

# 10. Korelacijska analiza

TEME

1.	Koeficijenti korelacije – podsjetnik
2.	Pearsonov koeficijent korelacije ( $r$ )
3.	Pearsonov point-biserijalni koeficijent korelacije ( $r_{pb}$ )
4.	Spearmanov $\rho$ i Kendalov tau-b koeficijenti korelacije
5.	Parcijalna korelacija

## 10.1. Koeficijenti korelacije – podsjetnik

### BIVARIJATNA KORELACIJA:

- 1. Pearsonov koeficijent korelacije** – izražava stupanj i smjer linearne povezanosti dviju kvantitativnih normalno distribuiranih varijabli

Ključne pretpostavke za primjenu Pearsonova koeficijenta korelacije:

- Kvantitativne varijable – obje varijable trebaju biti kontinuirane kvantitativne (no ne trebaju imati iste mjerne jedinice!)
- Normalnost distribucija – obje varijable trebaju biti normalno distribuirane
- Linearna veza – varijable trebaju biti u linearnoj vezi
- Homoskedastičnost – podaci trebaju biti jednako distribuirani oko regresijskog pravca

- 2. Pearsonov point-biserijalni koeficijent korelacije** – izražava stupanj i smjer povezanosti jedne dihotomne i jedne kvantitativne normalno distribuirane varijable

- 3. Spearmanov koeficijent korelacije**

- 4. Kendalov tau-b koeficijent korelacije**

Spearmanov  $\rho$  (rho) koeficijent i Kendallov tau-b koeficijent neparametrijski su koeficijenti korelacije, tj. primjereni za određivanje stupnja i smjera povezanosti ordinalnih varijabli ili kvantitativnih varijabli koje nisu normalno distribuirane. Računaju se na rangovima, a ne na izvornim podacima.

### MULTIVARIJATNA KORELACIJA:

- 5. Parcijalna korelacija**

**Svi navedeni koeficijenti korelacije u JASP-u se određuju iz modula *Regression*, odabirom procedure (*Classical*) *Correlation*.**

Prije odabira primjerenog koeficijenta korelacije ( $r$ ,  $r_{pb}$ ,  $\rho$ ,  $b$ ) uputno je u JASP-u definirati pojedine varijable kao 'Nominal', 'Ordinal' ili 'Scale', ovisno o primjerenosti.

## 10.2. Pearsonov koeficijent korelacije ( $r$ )

### Odabir koeficijenta

### Odabir hipoteze

### Odabir prikaza rezultata

Radimo na primjerima u datoteci pod nazivom:

10A\_bivarijatna\_korelacija.sav

Želimo li izračunati Pearsonov  $r$ , u osnovnom meniju procedure *Correlations* valja prvo odabrati varijable za koje računamo korelaciju te ih prebaciti u okvir 'Variables'. Moguće je odabrati i više od dvije varijable, u tom će slučaju biti izračunate sve interkorelacije selektiranih varijabli.

Za početak, odredimo Pearsonov koeficijent korelacije varijabli **INDKLER** (indeks klerikalizma na kojemu viši rezultat označava viši stupanj klerikalizma) i **INDETNO** (indeks etnocentrizma na kojemu viši rezultat označava viši stupanj etnocentrizma). Uz 5% rizika pri zaključivanju testirajmo hipotezu da je stupanj klerikalizma povezan sa stupnjem etnocentrizma.

Pearsonov  $r$  koeficijent korelacije bi već trebao biti označen, pa samo treba provjeriti određujemo li ispravan koeficijent na podacima koje imamo na uzorku.

Sample Correlation Coefficient

- Pearson's r
- Spearman's rho
- Kendall's tau-b

Ispod je potrebno odrediti koju hipotezu testiramo. Ako želimo testirati nultu ili neusmjerenu alternativnu hipotezu, označit ćemo 'Correlated' (već označeno po zadanim postavkama).

Ako želimo testirati usmjerenu alternativnu hipotezu koja govori o pozitivnoj povezanosti dviju varijabli, označit ćemo 'Correlated positively', a ako želimo testirati usmjerenu alternativnu hipotezu koja govori o negativnoj povezanosti dviju varijabli, označit ćemo 'Correlated negatively'. U našem slučaju ostavit ćemo ovako kako je zadano jer testiramo neusmjerenu alternativnu hipotezu.

Alt. Hypothesis

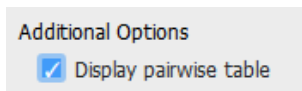
- Correlated
- Correlated positively
- Correlated negatively

U prikazu rezultata JASP po tvorničkoj postavci prikazuje **matricu korelacija** u kojoj se nalaze samo koeficijenti korelacije za odabrane varijable (bez p-vrijednosti ili veličine uzorka). Matrica korelacija korisna je kada želimo istovremeno prikazati interkorelacije većeg broja varijabli. No i tada bi u nju trebalo dodati podatke o p-vrijednosti i N-u.

Pearson's Correlations

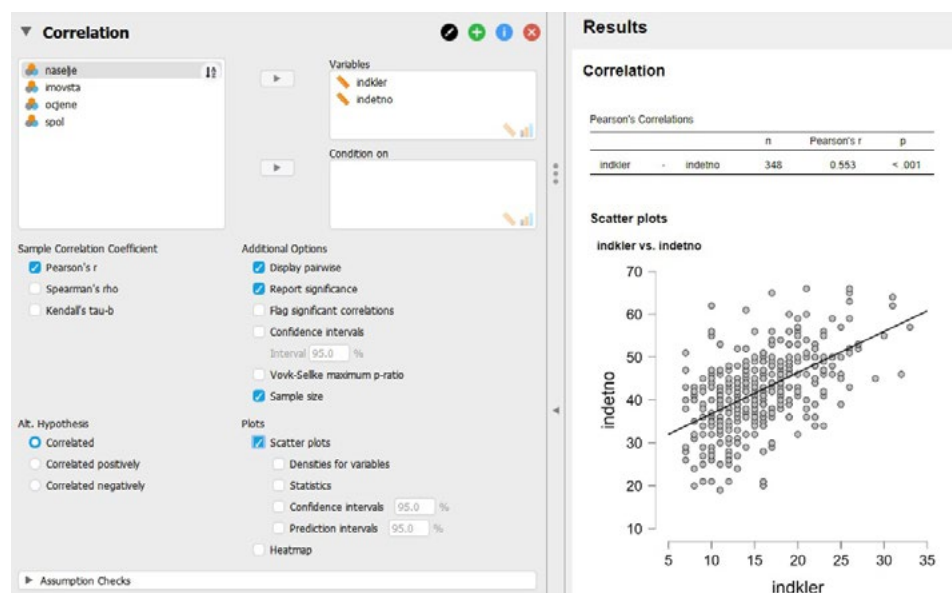
Variable		indkler	indetno
1. indkler	Pearson's r	—	—
	p-value	—	—
2. indetno	Pearson's r	0.553	—
	p-value	< .001	—

Ako želimo prikaz parova varijabli, tj. njihovih koeficijenata korelacije i pridruženih informacija, potrebno je aktivirati opciju 'Display pairwise' pod Additional Options. Ovakav je prikaz mnogo pregledniji u slučaju manjeg broja varijabli čije korelacije želimo odrediti.



U našem primjeru imamo samo dvije varijable za koje želimo odrediti iznos Pearsonova koeficijenta korelacije, pa će 'Display pairwise' prikaz biti pregledniji.

Da bismo dobili prikaz svih statističkih pokazatelja relevantnih za interpretaciju, potrebno je još pod Additional Options označiti 'Report Significance' kako bismo dobili odgovarajuću p-vrijednost za testiranu hipotezu te 'Sample size' za dobivanje informacije o veličinu uzorka na kojem je korelacija izračunata.



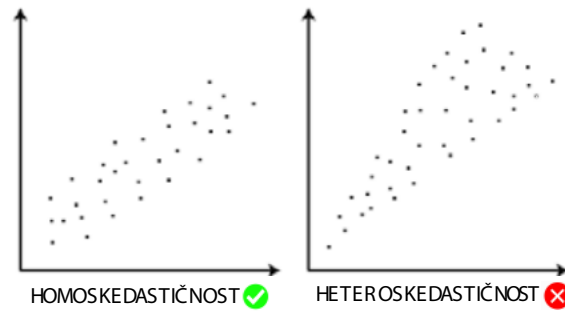
## Dijagram raspršenja

Pearsonov  $r$  koeficijent je korelacije koji iskazuje smjer i intenzitet linearne povezanosti varijabli. Pretpostavku o linearnoj povezanosti najjednostavnije je provjeriti dijagramom raspršenja (engl. *scatter plot*; *scattergram*)

**Dijagram raspršenja** (grafički prikaz podataka na dvije kvantitativne varijable istovremeno) koristan je za vizualnu inspekciju:

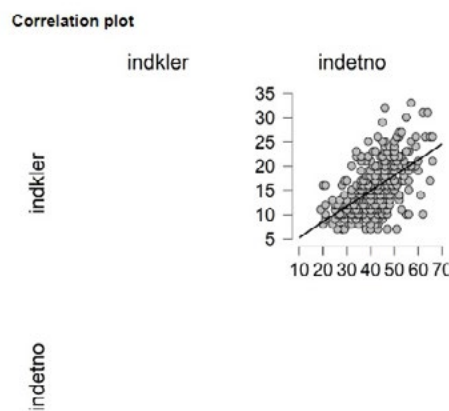
- linearnosti veze (prema obliku distribucije točaka), ali i
- smjera povezanosti (prema poziciji točaka i regresijskog pravca) te
- (samo načelno) jačine povezanosti dviju varijabli (prema udaljenosti točaka od regresijskog pravca), kao i
- pretpostavke o homoskedastičnosti (jednakom raspršenju oko pravca na svim vrijednostima varijabli).

## Podsjetnik: homoskedastičnost vs. heteroskedastičnost:



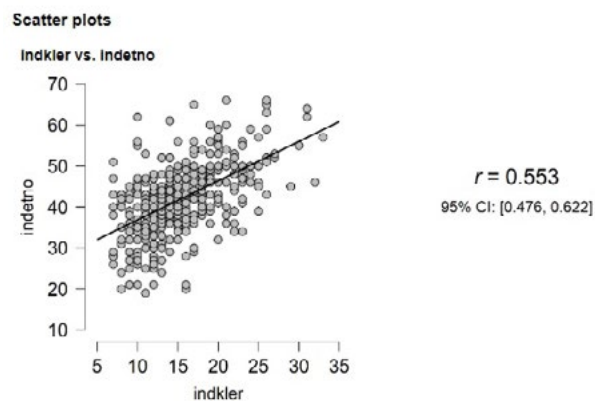
Dijagram raspršenja možemo zatražiti pod Plots, odabirom 'Scater plots'.

Dijagram raspršenja prikazat će se nešto drugačije (manje pregledno) ako isključimo 'Display pairwise' opciju:



Primijetite i da su varijable zamijenile osi, zato se čini da je oblik dijagrama raspršenja drugačiji na ove dvije slike, iako je riječ o istim podacima na obje slike.

Ako uz uključenu opciju 'Display pairwise' zatražimo dijagram raspršenja i još zatražimo pod 'Plots' prikaz opcije 'Statistics', dobivamo vrlo pregledan i informativan grafički prikaz:



Ovdje je još dodatno pod 'Additional options' uključena i opcija 'Confidence intervals'.

## Oznake statističke značajnosti korelacije za tablične prikaze rezultata

Uobičajeno je (*APA style*) da se u tekstu navodi iznos koeficijenta korelacije (i eventualno interval pouzdanosti) i egzaktna p-vrijednost. Ako se rezultati u nekom članku ili izvještaju prikazuju tablično, odnosno pomoću matrice korelacija, tada se razine značajnosti (5% / 1% / 0,1%) označavaju zvjezdicama (\* / \*\* / \*\*\*). Te zvjezdice dobivamo označavanjem opcije 'Flag significant correlations'. Ako u tablici koristimo zvjezdice, nije potrebno u istoj tablici prikazivati i p-vrijednosti jer tada imamo redundanciju informacija.

### Additional Options

- Display pairwise table
- Report significance
- Flag significant correlations

Primjer tablice u kojoj se koriste zvjezdice kao oznake statističke značajnosti koeficijenta korelacije:

Table 1

*Means, standard deviations, and correlations with confidence intervals*

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
1. rating	64.63	12.17						
2. complaints	66.60	13.31	.83** [.66, .91]					
3. privileges	53.13	12.24	.43* [.08, .68]	.56** [.25, .76]				
4. learning	56.37	11.74	.62** [.34, .80]	.60** [.30, .79]	.49** [.16, .72]			
5. raises	64.63	10.40	.59** [.29, .78]	.67** [.41, .83]	.45* [.10, .69]	.64** [.36, .81]		
6. critical	74.77	9.89	.16 [-.22, .49]	.19 [-.19, .51]	.15 [-.22, .48]	.12 [-.25, .46]	.38* [.02, .65]	
7. advance	42.93	10.29	.16 [-.22, .49]	.22 [-.15, .54]	.34 [-.02, .63]	.53** [.21, .75]	.57** [.27, .77]	.28 [-.09, .58]

*Note.* *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). \* indicates  $p < .05$ . \*\* indicates  $p < .01$ .

Izvor tablice: Stanley i Spence (2018).

Kod većih uzoraka i slaba povezanost može doseći statističku značajnost, stoga se u interpretaciji uvijek valja dodatno osvrnuti na jačinu povezanosti.

## Interpretacija jačine povezanosti varijabli

Ovdje nam može pomoći **Guilfordova tablica**, iako završna interpretacija jačine povezanosti ovisi i o teorijskim očekivanjima (npr. ako ne očekujemo da su dvije varijable uopće povezane, a dobijemo za njih  $r = 0,3$ , tada to možemo interpretirati kao iznenađujuće jaku vezu. No, ako isti iznos koeficijenta korelacije dobijemo za dvije varijable za koje očekujemo mnogo snažniju vezu, npr. oko  $0,8$ ; u tom slučaju nam se  $r = 0,3$  može činiti vrlo slabom vezom).

r	povezanost
±0,20 ili manji	slaba, gotovo neznatna
±0,21 – ±0,40	niska
±0,41 – ±0,70	umjerena
±0,71 – ±0,90	visoka (izrazita)
±0,91 – ±1,00	vrlo visoka

Vidi i: van Aswegen i Engelbrecht (2009: slika 1).

Za koeficijent korelacije također se može odrediti **interval pouzdanosti**. Naime, sam iznos koeficijenta korelacije određenog na podacima prikupljenim na nekom uzorku zapravo je **statistik** (obilježje uzorka) te je za njega moguće, uz određenu razinu rizika, procijeniti koliko iznosi u populaciji koju taj uzorak reprezentira. Razinu pouzdanosti u konstrukciji intervala usklađujemo s odabranom razinom rizika testiranja hipoteze o korelaciji dviju varijabli. U našem primjeru hipotezu smo testirali uz 5% rizika, pa ćemo odrediti 95% interval pouzdanosti za taj koeficijent korelacije.

Confidence intervals  
Interval  %

Rezultat se prikazuje u tablici s koeficijentom korelacije, odvojeno donja i gornja granica intervala pouzdanosti.

Pearson's Correlations

	n	Pearson's r	p	Lower 95% CI	Upper 95% CI
indkier - indetno	348	0.553	< .001	0.476	0.622

### Pearsonov $r$ – VAŽNO:

Nakon provedbe testiranja, u izvještaju/radu/zadaći potrebno je navesti sljedeće informacije:

- Navesti koju hipotezu testiramo
- Vrijednost koeficijenta korelacije
- Egzaktnu p-vrijednost
- Veličinu uzorka na kojemu je koeficijent izračunat
- 95%-tni (ili 99%-tni) interval pouzdanosti
- U slučaju da je Pearsonov koeficijent korelacije statistički značajan, u interpretaciji je potrebno komentirati smjer i jačinu povezanosti dviju varijabli.

**Kako zapisati koeficijent korelacije u tekstu?**

#1 npr.  $r = 0,553$ ;  $p < 0,001$ ;  $N = 348$ ; 95% CI: 0,476-0,622 ili

#2 npr.  $r = 0,553$ ;  $p < 0,001$ ;  $N = 348$ ; 95% CI [0,476, 0,622]

**Za ovaj zadatak možemo napisati rješenje:**

Uz 5% rizika zaključujemo da je stupanj klerikalizma umjereno pozitivno povezan sa stupnjem etnocentrizma ( $r = 0,553$ ;  $p < 0,001$ ;  $N = 348$ ; 95% CI: 0,476-0,622).

**Dodatak: Kako izgleda matrica korelacija s vrijednostima iznad i ispod dijagonale (u ovom slučaju za muške i ženske sudionike)?**

	Cross-Correlations					
	1	2	3	4	5	6
(1) Variable A		.78**	.74**	-.15**	-.21**	-.20**
(2) Variable B	.70**		.79**	-.22**	-.26**	-.25**
(3) Variable C	.70**	.70**		-.19**	-.21**	-.22**
(4) Variable D	-.18**	-.29**	-.11		.82**	.80**
(5) Variable E	-.26**	-.31**	-.17*	.82**		.88**
(6) Variable F	-.23**	-.19*	-.15*	.80**	.82**	

*Notes.* Zero-order correlation coefficients in the male sample are shown below the main diagonal, while coefficients in the female sample are presented above it; \*  $p < .05$ . \*\*  $p < .01$ .

Preuzeto i prilagođeno iz Koletić i sur. (2019).



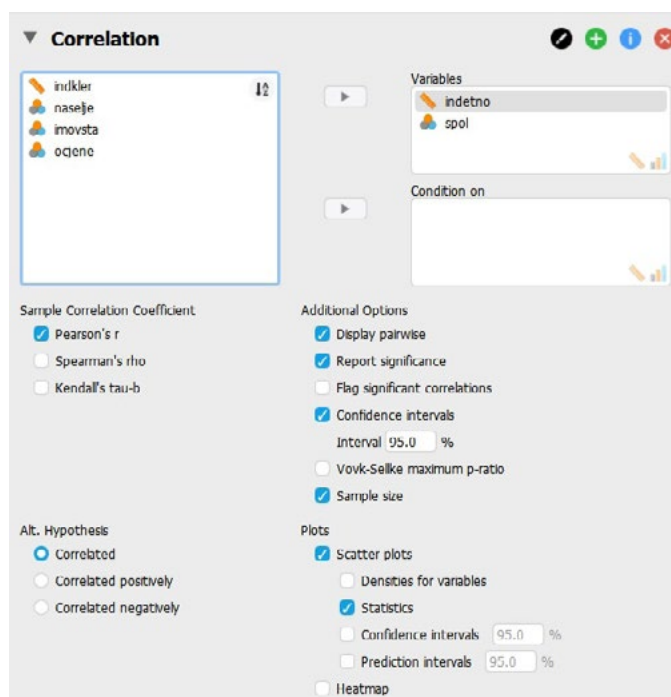
### 10.3. Pearsonov point- biserijalni koeficijent korelacije ( $r_{pb}$ )

Radimo na primjerima u datoteci pod nazivom:  
10A\_bivarijatna\_korelacija.sav

Nema zasebne procedure za određivanje Pearsonova point-biserijalnog koeficijenta korelacije, već se on automatski određuje kad u okvir 'Variables' stavimo jednu kvantitativnu i jednu dihotomnu nominalnu varijablu te zatražimo izračunavanje Pearsonova koeficijenta korelacije.

Primjer: Uz 1% rizika testirajte hipotezu da je SPOL (1 = ženski; 2 = muški) povezan sa stupnjem etnocentrizma (varijabla INDETNO – indeks etnocentrizma na kojemu viši rezultat označava veću sklonost etnocentrizmu).

Označit ćemo iste opcije koje daju potreban i pregledan prikaz rezultata kao i kod "običnog" Pearsonova koeficijenta korelacije:



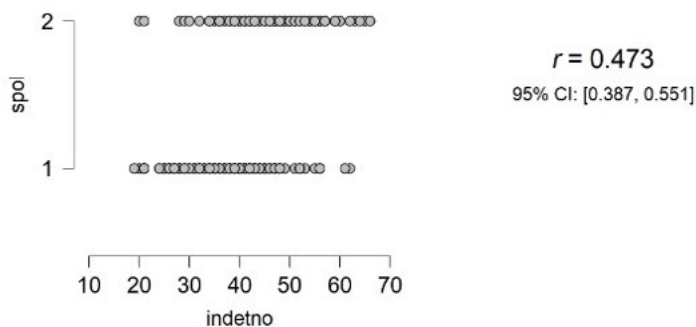
Napomena: hipotezu testiramo uz 1% rizika, pa ćemo zatražiti 99%-tni interval pouzdanosti za koeficijent korelacije.

Dobivamo sljedeće rezultate:

Pearson's Correlations		n	Pearson's r	p	Lower 95% CI	Upper 95% CI
indetno	- spol	348	0.473	< .001	0.387	0.551

#### Scatter plots

indetno vs. spol



## Interpretacija $r_{pb}$ -a

Za pravilnu interpretaciju point-biserijalnog koeficijenta korelacije moramo znati koja numerička oznaka na dihotomnoj varijabli označava koju grupu. Osobito je praktično koristiti kodove 0 i 1.

U našem slučaju nemamo kodove 0 i 1, već 1 i 2, ali logika interpretacije ostaje ista:

- ako je predznak koeficijenta korelacije **pozitivan**, to znači da **grupa s većim kodom postiže više rezultate** na kvantitativnoj varijabli (odnosno da grupa s manjim kodom postiže niže rezultate na kvantitativnoj varijabli);
- ako je predznak koeficijenta korelacije **negativan**, to znači da **grupa s manjim kodom postiže više rezultate** na kvantitativnoj varijabli (odnosno da grupa s većim kodom postiže niže rezultate na kvantitativnoj varijabli).

U interpretaciji  $r_{pb}$ -a nećemo, dakle, govoriti o smjeru povezanosti, već moramo objasniti predznak koeficijenta tako što ćemo navesti koja grupa postiže više rezultate na kvantitativnoj varijabli.

U našem slučaju predznak koeficijenta korelacije je pozitivan ( $r_{pb} = 0,473$ ), što znači da muškarci (grupa s većim kodom, kod 2) postižu više rezultate na indeksu etnocentrizma od žena.

Dodatno, valja imati na umu da vrijednost point-biserijalnog koeficijenta korelacije ne može postići vrijednost 1, već u slučaju simetrične raspodjele dihotomne varijable (npr., 50% muškarci, 50% žene) maksimalno 0,811, dok je kod asimetričnih raspodjela dihotomne varijable maksimum u pravilu niži. Zato Guilfordova tablica nije u potpunosti primjenjiva za interpretaciju jačine povezanosti dviju varijabli, već se može koristiti samo načelno, imajući na umu da je primjerice korelacija 0,79 već vrlo visoka za  $r_{pb}$ .

Načelno su i moguće usporedbe više point-biserijalnih koeficijenata u smislu koja je veza jača, a koja slabija (npr. spol s etnocentrizmom

ili spol s klerikalizmom), ali pri tome treba imati na umu da su takve usporedbe opravdane kad su omjeri proporcija dihotomnih kategorija slični (npr. približno 50/50% ili 70/30%) jer iznos  $r_{pb}$ -a ovisi i o odnosu proporcija dihotomnih varijabli.

Prema iznosu p-vrijednosti u našem primjeru ( $p < 0,001$ ) možemo zaključiti da je povezanost varijabli spol i INDETNO statistički značajna uz 1% rizika.

**Za ovaj zadatak možemo napisati rješenje:**

Testirali smo alternativnu hipotezu da je spol povezan sa stupnjem etnocentrizma. Uz 1% rizika zaključujemo da je spol statistički značajno povezan sa stupnjem etnocentrizma ( $r_{pb} = 0,473$ ;  $p < 0,001$ ;  $N = 348$ ; 99% CI: 0,359 - 0,573) na način da muškarci postižu više rezultate na indeksu etnocentrizma od žena.

**Pearsonov  $r_{pb}$  VAŽNO:**

Nakon provedbe testiranja, u izvještaju/radu/zadaći potrebno je navesti sljedeće informacije:

- a. Navesti koju hipotezu testiramo
- b. Vrijednost koeficijenta korelacije
- c. Egzaktnu p-vrijednost
- d. Veličinu uzorka na kojemu je koeficijent izračunat
- e. U slučaju da je Pearsonov point-biserijalni koeficijent korelacije statistički značajan, u interpretaciji je potrebno komentirati koja grupa postiže više rezultate na kvantitativnoj varijabli.

**Kako zapisati koeficijent korelacije u tekstu?**

npr.  $r_{pb} = 0,473$ ;  $p < 0,001$ ;  $N = 348$ ; 99% CI: 0,359 - 0,573

## 10.4. Spearmanov $\rho$ i Kendallov tau-b koeficijenti korelacije

Radimo na primjerima u datoteci pod nazivom:  
10A\_bivarijatna\_korelacija.sav

Spearmanov  $\rho$  koeficijent i Kendallov tau-b koeficijenti primjereni su u slučaju da su varijable čiju povezanost želimo izračunati ordinalne. Dodatno, Spearmanov  $\rho$  koeficijent možemo koristiti i na kvantitativnim varijablama za koje nisu zadovoljene osnovne pretpostavke za provedbu Pearsonove korelacije.

Kao i Pearsonov  $r$ , ovi su neparametrijski koeficijenti korelacije simetrični (postiću vrijednosti od -1 do 1).

Kendallov tau-b koeficijent primjereniji je no Spearmanov  $\rho$  u slučaju većeg broja vezanih rangova (što se događa kada više ispitanika na nekoj varijabli ima jednak rezultat, pa tada imaju i jednake rangove kad se podaci na toj varijabli rangiraju). Naime, u sociološkim istraživanjima, u kojima se varijable mjere skalama malog raspona varijacija, a broj ispitanika je velik, vezani rangovi najčešće jesu problem.

Oba koeficijenta računaju se na rangovima, a ne na izvornim vrijednostima varijabli. Rangiranje nije potrebno ručno vršiti prije računanja neparametrijskih koeficijenata korelacije, jer sam program automatski računa Spearmanov i Kendallov koeficijent na rangovima.

### Primjer 1.

Testirajte, uz 5% rizika, hipotezu da su veličina mjesta stanovanja i stupanj klerikalizma pozitivno povezani.

U navedenom primjeru imamo jednu ordinalnu varijablu s pet kategorija (NASELJE) te jednu kvantitativnu, stoga su primjereni neparametrijski koeficijenti korelacije.

Paziti: testira se usmjerena hipoteza:

The image shows the 'Sample Correlation Coefficient' dialog box in SPSS. It is divided into several sections: 'Sample Correlation Coefficient' with radio buttons for 'Pearson's r', 'Spearman's rho', and 'Kendal's tau-b'; 'Additional Options' with checkboxes for 'Display pairwise', 'Report significance', 'Flag significant correlations', 'Confidence intervals', 'Vovk-Selke maximum p-ratio', and 'Sample size'; 'Alt. Hypothesis' with radio buttons for 'Correlated', 'Correlated positively', and 'Correlated negatively'; and 'Plots' with checkboxes for 'Scatter plots', 'Densities for variables', 'Statistics', 'Confidence intervals', 'Prediction intervals', and 'Heatmap'. The 'Correlated positively' option is circled in red. The 'Confidence intervals' and 'Prediction intervals' are both set to 95.0%.

Ako zatražimo izračun oba koeficijenta, primijetit ćemo da je Spearmanov koeficijent nešto većeg iznosa od Kendallova, no oba su statistički značajna uz 5% rizika (iako ne bi bili da smo testirali neusmjereno!).

Correlation Table

			Spearman		Kendall		
			n	rho	p	tau B	p
naselje	-	indkler	348	0.096	0.038	0.074	0.037

Note. All tests one-tailed, for positive correlation

Interpretacija bi se trebala odnositi na vezu između rangova dviju varijabli, a ne vezu originalnih rezultata (iako to često u primjeni ovih koeficijenata izostaje).

**Interpretacija može glasiti (statistički preciznije):**

Uz 5% rizika zaključujemo da su rang veličine mjesta stanovanja i rang rezultata na indeksu klerikalizma slabo pozitivno povezani ( $\rho = 0,096$ ;  $p = 0,038$ ;  $N = 348$ ).

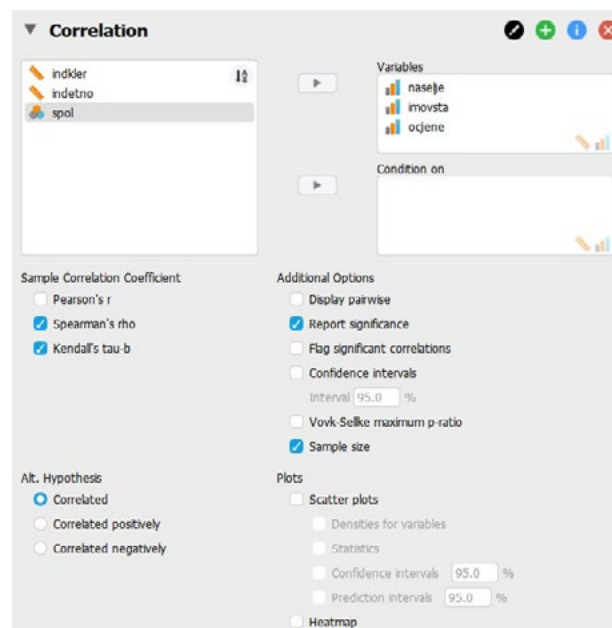
Interpretacija može također glasiti

(statistički manje precizno, ali nije neuobičajeno):

Uz 5% rizika zaključujemo da su veličina mjesta stanovanja i sklonost klerikalizmu slabo pozitivno povezani ( $\rho = 0,096$ ;  $p = 0,038$ ;  $N = 348$ ).

**Primjer 2:** Odredimo interkorelacije triju ordinalnih varijabli:

NASELJE (veličina naselja), IMOVSTA (samoprocjena imovinskog stanja) i OCJENE testiranjem nultih hipoteza uz 5% rizika.



S obzirom na to da sad imamo 3 varijable, i određujemo tri koeficijenta korelacije, ostavit ćemo zadani prikaz korelacijske matrice, tj. nećemo označiti opciju 'Display Pairwise'.

No, zatražit ćemo prikaz p-vrijednosti ('Report significance') i veličinu uzorka ('Sample size').

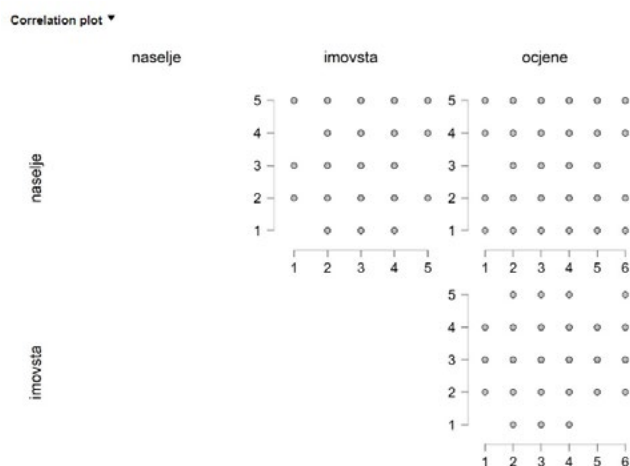
## Prikaz rezultata:

### Correlation

Variable		naselje	imovsta	ocjene
1. naselje	n	—	—	—
	Spearman's rho	—	—	—
	p-value	—	—	—
	Kendall's Tau B	—	—	—
	p-value	—	—	—
2. imovsta	n	348	—	—
	Spearman's rho	0.136	—	—
	p-value	0.011	—	—
	Kendall's Tau B	0.119	—	—
	p-value	0.011	—	—
3. ocjene	n	348	348	—
	Spearman's rho	0.121	0.053	—
	p-value	0.024	0.321	—
	Kendall's Tau B	0.099	0.046	—
	p-value	0.024	0.316	—

U ovom slučaju s isključivo ordinalnim varijablama relativno malog raspona varijacija (svaka varijabla ima 5 kategorija), Kendallov tau-b koeficijent korelacije primjereniji je od Spearmanova  $\rho$  zbog velike količine vezanih rangova.

Problem vezanih rangova neizravno je razvidan i na dijagramima raspršenja jer su pokrivena gotovo sve kombinacije vrijednosti unutar pojedinih parova varijabli, a ispitanika je mnogo više (348):



### Za ovaj zadatak možemo napisati rješenje:

Testirali smo nulte hipoteze o međupovezanostima veličine naselja, samoprocjene imovinskog stanja i ocjena. Uz 5% rizika zaključujemo da je veličina naselja slabo pozitivno povezana s imovinskim stanjem ( $b = 0,119$ ;  $p = 0,011$ ;  $N = 348$ ) te, također pozitivno, ali još slabije s ocjenama ( $b = 0,099$ ;  $p = 0,034$ ;  $N = 348$ ), dok ocjene i imovinsko stanje nisu statistički značajno povezani ( $b = 0,046$ ;  $p = 0,316$ ;  $N = 348$ ).

### Spearmanov rho i Kendallov tau-b – VAŽNO:

Nakon provedbe testiranja, u izvještaju/radu/zadaći potrebno je navesti sljedeće informacije:

- Navesti koju hipotezu testiramo
- Vrijednost koeficijenta korelacije
- Egzaktnu p-vrijednost
- Veličinu uzorka na kojemu je koeficijent izračunat
- U slučaju da je testirani koeficijent korelacije statistički značajan, u interpretaciji je potrebno komentirati smjer i jačinu povezanosti rangova dviju varijabli.

## 10.5. Parcijalna korelacija

Radimo na primjerima u datoteci pod nazivom:  
10B\_parcijalna\_korelacija.sav

Određivanje parcijalne korelacije, odnosno korelacije dviju varijabli uz isključenje utjecaja treće varijable, zatražit ćemo tako da:

- u okvir 'Variables' unesemo varijable čiju parcijalnu korelaciju želimo odrediti (dvije kvantitativne varijable);
- u okvir 'Condition on' unesemo varijablu čiji utjecaj želimo kontrolirati, tj. isključiti iz veze prvih dviju varijabli;
- zatražimo izračunavanje Pearsonova koeficijenta korelacije.

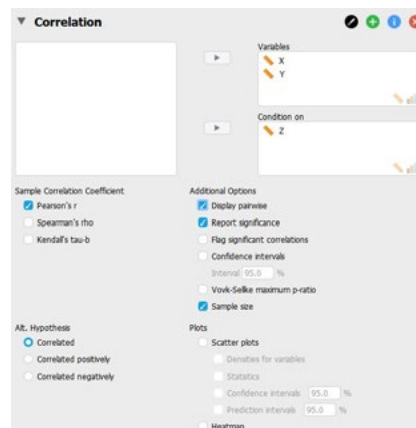
### Dvije tipične vrste parcijalizacije:

1. Lažna korelacija – treća (kontrolirana) varijabla je visoko korelirana s prve dvije i uzorkuje njihovu "lažnu" jaku vezu. Takva treća varijabla (sada je nazivamo zbunjujućom varijablom ili *confunderom*) smanjuje parcijalnu korelaciju u odnosu na izravnu korelaciju.
2. Supresija – treća (kontrolirana) varijabla u različitim je vezama (različitog predznaka) s prvom i drugom varijablom te potiskuje (pokriva) njihovu izravnu povezanost. Takva treća varijabla (sada je nazivamo supresorom) povećava parcijalnu korelaciju u odnosu na izravnu korelaciju.

### Primjer 1

Odredite parcijalnu korelaciju varijabli X i Y uz isključenje utjecaja varijable Z.

Označit ćemo iste opcije koje daju potreban i pregledan prikaz rezultata kao i kod "običnog" Pearsonova koeficijenta korelacije (osim intervala pouzdanosti koji se ne može odrediti za koeficijent parcijalne korelacije).



Prikaz rezultata:

### Correlation

Pearson's Partial Correlations

	n	Pearson's r	p
X - Y	100	0.037	0.715
Conditioned on variables: Z			

Ispod tablice s rezultatima naznačeno je da je riječ o parcijalnoj korelaciji varijabli navedenih u tablici (X i Y) uz isključenje utjecaja ('Conditioned on variables:') varijable Z.

**Za ovaj zadatak možemo napisati rješenje:**

Testirali smo nultu hipotezu o povezanosti varijabli X i Y uz isključenje utjecaja varijable Z na njihovu vezu. Uz 5% rizika zaključujemo da parcijalna korelacija varijabli X i Y uz isključenje utjecaja varijable Z nije statistički značajna ( $r_{XY.Z} = 0,037$ ;  $p = 0,715$ ;  $N = 100$ ).

**Parcijalna korelacija – VAŽNO:**

Nakon provedbe testiranja, u izvještaju/radu/zadaći potrebno je navesti sljedeće informacije:

- a. Navesti koju hipotezu testiramo (vezu koje dvije varijable testiramo uz isključenje koje treće varijable)
- b. Vrijednost koeficijenta korelacije
- c. Egzaktnu p-vrijednost
- d. Veličinu uzorka na kojemu je koeficijent izračunat
- e. U slučaju da je testirani koeficijent korelacije statistički značajan, u interpretaciji je potrebno komentirati smjer i jačinu povezanosti dviju varijabli uz isključenje utjecaja treće.

**Literatura**

- Goss-Sampson, M. A. (2019). *Statistička analiza u JASP programu: vodič za studente*. Korelaciona analiza (str. 57-62). URL: [http://static.jasp-stats.org/Manuals/Statistic%cc%8cka\\_analiza\\_u\\_JASP\\_programu\\_v0.10.2.pdf](http://static.jasp-stats.org/Manuals/Statistic%cc%8cka_analiza_u_JASP_programu_v0.10.2.pdf)
- Koletić, G., Štulhofer, A., Tomić, I., i Knežević Čuča, J. (2019). Associations between Croatian Adolescents' Use of Sexually Explicit Material and Risky Sexual Behavior: A Latent Growth Curve Modeling Approach. *International Journal of Sexual Health*, 31(1): 77-91. <https://doi.org/10.1080/19317611.2019.1567640>
- Navarro, D. J., Foxcroft, D.R., i Faulkenberry, T.J. (2019). *Learning Statistics with JASP: A Tutorial for Psychology Students and Other Beginners*. Poglavlje 11. Correlation and linear regression. URL: <https://tomfaulkenberry.github.io/JASPbook/chapters/chapter11.pdf>
- Stanley, D. J. i Spence, J. R. (2018). apaTables. URL: <https://dstanley4.github.io/apaTables/articles/apaTables.html> (11.2.2022.)
- van Aswegen i Engelbrecht (2009). The relationship between transformational leadership, integrity and an ethical climate in organisations. *SA Journal of Human Resource Management*, 7(1): 1-9. <https://sajhrm.co.za/index.php/sajhrm/article/view/175>