

9.

(Još neki)

nepara-
metrijski
testovi

TEME

1.	Mann-Whitneyev test
2.	Wilcoxonov test
3.	Kruskall-Wallisov test

U ovom su priručniku na drugim mjestima opisani neki neparametrijski postupci, a to su hi-kvadrat test te dva koeficijenta korelacije (Spearmanov i Kendalov tau-b). U ovom poglavlju fokusiramo se na testove razlika koji obično služe kao neparametrijske zamjene za t-testove i analizu varijance.

9.1. Mann-Whitneyev U-test

Radimo na primjerima u datoteci pod nazivom:

09_neparametrijski_testovi.sav

Mann-Whitneyev U-test neparametrijska je zamjena za t-test za nezavisne uzorke.

Koristimo ga u istim situacijama u kojima bismo koristili t-test za nezavisne uzorke (jedna kvantitativna varijabla kao zavisna i jedna nominalna dihotomna varijabla kao nezavisna), ali u slučajevima kad nisu zadovoljene pretpostavke za provedbu t-testa, odnosno ako:

- zavisna varijabla jest kvantitativna, ali nije normalno distribuirana (provjeriti npr. Shapiro-Wilkovim testom u modulu za deskriptivnu statistiku) ili
- zavisna varijabla zapravo nije kvantitativna, već ordinalna.

Primjer: Testirajte uz 5% rizika razlikuju li se muškarci i žene statistički značajno u važnosti koju pridaju rodnoj ravnopravnosti (varijabla VRIJED_RAVNO). Primijenite odgovarajući statistički test s obzirom na normalnost distribucije frekvencija varijable VRIJED_RAVNO.

Imati na umu da vrijednost 1 na varijabli VRIJED_RAVNO znači "uopće mi nije važno", a vrijednost 5 "izrazito mi je važno".

Prvo ćemo testirati normalnost distribucije varijable važnost rodne ravnopravnosti, kroz Descriptives / Shapiro-Wilk test.

Descriptive Statistics	
	vrijed_ravno
Valid	165
Missing	0
Shapiro-Wilk	0.899
P-value of Shapiro-Wilk	< .001

Iz dobivenih rezultata zaključujemo da distribucija frekvencija varijable VRIJED_RAVNO statistički značajno odstupa od normalne ($p < 0,001$), što znači da za rješavanje zadanog zadatka zapravo ne smijemo koristiti t-test, već njegovu neparametrijsku zamjenu: Mann-Whitneyev U-test.

Do ovog testa dolazimo na jednak način kao da želimo provesti t-test za nezavisne uzorke, dakle: *T-Tests / (Classical) Independent Samples T-Test*).

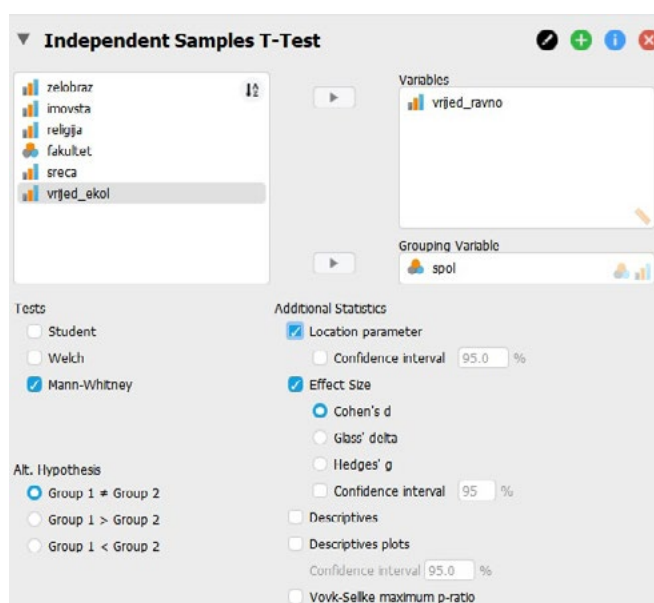
Odgovarajuće varijable premjestimo u odgovarajuće okvire: varijablu VRIJED_RAVNO u okvir *Variables*, a SPOL u okvir *Grouping Variable*.

No, pod *Tests* ćemo maknuti zadano označen Studentov t-test, a umjesto njega označiti Mann-Whitney.

Možemo još označiti i *Location parameter* čime ćemo dobiti tzv. Hodges-Lehmanovu procjenu, odnosno razliku medijana dviju grupa.

Označimo još i *Effect Size* čime ćemo dobiti rang-biserijalnu korelaciju (r_B) koju možemo interpretirati poput Pearsonova koeficijenta korelacije, a služi kao pokazatelj veličine učinka nezavisne na zavisnu varijablu u Mann-Whitneyevu U-testu.

Pripazimo također i na smjer hipoteze koju testiramo. U našem zadatku testiramo nultu hipotezu, što je statistički gledano jednako kao da testiramo neusmjerenu alternativnu hipotezu. Zato pod *Alt. Hypothesis* treba ostati označeno *Group 1 ≠ Group 2*.



Dobivamo sljedeće rezultate:

Independent Samples T-Test

Independent Samples T-Test					
	W	df	p	Hodges-Lehmann Estimate	Rank-Biserial Correlation
vrijed_ravno	2108.000		< .001	-1.000	-0.366

Note. For the Mann-Whitney test, effect size is given by the rank biserial correlation.
Note. Mann-Whitney U test.

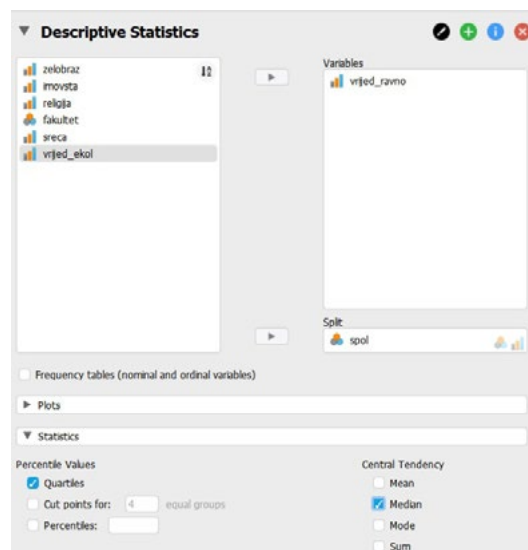
Iako u naslovu tablice piše *Independent Samples T-Test*, imajte na umu da ovo NISU rezultati t-testa! Na to nas upozoravaju i napomene ispod tablice.

P-vrijednost ($p < 0,001$) ukazuje na odbacivanje nulte hipoteze (uz 5% rizika) o nepostojanju razlike u rezultatima muškaraca i žena na varijabli VRIJED_RAVNO i prihvatanju alternativne.

Kod interpretacije rezultata neparametrijskih testova koji služe kao zamjene za bilo koji parametrijski test koji testira razliku u aritmetičkim sredinama, uobičajeno je kao deskriptivne pokazatelje navoditi MEDIJANE pojedinih grupa na nezavisnoj varijabli, a ne

njihove aritmetičke sredine. Kao pokazatelj disperzije možemo navesti prvi i treći kvartil.

Stoga, moramo još zatražiti određivanje medijana muškaraca i žena na varijabli VRIJED_RAVNO te odgovarajuće kvartile.



To ćemo učiniti ponovno kroz modul *Descriptives* na način da varijablu VRIJED_RAVNO uvedemo u okvir *Variables*, a varijablu SPOL u okvir *Split*, jer želimo odvojene rezultate za svaku kategoriju varijable SPOL.

Pod *Central Tendency* označimo Median, a pod *Percentile Values* označimo Quartiles.

Dobivamo rezultate koji će nam trebati za interpretaciju Mann-Whitneyeva U-testa (medijan na varijabli VRIJED_RAVNO za muškarce iznosi 3, a za žene 4):

Descriptive Statistics

	vrijed_ravno	
	muški	ženski
Valid	95	70
Missing	0	0
Median	3,000	4,000
25th percentile	2,000	3,000
50th percentile	3,000	4,000
75th percentile	4,000	5,000

Za ovaj zadatak možemo napisati sljedeće rješenje:

Primjenom Shapiro-Wilkova testa zaključili smo da varijabla važnost rodne ravnopravnosti u populaciji nije distribuirana normalno ($p < 0,001$), stoga smo hipotezu o razlici u rezultatima muškaraca i žena na toj varijabli testirali neparametrijskim Mann-Whitneyevim U-testom.

Rezultati Mann-Whitney U-testa pokazuju da se muškarci i žene statistički značajno razlikuju u rezultatima na varijabli važnost rodne ravnopravnosti ($W = 2108,000$; $p < 0,001$; $r_B = -0,366$), pri čemu žene pridaju rodnoj ravnopravnosti veću važnost ($Mdn = 4$; $Q_1 - Q_3: 3 - 5$) od muškaraca ($Mdn = 3$; $Q_1 - Q_3: 2 - 4$) te da je riječ o razlici umjerene veličine.

9.2. Wilcoxonov test

Radimo na primjerima u datoteci pod nazivom:
09_neparametrijski_testovi.sav

Wilcoxonov test (zove se još i Wilcoxonov test ekvivalentnih parova) neparametrijska je zamjena za t-test za zavisne uzorke.

Koristimo ga u istim situacijama u kojima bismo koristili t-test za zavisne uzorke (dvije kvantitativne varijable s istom skalom procjene mjerene na istom uzorku ili jedna kvantitativna varijabla mjerena dva puta / na dva zavisna uzorka), ali u slučajevima kad nisu zadovoljene pretpostavke za provedbu t-testa, odnosno ako:

- varijable jesu kvantitativne, ali jedna ili obje nisu normalno distribuirane (provjeriti npr. Shapiro-Wilkovim testom u modulu za deskriptivnu statistiku) ili
- varijable zapravo nisu kvantitativne, već ordinalne.

Primjer: Uz 5% rizika testirajte pridaje li se jednaka važnost vrijednosti rodne ravnopravnosti (varijabla VRIJED_RAVNO) i ekološkim vrijednostima (varijabla VRIJED_EKOL). Primijenite odgovarajući statistički test s obzirom na normalnost distribucije frekvencija varijabli VRIJED_RAVNO i VRIJED_EKOL.

Imati na umu da na obje varijable vrijednost 1 znači "uopće mi nije važno", a vrijednost 5 "izrazito mi je važno".

Prvo ćemo testirati normalnost distribucija objiju varijabli kroz Descriptives / Shapiro-Wilk test.

Descriptive Statistics

	vrijed_ekol	vrijed_ravno
Valid	165	165
Missing	0	0
Shapiro-Wilk	0.909	0.899
P-value of Shapiro-Wilk	< .001	< .001

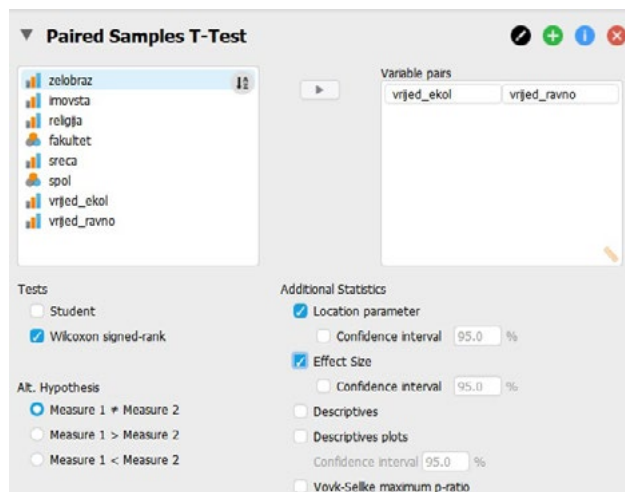
Iz dobivenih rezultata zaključujemo da distribucije frekvencija i varijable VRIJED_RAVNO ($p < 0,001$) i varijable VRIJED_EKOL ($p < 0,001$) statistički značajno odstupaju od normalne, što znači da za rješavanje ovog zadatka ne smijemo koristiti t-test za zavisne uzorke, već njegovu neparametrijsku zamjenu: Wilcoxonov test.

Do ovog testa dolazimo na jednak način kao da želimo provesti t-test za zavisne uzorke, dakle: *T-Tests / (Classical) Paired Samples T-Test*).

Obje varijable premjestimo u okvir *Variable pairs*.

No, pod *Tests* ćemo maknuti zadano označen Studentov t-test, a umjesto njega označiti Wilcoxonov test (*Wilcoxon signed-rank*). Možemo još označiti i *Location parameter* čime ćemo dobiti tzv. Hodges-Lehmanovu procjenu, odnosno razliku medijana dviju varijabli.

Označimo još i *Effect Size* čime ćemo dobiti rang-biserijalnu korelaciju (r_B) koju možemo interpretirati poput Pearsonova koeficijenta korelacije, a služi kao pokazatelj veličine učinka. Pripazimo također i na smjer hipoteze koju testiramo. U našem zadatku testiramo nultu (ili neusmjerenu alternativnu) hipotezu.



Dobivamo sljedeće rezultate:

Paired Samples T-Test

Measure 1	Measure 2	W	df	p	Hodges-Lehmann Estimate	Rank-Biserial Correlation
vrjed_ekol	- vrjed_ravno	3575.000		0.442	-5.832e-5	-0.077

Note. Wilcoxon signed-rank test.

Iako u naslovu tablice piše *Paired Samples T-Test*, imajte na umu da ovo NISU rezultati t-testa! Na to nas upozorava i napomena ispod tablice.

P-vrijednost ($p = 0,442$) ukazuje na prihvaćanje nulte hipoteze (uz 5% rizika) o nepostojanju razlike o važnost vrijednosti rodne ravnopravnosti i ekoloških vrijednosti.

Kako smo već naveli, kod interpretacije rezultata neparametrijskih testova koji služe kao zamjene za bilo koji parametrijski test koji testira razliku u aritmetičkim sredinama, uobičajeno je kao deskriptivne pokazatelje navoditi MEDIJANE pojedinih varijabli, a ne njihove aritmetičke sredine.

Stoga, moramo još zatražiti određivanje medijana varijabli VRIJED_RAVNO i VRIJED_EKOL i njihovih kvartila (prvog i trećeg).

To ćemo učiniti ponovno kroz modul *Descriptives* na način da varijable VRIJED_RAVNO i VRIJED_EKOL uvedemo u okvir *Variables* te pod *Central Tendency* označimo Median.

Descriptive Statistics		
	vrjed_ravno	vrjed_ekol
Valid	165	165
Missing	0	0
Median	4.000	3.000
25th percentile	2.000	2.000
50th percentile	4.000	3.000
75th percentile	4.000	4.000

Iako vidimo razliku u medijanama ovih dviju varijabli na uzorku, ne možemo na temelju Wilcoxonova testa govoriti o statistički značajnoj razlici u rezultatima na ove dvije varijable.

Za ovaj zadatak možemo napisati sljedeće rješenje:

Primjenom Shapiro-Wilkovih testova zaključili smo da niti varijabla važnost vrijednosti rodne ravnopravnosti ($p < 0,001$) niti varijabla važnost ekoloških vrijednosti ($p < 0,001$) u populaciji nisu distribuirane normalno, stoga smo hipotezu o razlici važnosti vrijednosti rodne ravnopravnosti i ekoloških vrijednosti testirali neparametrijskim Wilcoxonovim testom.

Rezultati Wilcoxonova testa pokazuju da se važnost rodne ravnopravnosti ($Mdn = 4$; $Q_1 - Q_3: 2 - 4$) i važnost ekoloških vrijednosti ($Mdn = 3$; $Q_1 - Q_3: 2 - 4$) ne razlikuju statistički značajno ($W = 3575,000$; $p = 0,442$).

**Napomena: s obzirom na to da razlika nije statistički značajna, nema potrebe prikazati iznos koeficijenta rang-biserijalne korelacije (r_B). Da razlika jest značajna, dodali bismo ga u zagradu s rezultatima testa i interpretirali.*

9.3. Kruskall-Wallisov test

Radimo na primjerima u datoteci pod nazivom:
09_neparametrijski_testovi.sav

Kruskall-Wallisov test neparametrijska je zamjena za analizu varijance.

Koristimo ga u istim situacijama u kojima bismo koristili analizu varijance (jedna kvantitativna varijabla kao zavisna i jedna nominalna višetomna varijabla kao nezavisna), ali u slučajevima kad nisu zadovoljene pretpostavke za provedbu ANOVA-e, odnosno ako:

- zavisna varijabla jest kvantitativna, ali nije normalno distribuirana (provjeriti npr. Shapiro-Wilkovim testom u modulu za deskriptivnu statistiku) ili
- zavisna varijabla zapravo nije kvantitativna, već ordinalna.

Primjer: Testirajte uz 5% rizika razlikuju li se studenti statistički značajno u važnosti koju pridaju rodnoj ravnopravnosti (varijabla VRIJED_RAVNO) s obzirom na područje studiranja (varijabla FAKULTET). Primijenite odgovarajući statistički test s obzirom na normalnost distribucije frekvencija varijable VRIJED_RAVNO.

Imati na umu da vrijednost 1 na varijabli VRIJED_RAVNO znači "uopće mi nije važno", a vrijednost 5 "izrazito mi je važno", dok su vrijednosti na varijabli FAKULTET sljedeće:

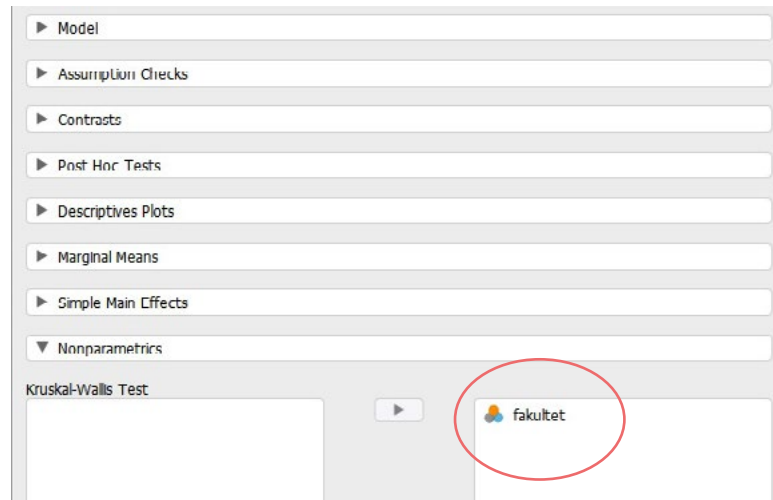
- 1 = "društveno-humanistički"
- 2 = "tehnički"
- 3 = "medicinsko-biotehnološki"

Trebali bismo prvo testirati normalnost distribucije varijable važnost rodne ravnopravnosti, kroz Descriptives / Shapiro-Wilk test, no kako smo to već napravili u prethodnim poglavljima, znamo da ona nije normalno distribuirana ($p < 0,001$), što znači da za rješavanje zadanog zadatka zapravo ne smijemo koristiti ANOVA-u, već njegovu neparametrijsku zamjenu: Kruskall-Wallisov test.

Do ovog testa dolazimo na jednak način kao da želimo provesti analizu varijance, dakle: ANOVA / (Classical) ANOVA.

Odgovarajuće varijable premjestimo u odgovarajuće okvire: varijablu VRIJED_RAVNO u okvir *Dependent Variable*, a FAKULTET u okvir *Fixed Factors*.

Da bismo odabrali provedbu Kruskall-Wallisova testa, idemo na zadnju opciju u popisu: *Nonparametrics* te nezavisnu varijablu (ovdje: FAKULTET) premjestimo u desni okvir.



U prozoru s rezultatima u jednoj ćemo tablici dobiti rezultate analize varijance (koji nas trenutno ne interesiraju), a u drugoj rezultate Kruskal-Wallisova testa.

Kruskal-Wallis Test

Kruskal-Wallis Test

Factor	Statistic	df	p
fakultet	16.077	2	< .001

Napomena: *Statistic* čija se vrijednost testira jest hi-kvadrat testovni statistik (ovdje: $\chi^2 = 16,077$).

Iznos p-vrijednosti ($p < 0,001$) ukazuje na odbacivanje nulte hipoteze i prihvaćanje alternativne.

Zasad, dakle, samo možemo zaključiti da se studenti međusobno statistički značajno razlikuju u važnosti koju pridaju rodnoj ravnopravnosti s obzirom na područje studiranja.

No, zasad još ne znamo:

- A) koje se skupine prema području studiranja međusobno razlikuju,
- B) na koji se način razlikuju (kome je važnija, a kome manja važna rodna ravnopravnost).

Prvu informaciju bismo u analizi varijance dobili provedbom odgovarajućih *post hoc* testova, a drugu interpretacijom aritmetičkih sredina na varijabli VRIJED_RAVNO svake od grupa na varijabli FAKULTET.

No, kod Kruskall-Wallisova testa situacija je nešto drugačija.

A) *Post hoc* testovi ne postoje, barem ne inkorporirani u samu proceduru Kruskall-Wallisova testa. Morat ćemo ih zamijeniti provedbom odvojenih Mann-Whitneyevih U-testova za sve parove grupa na nezavisnoj varijabli (FAKULTET), ali još dodatno uz određenu korekciju za višestruke usporedbe – Bonferronijevu korekciju.

B) Kao pokazatelj centralne tendencije, umjesto aritmetičkih sredina koje koristimo u ANOVA-i, ovdje ćemo zatražiti određivanje MEDIJANA i kvartila varijable VRIJED_RAVNO za svaku od grupa iz varijable FAKULTET te ćemo njih koristiti za utvrđivanje kome je važnija, a kome manja važna rodna ravnopravnost.

Morat ćemo provesti tri Mann-Whitneyeva U-testa kako bismo međusobno usporedili sve tri skupine varijable FAKULTET na varijabli VRIJED_RAVNO:

1. društveno-humanistički i tehnički
2. društveno-humanistički i medicinsko-biotehnološki
3. medicinsko-biotehnološki i tehnički

S obzirom na to da je uvijek logičnije prvo pogledati deskripciju varijabli, odredit ćemo prvo medijane na varijabli VRIJED_RAVNO za svaku od grupa iz varijable FAKULTET.

Descriptive Statistics

	vrijed_ravno		
	društ. hum.	tehnički	med. biotech.
Valid	55	55	55
Missing	0	0	0
Median	3.000	4.000	3.000
25th percentile	2.000	3.000	2.000
50th percentile	3.000	4.000	3.000
75th percentile	4.000	5.000	4.000

Idemo u Descriptives, varijablu VRIJED_RAVNO premjestimo u okvir *Variables*, varijablu FAKULTET u okvir *Split*, a pod *Central Tendency* označimo Median. Dobivamo sljedeće rezultate iz kojih je vidljivo da medijan važnosti rodne ravnopravnosti za studente društveno-humanističkih i medicinsko-biotehnoloških studenata iznosi 3, dok za one s tehničkih fakulteta iznosi 4.

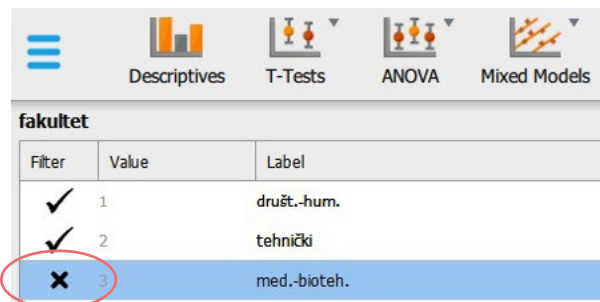
Mann-Whitneyevi U-testovi s Bonferronijevom korekcijom kao zamjena za *post hoc* testove u Kruskall-Wallisovu testu

Potom prelazimo na testiranje razlika u parovima skupina fakulteta: provodimo **Mann-Whitneyeve U-testove s Bonferronijevom korekcijom kao zamjene za *post hoc* testove u Kruskall-Wallisovu testu.**

Da bismo proveli tri Mann-Whitneyeva U-testa, morat ćemo koristiti filtere.

1. Neka nam prva usporedba bude za studente društveno-humanističkih i tehničkih fakulteta.

Uz pomoć filtera isključimo kategoriju "medicinsko-biotehnološki" na varijabli FAKULTET, a ostavljamo one skupine koje uspoređujemo.



Potom idemo u: *T-Tests / (Classical) Independent Samples T-Test*. Odgovarajuće varijable premjestimo u odgovarajuće okvire: varijablu VRIJED_RAVNO u okvir *Variables*, a FAKULTET u okvir *Grouping Variable*.

Pod *Tests* maknemo Studentov t-test, a označimo Mann-Whitney. Označimo *Location parameter* i *Effect Size*. Pazite da testirate neusmjereno.

Dobivamo sljedeće rezultate.

Independent Samples T-Test					
	W	df	p	Hodges-Lehmann Estimate	Rank-Biserial Correlation
vrijed_ravno	1143.000		0.022	-1.000	-0.244

Note. For the Mann-Whitney test, effect size is given by the rank biserial correlation.
Note. Mann-Whitney U test.

Bonferronijeva korekcija za višestruke usporedbe

Da bismo proveli još i Bonferronijevu korekciju za višestruke usporedbe, potrebno je **odabranu razinu rizika** uz koju testiramo hipotezu (ovdje je to 5%, odnosno $\alpha = 0,05$) **podijeliti s brojem usporedbi** koje ćemo napraviti, a ovdje je to 3.

Bonferronijevu korekciju učiniti ćemo "ručno": $0,05 / 3 = 0,017$.

To znači da **p-vrijednost u svakom provedenom Mann-Whitneyevu U-testu mora biti MANJA OD 0,017 da bi razlika** između dviju grupa fakulteta na varijabli važnost rodne ravnopravnosti **bila statistički značajna uz 5% rizika!**

U ovom bi slučaju p-vrijednost za razliku u važnosti rodne ravnopravnosti između studenata društveno-humanističkih i tehničkih fakulteta ($p = 0,022$) bez Bonferronijeve korekcije bila statistički značajna uz 5% rizika (jer je manja od 0,05), no uz Bonferronijevu korekciju NIJE statistički značajna uz 5% rizika (jer nije manja od 0,017).

Stoga na temelju ovog testa zaključujemo da se studenti društveno-humanističkih i tehničkih fakulteta međusobno statistički značajno **ne razlikuju u važnosti rodne ravnopravnosti**.

2. Druga usporedba bit će za studente društveno-humanističkih i medicinsko-biotehnoških fakulteta. Namjestimo filter:



Provedemo Mann-Whitneyev U-test:

Independent Samples T-Test

Independent Samples T-Test					
	W	df	p	Hodges-Lehmann Estimate	Rank-Biserial Correlation
vrijed_ravno	1781.000		0.100	7.446e-5	0.178

Note. For the Mann-Whitney test, effect size is given by the rank biserial correlation.
Note. Mann-Whitney U test.

Dobivena p-vrijednost ($p = 0,100$) veća je od kritične uz Bonferronijevu korekciju ($0,017$), stoga zaključujemo da razlika u važnosti rodne ravnopravnosti između studenata društveno-humanističkih i medicinsko-biotehnoloških fakulteta nije statistički značajna uz 5% rizika.

3. Treća usporedba bit će za studente medicinsko-biotehnoloških i tehničkih fakulteta. Namjestimo filter:

fakultet		
Filter	Value	Label
<input checked="" type="checkbox"/>		društ.-hum.]
<input checked="" type="checkbox"/>	2	tehnički
<input checked="" type="checkbox"/>	3	med. biotech.

Provedemo Mann-Whitneyev U-test:

Independent Samples T-Test

Independent Samples T-Test					
	W	df	p	Hodges-Lehmann Estimate	Rank-Biserial Correlation
vrijed_ravno	2163.500		< .001	1.000	0.430

Note. For the Mann-Whitney test, effect size is given by the rank biserial correlation.
Note. Mann-Whitney U test.

Dobivena p-vrijednost ($p < 0,001$) je manja od kritične uz Bonferronijevu korekciju ($0,017$), stoga zaključujemo da razlika u važnosti rodne ravnopravnosti između studenata medicinsko-biotehnoloških i tehničkih fakulteta jest statistički značajna uz 5% rizika.

Gotovi smo s procedurom i možemo napisati **rješenje zadatka i interpretaciju dobivenih nalaza**.

Primjenom Shapiro-Wilkova testa zaključili smo da varijabla važnost rodne ravnopravnosti u populaciji nije distribuirana normalno ($p < 0,001$), stoga smo hipotezu o razlici važnosti rodne ravnopravnosti s obzirom na područje studiranja testirali neparametrijskim Kruskal-Wallisovim testom.

Rezultati Kruskal-Wallisova testa pokazuju da se važnost rodne ravnopravnosti statistički značajno razlikuje između studenata triju skupina fakulteta ($\chi^2 = 16,077$; $df = 2$; $p < 0,001$). Za usporedbu pojedinih grupa korišteni su Mann-Whitneyevi U-testovi s Bonferronijevom korekcijom za razinu statističke značajnosti od $\alpha = 0,017$ ($0,05/3$). Studentima s tehničkih fakulteta rodna je ravnopravnost važnija ($Mdn = 4$; $Q_1 - Q_3: 3 - 5$) nego studentima s medicinsko-biotehnoloških fakulteta ($Mdn = 3$; $Q_1 - Q_3: 2 - 4$) ($W = 2163,500$; $p < 0,001$). Između ostalih parova skupina fakulteta nije utvrđena statistički značajna razlika u važnosti rodne ravnopravnosti.

Literatura

- Goss-Sampson, M. A. (2019). *Statistička analiza u JASP programu: vodič za studente*.
URL: http://static.jasp-stats.org/Manuals/Statistic%cc%8cka_analiza_u_JASP_programu_v0.10.2.pdf
Man-Vitnejev test (str. 50-51)
Vilkoksonov test ekvivalentnih parova (str. 55-56)