

6. Kontingen- cijske tablice i χ^2 test

TEME

- | | |
|----|---|
| 1. | Kontingencijska tablica |
| 2. | Provedba hi-kvadrat testa za nezavisnost dviju nominalnih varijabli |
| 3. | Testiranje oblika distribucije hi-kvadrat testom |

6.1. Kontingencijska tablica

Podsjetnik! Što je kontingencijska tablica?

Kontingencijska tablica prikazuje učestalost pojavljivanja svake moguće kombinacije vrijednosti dviju nominalnih (ili nominaliziranih) varijabli.

Podsjetnik! Kako izgleda kontingencijska tablica?

		ZELOBRAZ			
spol		fakultet	magisterij	doktorat	Total
muški	Count	52.000	19.000	16.000	87.000
	Expected count	43.764	16.873	26.364	87.000
	% within row	59.770%	21.839%	18.391%	100.000%
	% within column	62.651%	59.375%	32.000%	52.727%
	% of total	31.515%	11.515%	9.697%	52.727%
ženski	Count	31.000	13.000	34.000	78.000
	Expected count	39.236	15.127	23.636	78.000
	% within row	39.744%	16.667%	43.590%	100.000%
	% within column	37.349%	40.625%	68.000%	47.273%
	% of total	18.788%	7.879%	20.606%	47.273%
Total	Count	83.000	32.000	50.000	165.000
	Expected count	83.000	32.000	50.000	165.000
	% within row	50.303%	19.394%	30.303%	100.000%
	% within column	100.000%	100.000%	100.000%	100.000%
	% of total	50.303%	19.394%	30.303%	100.000%

Kontingencijska tablica s varijablom SPOL kao varijablom retka i varijablom ZELOBRAZ (željena razina obrazovanja) kao varijablom stupca. Tablica prikazuje empirijske frekvencije (*Count*), teorijske frekvencije (*Expected count*), postotke retka (*% within row*), postotke stupca (*% within column*) te ukupne postotke (*% of total*).

Interpretacija svih vrijednosti u ćeliji 'muški'/'fakultet':

- 52 muškarca u uzorku žele završiti fakultet
- pod pretpostavkom da su varijable SPOL i ZELOBRAZ u populaciji nezavisne, u ovom uzorku očekujemo 43,76 muškaraca koji žele završiti fakultet
- od svih muškaraca u uzorku, njih 59,77% želi završiti fakultet
- od svih ispitanika u uzorku koji žele završiti fakultet, njih 62,65% su muškarci
- od svih ispitanika u uzorku, 31,52% su muškarci koji žele završiti fakultet

Kako interpretiramo kontingencijsku tablicu? Odredimo i interpretiramo **postotke nezavisne varijable***.

Pravilo za odabir postotaka koje ćemo izračunati i interpretirati jest:

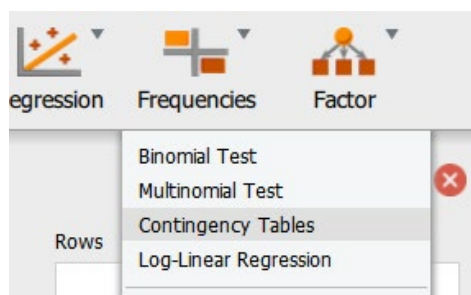
- Ako je nezavisna varijabla varijabla retka, za interpretaciju koristimo postotke retka (uspoređujemo postotke retka u istom stupcu);
- Ako je nezavisna varijabla varijabla stupca, za interpretaciju koristimo postotke stupca (uspoređujemo postotke stupca u istom redu).

* U ovom priručniku koristimo termine "zavisna" i "nezavisna" varijabla u kontekstu korelacijskih nacrtava istraživanja pri čemu zavisnost proizlazi iz teorijskih pretpostavki o odnosima između varijabli, a ne iz metodoloških preduvjeta za utvrđivanje zavisnosti, poput onih u eksperimentalnim nacrtima istraživanja. Drugim riječima, iako koristimo termine "zavisna" i "nezavisna" varijabla, ne donosimo kauzalne zaključke.

6.1.1. Procedura izrade kontingencijske tablice

Radimo na primjerima u datoteci pod nazivom:
06A_hi-kvadrat_test.sav

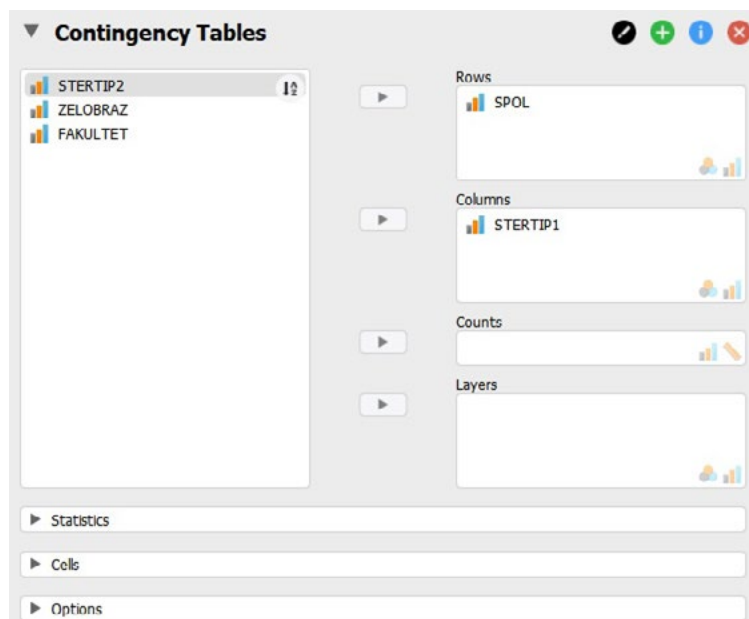
Nakon što ste učitali podatke u JASP, potrebno je pokrenuti proceduru 'Contingency Tables' koja se nalazi u izborniku 'Frequencies'.



Potom odabiremo nominalne (ili ordinalne) varijable koje ćemo uvesti kao varijablu retka i varijablu stupca.

Za primjer ćemo napraviti kontingencijsku tablicu varijabli *spol* i *stertip1* (dihotomna procjena stereotipa "Francuzi su dobri ljubavnici": istinito/neistinito)

Postavit ćemo varijablu *spol* kao varijablu retka, a varijablu *stertip1* kao varijablu stupca (mogli smo i obrnuto, nema pravila).



Napomena: neki istraživači vole nezavisnu varijablu postaviti kao varijablu retka, no to nije obavezno, a odabir koja će varijabla biti u retku, a koja u stupcu može ovisiti i o broju kategorija pojedine varijable te o preferiranom izgledu tablice (želimo li više redaka ili više stupaca).

6.1.2. Istovremena izrada više kontingencijskih tablica

U prozore Rows i Columns možemo staviti i više varijabli. Izvršimo li sljedeći odabir varijabli redaka i stupaca:



dobit ćemo četiri kontingencijske tablice u kojima će biti križane varijable:

- *spol* × *stertip1*
- *spol* × *stertip2*
- *zelobraz* × *stertip1*
- *zelobraz* × *stertip2*

Kontingencijske tablice kreiraju se u sličnim slučajevima tako da se **svaka odabrana varijabla retka križa sa svakom odabranom varijablom stupca.**

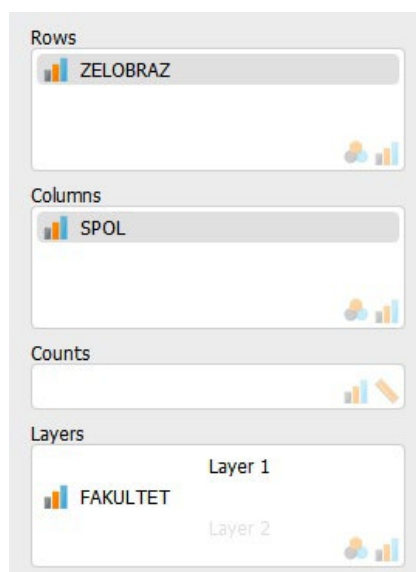
6.1.3. Kontingencijske tablice poduzoraka (kontrola treće varijable)

Želimo li križati dvije varijable **uz istodobnu kontrolu utjecaja treće varijable**, koristit ćemo prozor označen s **Layer**. Pomoću te opcije izradit će se zasebna kontingencijska tablica za svaku vrijednost varijable u Layer prozoru.

Na primjer, imamo tri varijable:

- *zelobraz*, s kategorijama 1 = fakultet, 2 = magisterij i 3 = doktorat;
- *spol*, s kategorijama 1 = muško i 2 = žensko;
- *fakultet*, s kategorijama 1 = društveno–humanistički, 2 = tehnički i 3 = medicinsko–biotehnološki.

Izvršimo li sljedeći odabir varijabli:



...dobit ćemo tri kontingencijske tablice, **po jednu za svaku kategoriju layer varijable (fakultet).**

U prvoj kontingencijskoj tablici bit će uzeti u obzir samo oni ispitanici koji na varijabli *fakultet* imaju vrijednost 1 (društveno-humanistički fakulteti), u drugoj samo oni koji na toj varijabli imaju vrijednost 2 (tehnički fakulteti), a u trećoj tablici samo oni ispitanici koji na varijabli *fakultet* imaju vrijednost 3 (medicinsko-biotehnoški fakulteti).

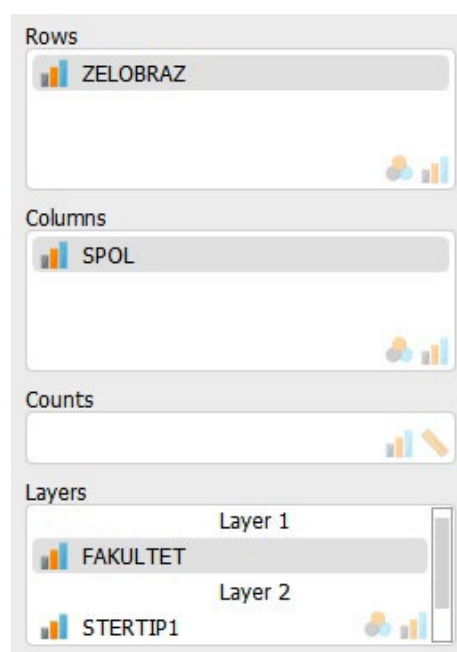
Contingency Tables

FAKULTET	ZELOBRAZ	SPOL		Total
		muški	ženski	
društ.-humanistički	fakultet	6	14	20
	magisterij	6	10	16
	doktorat	8	11	19
	Total	20	35	55
tehnički	fakultet	33	5	38
	magisterij	6	0	6
	doktorat	7	4	11
	Total	46	9	55
medicinsko-bioteh.	fakultet	13	12	25
	magisterij	7	3	10
	doktorat	1	19	20
	Total	21	34	55
Total	fakultet	52	31	83
	magisterij	19	13	32
	doktorat	16	34	50
	Total	87	78	165

6.1.4. Kontingencijske tablice poduzoraka (kontrola više od jedne varijable)

Moguće je odabrati i više od jedne *layer* varijable.

Ako na primjer dodamo varijablu *stertip1* kao drugu *layer* varijablu:



... dobit ćemo kontingencijske tablice za svaki *fakultet* unutar svake kategorije odgovora na varijabli *stertip*:

Contingency Tables ▾

STERTIP1	FAKULTET	ZELOBRAZ	SPOL		Total
			muški	ženski	
istinito	društ.-humanistički	fakultet	0	1	1
		magisterij	0	1	1
		doktorat	1	1	2
		Total	1	3	4
tehnički	tehnički	fakultet	6	1	7
		magisterij	0	0	0
		doktorat	0	0	0
		Total	6	1	7
medicinsko-bioteh.	medicinsko-bioteh.	fakultet	1	1	2
		magisterij	0	0	0
		doktorat	0	2	2
		Total	1	3	4
Total	Total	fakultet	7	3	10
		magisterij	0	1	1
		doktorat	1	3	4
		Total	8	7	15
neistinito	društ.-humanistički	fakultet	6	13	19
		magisterij	6	9	15
		doktorat	7	10	17
		Total	19	32	51
tehnički	tehnički	fakultet	25	4	30
		magisterij	6	0	6
		doktorat	7	4	11
		Total	39	8	47
medicinsko-bioteh.	medicinsko-bioteh.	fakultet	12	11	23
		magisterij	7	3	10
		doktorat	1	17	18
		Total	20	31	51
Total	Total	fakultet	44	28	72
		magisterij	19	12	31
		doktorat	15	31	46
		Total	78	71	149

... kao i za sve podatke zajedno (u nastavku tablice):

Total	društ.-humanistički	fakultet	6	14	20	
		magisterij	6	10	16	
		doktorat	8	11	19	
Total	Total	Total	20	35	55	
		tehnički	fakultet	33	5	38
			magisterij	6	0	6
doktorat	7		4	11		
Total	Total	Total	46	9	55	
		medicinsko-bioteh.	fakultet	13	12	25
			magisterij	7	3	10
doktorat	1		19	20		
Total	Total	Total	21	34	55	
		Total	fakultet	52	31	83
			magisterij	19	13	32
doktorat	16		34	50		
Total	Total	Total	87	78	165	

6.1.5. Odabir elemenata ćelija kontingencijske tablice

Nakon selektiranja varijabli koje će tvoriti kontingencijsku tablicu, koristeći izbornik 'Cells', odabiremo koje će elemente sadržavati ćelije kontingencijske tablice.

▾ Cells

<p>Counts</p> <p><input type="checkbox"/> Expected</p>	<p>Percentages</p> <p><input type="checkbox"/> Row</p> <p><input type="checkbox"/> Column</p> <p><input type="checkbox"/> Total</p>
--	---

Counts (frekvencije):

- **Expected** – teorijske (očekivane) frekvencije ako su varijabla retka i varijabla stupca međusobno nezavisne;

Percentages (postoci):

- **Row** – postoci retka;
- **Column** – postoci stupca;
- **Total** – ukupni postoci.

Napomena 1: prikaz empirijskih frekvencija (u SPSS-u: *observed counts*, u JASP-u samo *counts*) u kontingencijskoj tablici u ovoj je verziji JASP-a zadan i ne može se isključiti (za razliku od SPSS-a).

Napomena 2: JASP zasad ne nudi opciju prikaza reziduala pojedinih ćelija (standardizirane razlike između empirijskih i teorijskih frekvencija), korisnih za interpretaciju doprinosa pojedinih ćelija povezanosti varijabli koje su u kontingencijskoj tablici.

6.2. Provedba hi-kvadrat testa za nezavisnost dviju nominalnih varijabli

Želimo li osim izrade kontingencijske tablice provesti i hi-kvadrat test ili neki drugi primjereni test nezavisnosti varijabli te izračunati koeficijente asocijacije relevantne za određeni tip kontingencijske tablice (2×2 , 4×2 , itd.), možemo to učiniti preko izbornika *Statistics* u sklopu procedure *Contingency tables*.

Statistički testovi od interesa (poznati iz gradiva OSS2):

- χ^2 – hi-kvadrat test nezavisnosti dviju nominalnih varijabli
- χ^2 **continuity correction** – hi-kvadrat test nezavisnosti dviju nominalnih varijabli uz Yatesovu korekciju
- **Fisherov egzaktni test** – za tablice 2×2 , pokreće se odabiranjem *Log odds ratio (2x2 only)* i odabirom hipoteze koju želimo testirati na varijablama u kontingencijskoj tablici.

Podsjetnik:

Yatesova korekcija koristi se u bilo kojem od ovih slučajeva:

- ▶ ako obje varijable imaju samo dvije kategorije (2×2 tablice)
- ▶ ako je broj ispitanika mali ($N < 40$)
- ▶ ako je teorijska frekvencija u bilo kojoj ćeliji manja od 5

Primjer ispisa testova nezavisnosti varijabli (križamo varijable *spol* i *stertip2*):

Contingency Tables ▼

		stertip2		
		istinito	neistinito	Total
muški	Count	7.000	80.000	87.000
	% within row	8.046 %	91.954 %	100.000 %
ženski	Count	3.000	75.000	78.000
	% within row	3.846 %	96.154 %	100.000 %
Total	Count	10.000	155.000	165.000
	% within row	6.061 %	93.939 %	100.000 %

→ kontingencijska tablica s empirijskim frekvencijama i postocima nezavisne varijable (*spol*).

Chi-Squared Tests

	Value	df	p
χ^2	1.274	1	0.259
χ^2 continuity correction	0.643	1	0.423
N	165		

→ Pearsonov hi-kvadrat test
→ Pearsonov hi-kvadrat test uz Yatesovu korekciju

Log Odds Ratio ▼

	Log Odds Ratio	95% Confidence Intervals		p
		Lower	Upper	
Odds ratio	0.783	-0.606	2.171	
Fisher's exact test	0.778	-0.743	2.605	0.336

→ Fisherov egzaktni test

Interpretacija kontingencijske tablice (deskriptivna razina analize):

Od svih muškaraca u uzorku, njih 8% smatra da je stereotip 2 ("Crnogorci su lijeni") istinit, dok od svih žena to smatra njih 3,8%.
ILI

Da su Crnogorci lijeni, smatra 8% muškaraca i 3,8% žena u uzorku.

Inferencijalna razina analize:

Nulta hipoteza: Prihvaćanje stereotipa da su Crnogorci lijeni nije povezano sa spolom.

Hipotezu ćemo testirati uz 5% rizika.

Interpretacija i zapis rezultata testova:

1. **Pearsonov hi-kvadrat test:** Uz 5% rizika zaključujemo da prihvaćanje stereotipa da su Crnogorci lijeni nije povezano sa spolom ($\chi^2 = 1,274$; $df = 1$; $p = 0,259$).
2. **Pearsonov hi-kvadrat test uz Yatesovu korekciju:** Uz 5% rizika zaključujemo da prihvaćanje stereotipa da su Crnogorci lijeni nije povezano sa spolom ($\chi^2 = 0,643$; $df = 1$; $p = 0,423$).
3. **Fisherov egzaktni test:** Uz 5% rizika zaključujemo da prihvaćanje stereotipa da su Crnogorci lijeni nije povezano sa spolom ($p = 0,336$).

Napomene:

- a) **p-vrijednost!** Koristeći statističke pakete za obradu i analizu podataka, testovni statistik ne uspoređujemo s kritičnim vrijednostima u određenim tablicama (za χ^2 -distribuciju, t-distribuciju, F-distribuciju itd.) kao što radimo u "ručnom računu", već za odluku o hipotezi koristimo p-vrijednost.

p-vrijednost jest vjerojatnost dobivanja testovnog statistika (ovdje je to izračunata hi-kvadrat vrijednost) **u odgovarajućoj sampling distribuciji** (ovdje hi-kvadrat distribucija uz $df = 1$) **pod uvjetom istinitosti nulte hipoteze.**

Ako je dobiveni testovni statistik jako malo vjerojatan za dobivanje u populaciji u kojoj je nulta hipoteza istinita (manji od razine rizika koju su odabrali), tada odbacujemo nultu hipotezu.

Tehnički je, dakle, potrebno samo provjeriti je li je naša dobivena p-vrijednost:

- manja od 0,01 (ako testiramo uz 1% rizika) ili
- manja od 0,05 (ako testiramo uz 5% rizika)

Ako je dobivena p-vrijednost manja od odabrane razine rizika (obično 0,01 ili 0,05), odbacujemo nultu hipotezu, prihvaćamo alternativnu i utvrđujemo da postoji statistički značajan učinak (povezanost ili razlika).

- b) Veličina uzorka i χ^2 vrijednost

Temeljem iznosa p-vrijednosti, odnosno vjerojatnosti dobivenog testovnog statistika, odbacujemo ili prihvaćamo naše hipoteze. Kao i kod drugih statističkih testova, vrijednost χ^2 testovnog statistika ovisi o veličini uzorka, tj. ako povećamo uzorak, povećavamo vjerojatnost odbacivanja nulte hipoteze (a time i vjerojatnost pogreške tipa alfa).

Stoga u slučaju odbacivanja nulte hipoteze uvijek valja izračunati odgovarajući mjeru asocijacije kako bismo dobili informaciju o jačini povezanosti između dvije varijable.

c) Kao i kod izrade kontingencijskih tablica, χ^2 test može se u JASP-u provesti odvojeno za pojedine poduzorke tako da ukupni uzorak podijelimo prema kategorijama neke nominalne ili ordinalne varijable.

Primjerice, testiramo povezanost između varijabli *spol* i *stertip2*, ali odvojeno po fakultetima. Tada bismo varijablu *fakultet* uveli u funkciju 'Layers' te osim zasebnih kontingencijskih tablica za svaku skupinu fakulteta, dobili također i zasebne rezultate provedenih testova za svaku skupinu fakulteta.

Chi-Squared Tests ▼

fakultet		Value	df	p
društ.-humanistički	X ²	2.672	1	0.102
	X ² continuity correction	1.405	1	0.236
	N	55		
tehnički	X ²	1.716	1	0.190
	X ² continuity correction	0.113	1	0.737
	N	55		
medicinsko-bioteh.	X ²	3.360	1	0.067
	X ² continuity correction	1.192	1	0.275
	N	55		
Total	X ²	1.274	1	0.259
	X ² continuity correction	0.643	1	0.423
	N	165		

Interpretacija ove tablice:

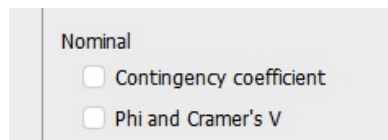
Uz 5% rizika zaključujemo da prihvaćanje stereotipa da su Crnogorci lijeni nije povezano sa spolom ni među studentima društveno-humanističkih fakulteta ($\chi^2 = 1,405$; $df = 1$; $p = 0,236$), ni među studentima tehničkih fakulteta ($\chi^2 = 0,113$; $df = 1$; $p = ,737$), kao ni među studentima medicinsko-biotehnoloških fakulteta ($\chi^2 = 1,192$; $df = 1$; $p = 0,275$).

6.2.1. Mjere asocijacije

U slučaju odbacivanja nulte hipoteze uvijek valja izračunati odgovarajući mjeru asocijacije kako bismo dobili informaciju o jačini povezanosti između dvije varijable.

U slučaju prihvatanje nulte hipoteze nema smisla izvještavati o iznosu mjere asocijacije jer smo već zaključili da veza između varijabli nije statistički značajna pa je njezin iznos na uzorku irelevantan.

Željene mjere asocijacije možete odabrati pod 'Statistics' unutar 'Contingency Tables' procedure:



Za nominalne varijable:

- **Contingency coefficient:** Pearsonov koeficijent kontingencije (za simetrične tablice npr. 3×3, 4×4...);
- **Phi coefficient and Cramer's V: Φ** —koeficijent (za tablice 2×2) i Cramerov V (za sve tablice);

Statistička značajnost koeficijenata asocijacije ne testira se zasebno, oni su statistički značajni (ili nisu) na istoj razini rizika kao i povezanost dviju varijabli u kontingencijskoj tablici.

Interpretacija mjera asocijacije:

— navedeni koeficijenti za nominalne varijable simetrični su (ne razlikuju zavisnu od nezavisne varijable) te postižu vrijednosti između 0 i 1 (maksimalna moguća vrijednost koeficijenta kontingencije ovisi o broju redaka i stupaca u kontingencijskoj tablici). Vrijednosti bliske nuli znače slabu povezanost, dok vrijednosti bliske jedinici znače jaku povezanost.

6.2.2.
Dodatne napomene
o izvještavanju
i interpretaciji

VAŽNO #1:

Nakon provedbe testiranja u izvještaju/radu/kolokviju/ispitu potrebno je navesti sljedeće informacije:

- a. Koju hipotezu testiramo i ishod testiranja
- b. Kontingencijsku tablicu s adekvatnim postocima ćelija (nužno ako smo utvrdili statističku značajnost)
- c. Iznos testovnog statistika (iznos hi-kvadrata)
- d. Broj stupnjeva slobode (navesti u zagradi pored znaka χ^2 ili zapisati odvojeno: npr. $df = 101$)
- e. Egzaktnu p-vrijednost (uvijek na tri decimale)

Npr. $\chi^2_{(101)} = 125,83; p = 0,060$ [nije značajno]
ili
 $\chi^2 = 125,83; df = 101; p = 0,060$ [nije značajno]

- f. Ako je testovni statistik značajan, potrebno je dodati i adekvatni koeficijent asocijacije:

Npr. $\chi^2_{(101)} = 125,83; p = 0,025; V = 0,234$

Napomena! Ako je p-vrijednost manja od 0,001, tako je i navodimo:

Npr. $\chi^2_{(101)} = 125,83; p < 0,001; V = 0,456$

VAŽNO #2:

χ^2 testom testiramo povezanost dvije nominalne varijable. Međutim, u interpretaciji možemo koristiti termine "razlika" ili "razlikuje se" kada interpretiramo odnos između pojedinih ćelija u kontingencijskoj tablici. Primjerice: "Utvrdili smo statistički značajnu povezanost između spola i glasačkih preferencija. Razlika je uočena kod desnih orijentiranih stranaka, odnosno veći postotak muškaraca preferira desno orijentirane stranke."

6.3. Testiranje oblika distribucije (engl. *goodness of fit test*) hi-kvadrat testom

Radimo na primjerima u datoteci pod nazivom:
06B_testiranje_oblika_distribucije.sav

Osim za testiranje nezavisnosti varijabli, hi-kvadrat test može se koristiti i za testiranje oblika distribucije varijabli (kao i za neke druge svrhe koje nisu obrađene u ovom priručniku).

Hi-kvadrat test omogućuje testiranje različitih hipoteza o obliku distribucije varijable te se njime može testirati razlikuje li se neka empirijska distribucija statistički značajno od normalne, pravokutne ili bilo koje druge teorijske ili empirijske distribucije.

Proceduri za testiranje oblika distribucije hi-kvadrat testom pristupa se odabirom *Frequencies* → *Multinomial Test*:



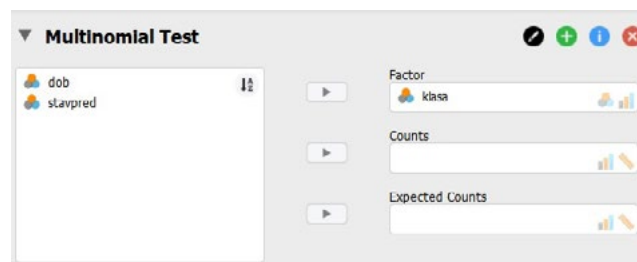
Otvorimo novu datoteku naziva '06B Testiranje oblika distribucije'.

Pretpostavimo da su u nekom istraživanju na slučajnom uzorku od 100 članova stranke X mjerene i sljedeće tri varijable:

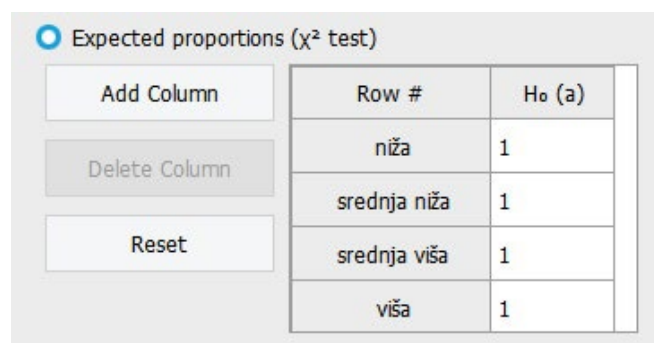
- *dob* – dob ispitanika;
- *stavpred* – stav o predsjedniku stranke;
- *klasa* – klasna pripadnost članova.

6.3.1. Testiranje distribucije na pravokutnost

Da bismo testirali hipotezu da se distribucija varijable *klasa* ne razlikuje statistički značajno od pravokutne distribucije (tj. da su proporcije klasa u populaciji članstva jednake), u glavnom meniju procedure Multinomial Test selektirat ćemo varijablu koje želimo testirati (*klasa*) te je prebaciti u prozor *Factor*:



Pod *Test Value* označit ćemo *Expected proportions* (χ^2 test). Zadano je već testiranje oblika distribucije na pravokutnost jer su u stupcu H_0 (a) postavljene vrijednosti jedinica koju označavaju jednake omjere frekvencija u svim kategorijama kao teorijske frekvencije.



Hipoteza koju testiramo:

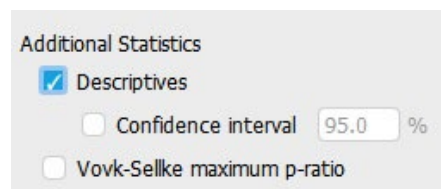
H_0 : U populaciji su sve kategorije varijable *klasa* jednako zastupljene.

ILI:

H_0 : Varijabla *klasa* u populaciji je distribuirana pravokutno.

Obje formulacije imaju u statističkom smislu jednako značenje.

Pod *Additional Statistics* možemo zatražiti i deskripciju (*Descriptives*)...



...čime ćemo dobiti tablicu s empirijskim (*Observed*) i teorijskim frekvencijama pod pretpostavkom hipoteze koju testiramo (*Expected: H₀ (a)*).

Descriptives		
klasa	Observed	Expected: H ₀ (a)
niža	21	25
srednja niža	25	25
srednja viša	26	25
viša	28	25

Za tu je tablicu zadan prikaz frekvencija, no može se promijeniti u prikaz proporcija.

Additional Statistics		Display
<input checked="" type="checkbox"/> Descriptives		<input type="radio"/> Counts
<input type="checkbox"/> Confidence interval	95.0 %	<input checked="" type="radio"/> Proportions

Descriptives		
klasa	Observed	Expected: H ₀ (a)
niža	0.210	0.250
srednja niža	0.250	0.250
srednja viša	0.260	0.250
viša	0.280	0.250

Rezultat testiranja hipoteze:

Multinomial Test			
	χ^2	df	p
H ₀ (a)	1.040	3	0.792

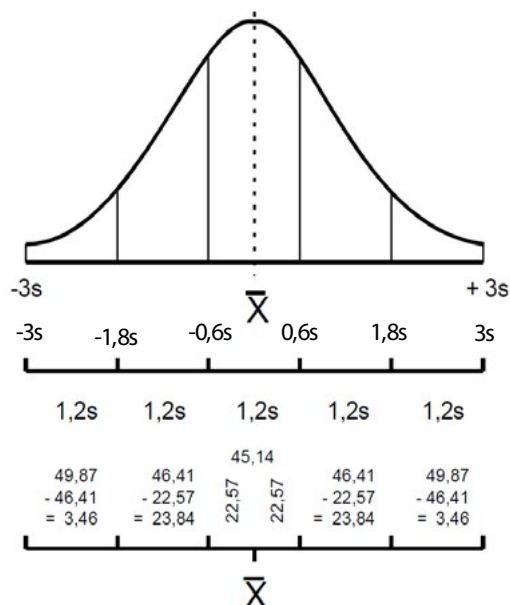
Hi-kvadrat za navedenu tablicu ($\chi^2 = 1,040$) uz 3 stupnja slobode nije statistički značajan ($p = 0,792$), pa ne možemo odbaciti (nultu) hipotezu da je distribucija varijable KLASA u populaciji članova stranke X pravokutna.

Zaključujemo, dakle – uz 5% rizika – da je distribucija varijable KLASA u populaciji članova stranke χ pravokutno distribuirana ($\chi^2 = 1,040$; $df = 3$; $p = 0,792$).

6.3.2. Testiranje distribucije na normalnost

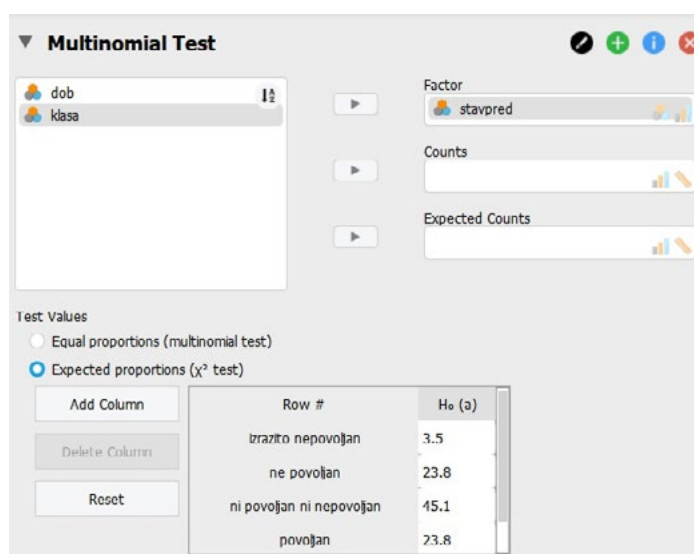
Preko hi-kvadrat testa možemo testirati i hipotezu da je distribucija neke varijable normalna. Da bismo testirali normalnost distribucije varijable *stavpred*, valja prvo odrediti teorijske frekvencije, odnosno frekvencije pojedinih kategorija koje bismo očekivali u normalnoj distribuciji.

Budući da varijabla *stavpred* ima 5 kategorija, očekivane proporcije možemo – preko tablice površine ispod normalne distribucije – odrediti na sljedeći način:



Da bismo proveli test, valja prvo selektirati testnu varijablu i premjestiti je u prozor *Factor*, a potom pod *Expected proportions* (χ^2 test) odgovarajućim redom unijeti očekivane frekvencije za pojedine kategorije testne varijable (svejedno je unesemo li frekvencije, postotke ili proporcije!).

JASP dopušta unos samo s jednom decimalom pa je potrebno zaokruživanje (ne zaboravite koristiti točku za decimalne vrijednosti, a ne zarez). Zbog toga je dobro unositi iznose u postotku umjesto u proporciji jer se tad zaokruživanjem na jednu decimalu gubi mnogo manje informacija.



Poredak vrijednosti je važan – on korespondira s uzlaznim (*ascending*) poretkom vrijednosti kategorija testne varijable. Prva vrijednost na listi korespondira s najmanjom vrijednošću kategorije testne varijable, a posljednja vrijednost korespondira s najvećom vrijednošću kategorije testne varijable. Sve se vrijednosti zbroje, a tada se svaka vrijednost podijeli s tim zbrojem kako bi se izračunala proporcija podataka koja se očekuju u korespondentnoj kategoriji. Npr. niz vrijednosti 3, 4, 5, 4 specificirat će očekivane proporcije 3/16, 4/16, 5/16 i 4/16.

Rezultati testa pokazuju da je distribucija varijable *stavpred* u populaciji članova stranke χ statistički značajno različita od normalne distribucije, na razini rizika 1% ($\chi^2 = 28,133$; $df = 4$; $p < ,001$):

Results

Multinomial Test

Multinomial Test

	χ^2	df	p
H ₀ (a)	28.133	4	< .001

Note. Chi-squared approximation may be incorrect

Descriptives

stavpred	Observed	Expected: H ₀ (a)
izrazito nepovoljan	12	4
ne povoljan	20	24
ni povoljan ni nepovoljan	40	45
povoljan	20	24
izrazito povoljan	8	4

6.3.3. Testiranje distribucije na zadani oblik

Pretpostavimo da želimo utvrditi je li dobna struktura članstva stranke X promijenjena u zadnjih godinu dana te da – s jedne strane – imamo empirijsku distribuciju varijable *dob* koja pokazuje sadašnju distribuciju članstva po dobnim skupinama dobivenu na uzorku, a s druge strane imamo sljedeće podatke o dobnjoj strukturi članstva prije godinu dana:

Dobna struktura članstva prije godinu dana

Dobna skupina	f
19 god. i manje	6
20–34 god.	22
35–49 god.	26
50–64 god.	36
65 god. i više	10

Testiranje ćemo provesti na isti način kao i kod testiranja normalnosti distribucije – nakon odabira testne varijable upisat ćemo očekivane frekvencije za pojedine kategorije:

Multinomial Test

Factor: dob

Counts: [empty]

Expected Counts: [empty]

Test Values

Equal proportions (multinomial test)

Expected proportions (χ^2 test)

Row #	H ₀ (a)
19 i manje	6
20-34	22
35-49	26
50-64	36

Results

Multinomial Test

Multinomial Test	χ^2	df	p
H ₀ (a)	4.857	4	0.302

dob	Observed	Expected: H ₀ (a)
19 i manje	4	6
20-34	19	22
35-49	24	26
50-64	37	36
65 i više	16	10

Hi-kvadrat test pokazuje da se distribucija varijable *dob* u populaciji članova stranke X ne razlikuje statistički značajno od distribucije članstva prema dobi od prije godinu dana ($\chi^2 = 4,857$; $df = 4$; $p = 0,302$).

Literatura

- Navarro, D.J., Foxcroft, D.R., i Faulkenberry, T.J. (2019). *Learning Statistics with JASP: A Tutorial for Psychology Students and Other Beginners*. Poglavlje 9.
URL: <https://tomfaulkenberry.github.io/JASPbook/chapters/chapter9.pdf>
- Goss-Sampson, M. A. (2019). *Statistička analiza u JASP programu: vodič za studente*. Poglavlja Bivarijatni hi-kvadrat test (str. 123-128) i Multinomni test i univarijatni hi-kvadrat test (str. 45).
URL: http://static.jasp-stats.org/Manuals/Statistic%cc%8cka_analiza_u_JASP_programu_v0.10.2.pdf