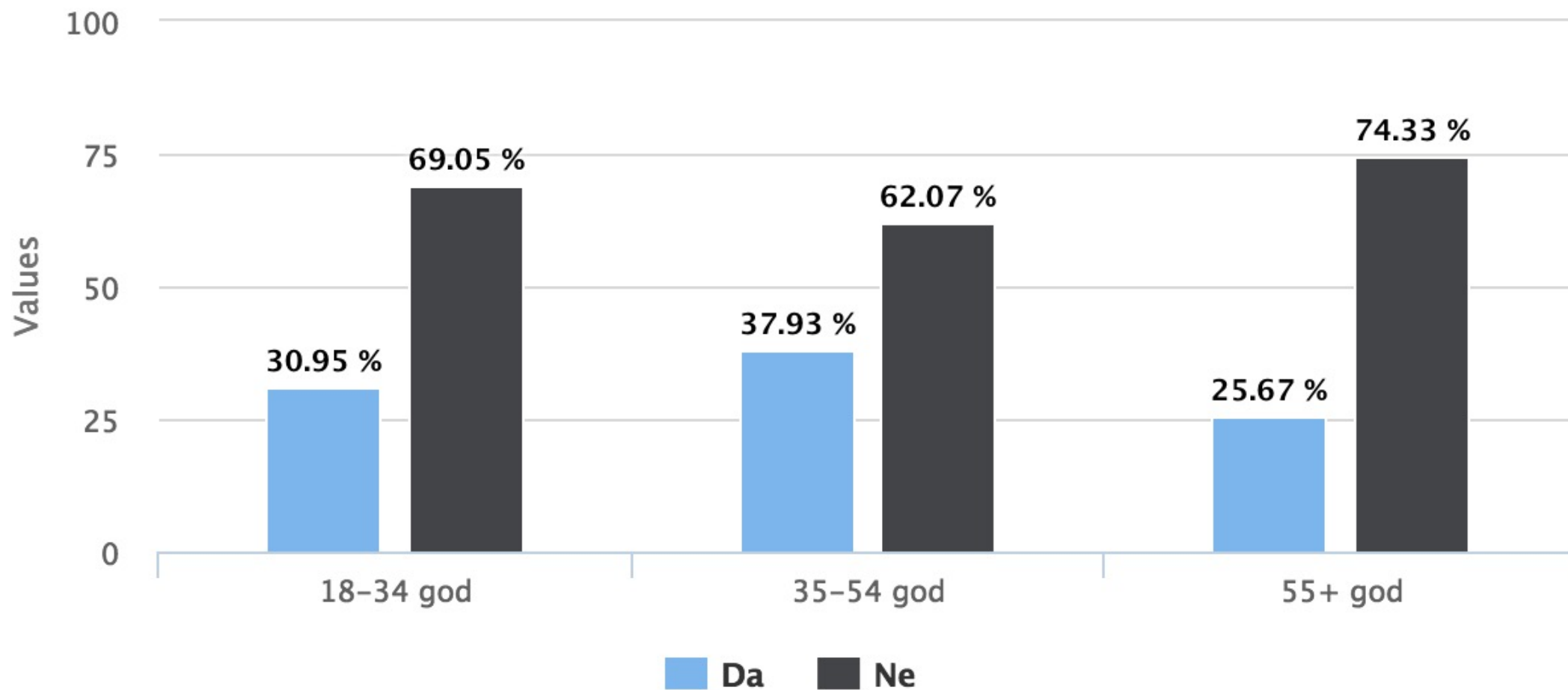


N (broj ispitanika) = 1133

$\chi^2(2) = 13.6496, p < 0.05$

$\chi^2(2) > 5.99$

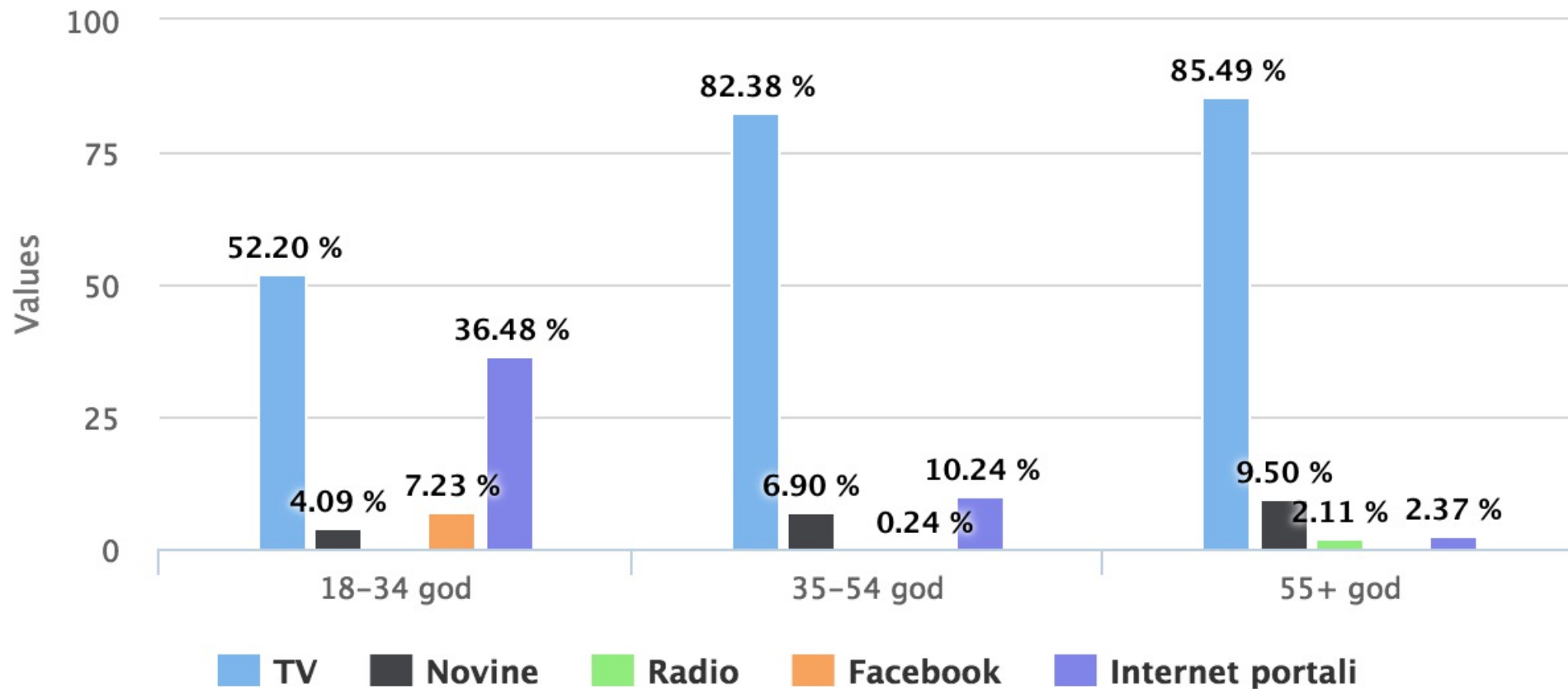
Ukrštanje je statistički značajno



Putem kojeg medija se prevashodno informišete o političkim dešavanjima? / Godine

N (broj ispitanika) = 1121
 $\chi^2(8) = 239.7361, p < 0.05$
 $\chi^2(8) > 15.51$

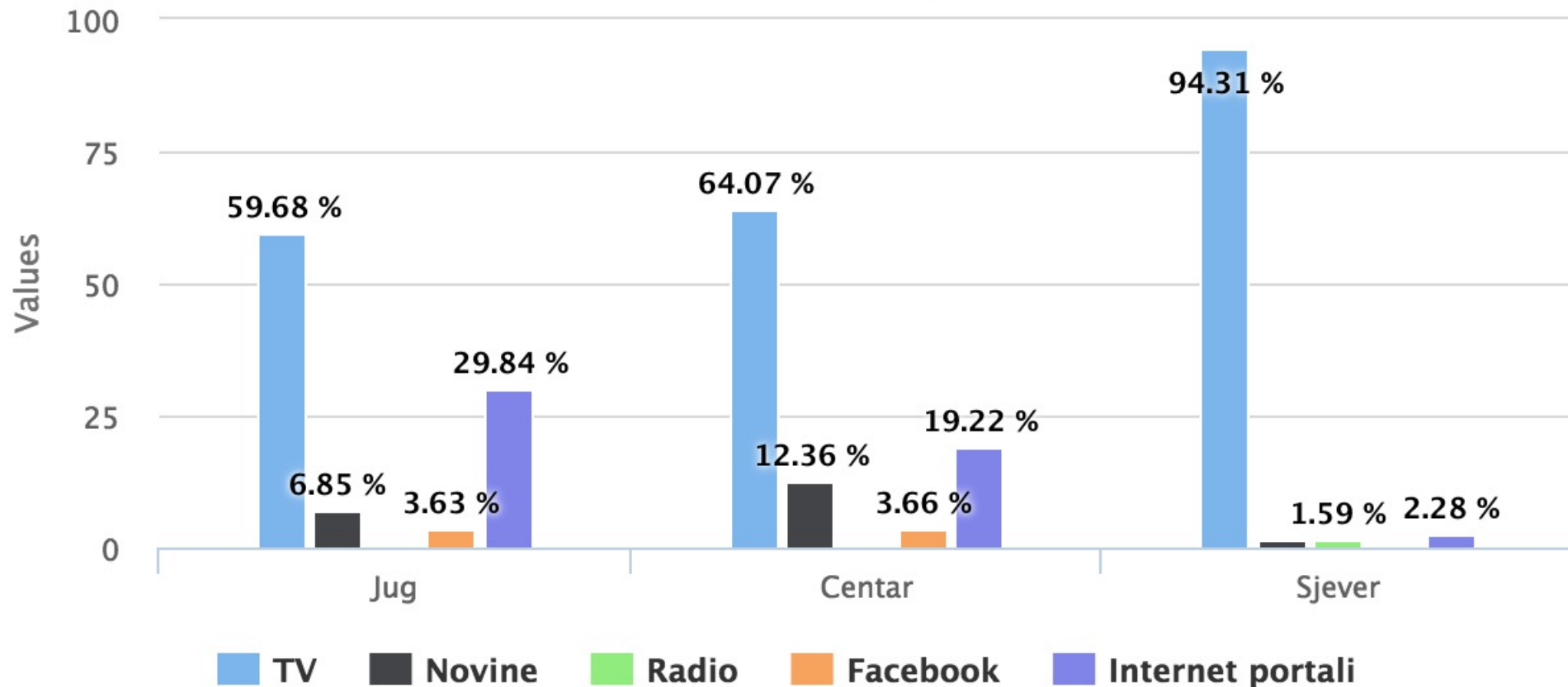
Ukrštanje je statistički značajno



Putem kojeg medija se prevashodno informišete o političkim dešavanjima? / Region

N (broj ispitanika) = 1121
 $\chi^2(8) = 180.5425, p < 0.05$
 $\chi^2(8) > 15.51$

Ukrštanje je statistički značajno



Neke od dostupnih baza u CG...



European
Values
Study



MNeS
montenegrin national election study