SINIŠA OPIĆ

UDK: 510.42
Pregledni članak / Review
Rukopis prihvaćen za tisak: 28. 6. 2023.
DOI: https://dx.doi.org/10.21857/moxpjhz4dm

Faktorijalni dizajni *Analize varijance* (*ANOVA*) u tipovima sume kvadrata 1, 2, 3 (*type one SS, type two SS, type three SS*) u nebalansiranim dizajnima – princip marginalnosti

Sažetak

Cilj je ovoga rada prikazati specifičnosti korištenja, ali i ograničenja faktorijalnih dizajna u kvantitativnim diferencijalnim nacrtima u području društvenih znanosti. Prikazane su specifičnosti dvosmjernih (*two way*) i višesmjernih (*n-way*) faktorijalnih dizajna u *Analizi varijance* (*ANOVA*) s obzirom na specifičnosti vrsta suma kvadrata (*type one SS, type two SS, type three sum of squares*). Testiran je nebalansiran model 2*4 dvosmjerne *ANOVA*-e s pomoću Tipa 3 sume kvadrata (*type three SS*) u kojem je izostao glavni učinak (*main effect*). Zbog principa marginalnosti s pomoću Tipa 3 sume kvadrata (*type three SS*) ne smiju se interpretirati glavni učinci, već se za glavne učinke nameće potreba partikularnog pristupa, jednosmjerne analize varijance (*one way ANOVA*) te dvosmjerne analize varijance (faktorijalni dizajn), ali ne s pomoću Tipa 3 sume kvadrata.

Ključne riječi: *ANOVA;* interakcijski učinci; marginalnost*;* suma kvadrata tipa 1, tipa 2, tipa 3

Uvod

U diferencijalnom nacrtu, univarijatnom pristupu, najčešće korišteni parametrijski testovi za testiranje razlika aritmetičkih sredina (subuzoraka) na zavisnoj varijabli jesu t-test (kad su dvije aritmetičke sredine) te analiza varijance (*ANOVA*) za testiranje više aritmetičkih sredina. Smisao računanja analize varijance jest omjer između međugrupne i unutargrupne varijance. Međugrupna varijanca je udaljenost svakoga zasebnog rezultata svih subuzoraka od zajedničke/ukupne aritmetičke sredine, dok je unutargrupna varijanca zapravo zaseban varijabilitet svakog subuzorka – sistemska ili rezidualna. Što je veći omjer međugrupne i untargrupne varijance, veća je vjerojatnost izostavljanja nul-hipoteze.

$$F_{omjer} = \frac{\text{Međugrupna varijanca (Mean square beetwen)}}{\text{Unutargrupna varijanca (Mean square within)}} .$$

U grafičkom prostoru to se vidi s pomoću prostora preklapanja distribucija (*intersekcija*), odnosno što su intersekcije subuzoraka veće, veća je vjerojatnost da pripadaju istoj populaciji, tj. vjerojatnost potvrđivanja nul-hipoteze.

Suma kvadrata (*Sum of squares*) zapravo je mjera varijacije, tj. ona je numerator formule za varijancu:

$$\sigma^2 = SS/N\text{-}1;\ SS = \Sigma(x\text{-}\bar{x})^2.$$

Ukupni varijabilitet SSt (*sum of squares total*) zapravo je umnožak zajedničke (*grand*) varijance i ukupnog broja ispitanika (točnije N-1), gdje je *grand* varijanca – ukupna varijanca svih rezultata:

$$SSt = \sigma^2(N\text{-}1).$$

U dvosmjernoj analizi varijance (*two way ANOVA*) dva su faktora, tj. tada se ukupno varijabilitet sastoji od objašnjene varijance modelom, tj. varijance prvog faktora ($\sigma_{f1}{}^2$) i drugog faktora ($\sigma_{f2}{}^2$) te varijance objašnjene interakcijom f1 i f2:

$$MS_{f1} = \frac{SS_{f1}}{df_{f1}} \quad \text{gdje je } df_{f1} = k\text{-}1.$$

$$MS_{f2} = \frac{SS_{f2}}{df_{f2}} \quad \text{gdje je } df_{f2} = k\text{-}1.$$

Tada je F omjer $F_{f1} = \frac{MS_{f1}}{MSE}$, gdje je MSE = $\frac{SSE}{dfE}$; SSE je rezidualna suma kvadrata odnosno varijanca koja nije objašnjena modelom (neobjašnjen varijabilitet).

Analogno tome slijedi:

$$F_{f2} = \frac{MS_{f2}}{MSE}$$

Interakcijski učinak f1*f2 definiran je sljedećom formulom:

$$MS_{f1*f2} = \frac{SS_{f1*f2}}{MSE}$$

Tada je:

$$F_{f1*f2} = \frac{MS_{f1*f2}}{MSE}$$

Faktorijalni dizajni

Faktorijalni dizajni (*two way, three way ili n-way*) u univarijatnom pristupu analize varijance (*ANOVA*) ili multivatijatnom – npr., multivarijatna analiza varijance (*MANOVA*) i multivarijatna analiza kovarijance (*MANCOVA*) – iznimno su vrijedni statistički postupci za determiniranje interakcijskih učinaka više nezavisnih varijabli na zavisnu varijablu (ili u multivarijatnom prostoru na zavisne varijable). Odnosi u društvenim znanostima iznimno su kompleksni te je vrijedno istražiti složene odnose više varijabli, tj. njihove interakcijske učinke.

Kad postoji jedna nezavisnu varijabla (n kategorija) i jedna zavisna varijabla, tada je riječ o *jednosmjernoj analizi varijance* (*ANOVA*), a kad su dvije nezavisne varijable i testira se njihov interakcijski učinak na zavisnu varijablu, tada je riječ o dvosmjernoj analizi varijance. Ovisno o broju kategorija dviju nezavisnih varijabli, imamo npr., 2*3 dvosmjernu analizu varijance (*two way ANOVA*). Prva nezavisna varijabla ima dvije kategorije, dok druga nezavisna ima tri kategorije (subuzoraka) ili, npr., 4*2 (prva nezavisna varijabla ima četiri kategorije, a druga dvije) i sl. Dakle, dvosmjerne testove (*two way*) označavamo ovako: $n_{kat;x^*}$ $n_{kat;y}$.

Najčešće se testiraju interakcijski učinci dviju nezavisnih varijabli (*two way*) na jednu zavisnu varijablu, no mogu biti i tri nezavisne varijable (*three way ANOVA*) ili više nezavisnih varijabli (*n way ANOVA*). Ako je više od jedne zavisne varijable, tada je riječ o *n-way* multivarijatnoj analizi varijance (*MANOVA*) ili kad je riječ o kontroli kovarijata *n-way* analizi kovarijance (*ANCOVA*). Svi *n way* >1 modeli nazivaju se faktorijalnim dizajnima.

Jedan je od preduvjeta za testove u području diferencijalnih nacrta (t-test, *ANOVA, MANOVA…*) – osim preduvjeta o normalnosti distribucije, homogenosti varijanci subuzoraka (*homoscedascitet*) duž zavisne varijable, zavisna varijabla kvantitativno najmanja na intervalnoj skali (ovaj uvjet se učestalo zanemaruje) – također i ista veličina uzoraka koji se uspoređuju. To je preduvjet i za jednosmjerne (*one way*) testove, ali velik utjecaj ima i u faktorijalnim dizajnima. Razlika u veličini subuzoraka (n) koji se uspoređuju razvidna je i u razlikama između deskriptivnih i procijenjenih marginalnih aritmetičkih sredina (*estimated marginal means*). To se u faktorijalnom modelu kontrolira s pomoću Tipa 3 sume kvadrata (*type three SS*) koji se odnosi kao da su subuzorci iste veličine (nisu različiti) pa se zbog toga koristi harmonijska aritmetička sredina. Stoga, kad postoje različite veličine subuzoraka, bolje je koristiti se Tipom 3 suma kvadrata (*type three SS*) koji je zapravo prilagođen sumi kvadrata svakoga zasebnog glavnog učinka (*main effect*), drugim glavnim učincima (*main effect*) te ujedno i interakcijskom učinku (ili učincima).

Zanimljivo je da navedene vrijednosti tri sume kvadrata (*type 1 SS, type 2 SS, type 3 SS*) u dvosmjernoj analizi varijance (*two way ANOVA*) ne daju iste rezultate u uvjetima kad postoje nebalansirane/neujednačene *(non-orthogonal)* vrijednosti podatka (varijanci, veličine n, izražene asimetričnosti, *skewness*).

Tip 1 suma kvadrata *(type one SS)* možemo nazvati i *sekvencijalnom (inkrementalnom)* jer su varijance varijable složene redom, odnosno poredak varijabli u modelu čini razliku u vrijednostima sume kvadrata. Tip 1 suma kvadrata *(type one SS)* svakog učinka određuje se oduzimanjem predviđenog zbroja kvadrata s učinkom u modelu od predviđenog zbroja kvadrata za prethodni model, ali ne uključujući učinak. U analizi varijance (*ANOVA)* Tip 1 suma kvadrata koristi se kad je riječ o podjednakim veličinama subuzoraka (nezavisnih varijabli) te je također prikladna *u polinomial* regresiji, gdje se niže rangirani učinak (*lower order – effect)* uvodi prije nego više rangiranog učinka (*higher order effect)*. Zapravo, glavni je nedostatak činjenica da Tip 1 suma kvadrata *designa* odnosno pojedinačnog učinka, ovisi o redoslijedu u kojem je učinak stavljen u model.

Tip 2 suma kvadrata *(type two SS) ili hijerarhijska* koristi se u modelima u kojima nema interakcije nezavisnih varijabli. Riječ je o suprotnosti od Tipa 1 sume kvadrata (*type one SS)* jer se kod Tipa 2 sume kvadrata *(type two SS)* dobije čisti učinak (effect), odnosno kontrolira se svaki drugi učinak (effect). Tip 2 sume kvadrata *(type two SS)* je ekvivalent Tipu 1 sume kvadrata za zadnji prediktor u modelu. To obuhvaća činjenicu da učinak ne ovisi o redoslijedu u modelu, već je nepromjenjiv. Tip 2 suma kvadrata može se koristiti i u faktorijalnim dizajnima (kad postoji interakcijski učinak), ali samo za determiniranje glavnih učinaka, a nipošto interakcijskih jer se varijanca glavnih učinaka (*main effect)* ne gubi u interakcijskim odnosima.

Tip 3 suma kvadrata (*type three SS*) je parcijalna, u osnovi slična Tipu 2 sumi kvadrata, ali prednost je u činjenici da je prikladna za korištenje kad je riječ o disproporciji veličine subuzorka nezavisne varijable, no nije prikladna kad postoje vrijednosti koje nedostaju (*missing files)*. Za razliku od Tipa 2 sume kvadrata, Tip 3 suma kvadrata (*type three SS*) prikladna je za modele s interakcijama jer su glavni učinci prilagođeni za interakcije. Postoje i tipovi 4, 5, 6 sume kvadrata, no smatra se da je Tip 3 suma kvadrata (*type three SS)* prikladniji (Foster, Barkus, Yavorsky, 2006.; Milliken and Johnson (2009.). Iako je u statističkim programima, prema *defaultu*, u faktorijalnoj analizi varijance (*ANOVA*) za postavljeni *Tip* 3 sume kvadrata, postoje i oprečna mišljenja. Tako, npr., Langsrud (2003.) upućuje na ograničenja navedenog Tipa 3 sume kvadrata (*type three SS),* gdje se parametri za svaki izraz zbrajaju na nulu kada se zbroje preko bilo kojeg indeksa:

$$\sum_{i=1}^{a} \alpha_i = 0 \ , \quad \sum_{j=1}^{b} \beta_j = 0 \ , \quad \sum_{i=1}^{a} \gamma_{ij} = 0 \ \ \forall \ j \ , \quad \sum_{j=1}^{b} \gamma_{ij} = 0 \ \ \forall \ i$$

Slično tome, Nelder i Lane (1995.), kritizirajući Tip 3 sumu kvadrata (*type three SS*), ističu da korespondira s nezanimljivom hipotezom – testiranje glavnih učinaka u prisustvu interakcija. Analizirajući komparativne studije, Lewsey et al. (1997., 2001.) upućuje da prosječno Tip 2 suma kvadrata (*type two SS*) producira jaču snagu testa.

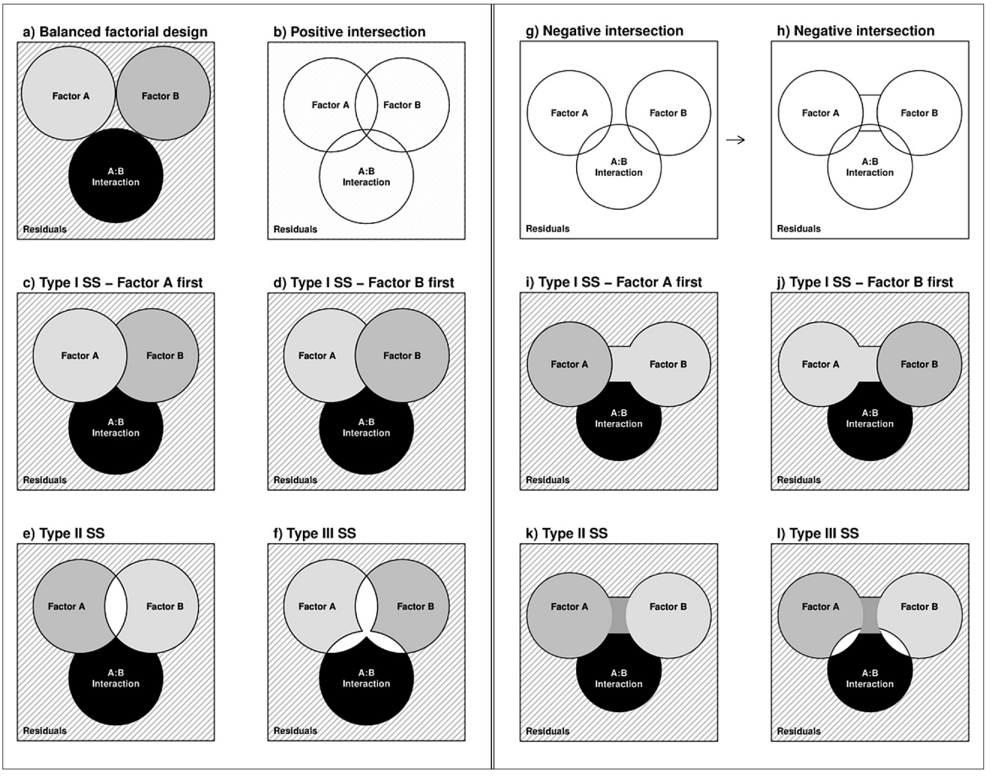
Štoviše, takav pristup bez dodatnih vrijednosti Kempthorne (1975.) naziva glupošću. No, s takvim pristupom teško se složiti jer je 1934. godine Yates postavio sva tri modela/tipa sume kvadrata u faktorijalnoj analizi varijance i jasno definirao specifičnosti korištenja. Takav pristup, uz dodatna objašnjenja (npr., Herr, 1986.) postavio je temelj u kvantitativnoj metodologiji i za korištenje Tipa 3 sume kvadrata (*type three SS*) u faktorijalnoj analizi varijance (*ANOVA*).

Aho (2013.) zorno prikazuje komparaciju svih triju suma kvadrata u faktorijalnoj analizi varijance (*ANOVA*) (Slika 1.).

	računanje	uvjetovanost
Tip 1 suma kvadrata (type one SS) - sekvencijalna		
A	$SS_A = SSE(\mu) - SSE(A)$	$SS(A \mid \mu)$
B	$SS_B = SSE(A) - SSE(A + B)$	$SS(B \mid A)$
$A{:}B$	$SS_{A:B} = SSE(A + B) - SSE(A^* B)$	$SS(A{:}B \mid A, B)$
Error	$SSE(A^* B)$	
Tip 2 suma kvadrata (type two SS) - parcijalna		
A	$SS_A = SSE(B) - SSE(A + B)$	$SS(A \mid B)$
B	$SS_B = SSE(A) - SSE(A + B)$	$SS(B \mid A)$
$A{:}B$	$SS_{A:B} = SSE(A + B) - SSE(A^*B)$	$SS(A{:}B \mid A, B)$
Error	$SSE(A^* B)$	
Tip 3 suma kvadrata (type three SS) - marginalna		
A	$SS_A = SSE(B + A{:}B) - SSE(A^*B)$	$SS(A \mid B, A{:}B)$
B	$SS_B = SSE(A + A{:}B) - SSE(A^*B)$	$SS(B \mid A, A{:}B)$
$A{:}B$	$SS_{A:B} = SSE(A + B) - SSE(A^*B)$	$SS(A{:}B \mid A, B)$
Error	$SSE(A^*B)$	

Slika 1. *Komparacija svih triju sume kvadrata (type 1 SS, type 2 SS, type 3 SS) u analizi varijance (ANOVA)*

Jedan od izazova u Tipu 3 sume kvadrata (*type three SS*) u faktorijalnim dizajnima je nebalansirani (*unbalanced)* dizajn kad su prisutne disproporcije veličine subuzoraka (n) i/ili prazne ćelije (*missing value*).



Slika 2. *Intersekcije u sumama kvadrata; balansirani i nebalansirani dizajni*
(Izvor: https://www.flutterbys.com.au/stats/tut/tut7.6a.html)

Tip 3 suma kvadrata *(type three SS)* daje konzervativne procjene te su šanse za procjenjivanje glavnih učinaka ravne nuli, dok rezultati Tipa 2 sume kvadrata (*type two SS)* jako variraju. Ako nema interakcije, rezultati pokazuju da su prednosti Tipa 2 sume kvadrata (*type two SS)* u odnosu na Tip 3 sume kvadrata (*type three SS)* relativno male. Možda je najbolje primijeniti Tip 3 sumu kvadrata (*type three SS)* kod neuravnoteženih skupova podataka kad god je interakcija zamisliva (Landsheer i van den Wittenboer, 2015.).

No, kad je riječ o testiranju glavnih učinaka *(main effect) u* faktorijalnim modelima, treba biti vrlo oprezan u zaključcima. Zapravo, glavne efekte u prisustvu interakcija ne treba interpretirati. Stewart-Oaten (1995., str. 2007.) našalio se s „Tipom 3 sume kvadrata (*type three SS)* jer je 'očito' najbolja za glavne efekte kada uopće nema smisla testirati glavne efekte". Slično tome, Smith and Cribbie (2014.) upućuju da Tip 3 suma kvadrata (*type three SS*) ima jaču statističku snagu, no suspektni su pouzdani glavni učinci.

Zapravo, u prisustvu interakcijskih učinaka uopće ne treba testirati glavne efekte zbog *Principa marginalnosti*, tj. determiniranje pojedinačnih zasebnih učinaka pojedinih varijabli marginalno je u prisustvu interakcijskih učinaka. To treba biti obvezno načelo u interpretaciji faktorijalnih modela u analizi varijance (*ANOVA*). Usporedo s tim Hector von Felten i Schmid (2010.) upućuju da nema mnogo smisla interpretirati glavne efekte (*main effect*) u prisustvu statistički značajnih interakcija.

Tip 1 sumu kvadrata (*type one SS*), Tip 2 sumu kvadrata (*type two SS*) i Tip 3 sumu kvadrata (*type three SS*) možemo matematički prikazati (Slika 3.):

	Tip 1 suma kvadrata (Type one SS)	Tip 2 suma kvadrata (Type two SS)	Tip 3 suma kvadrata (Type three SS)
A	SS(A)=R(1)-R(A)	SS(A|B)=R(B)-R(A,B)	SS(A|B,AB)=R(B,AB)-R(A,B,AB)
B	SS(B|A)=R(A)-R(A,B)	SS(B|A)=R(A)-R(A,B)	SS(B|A,AB)=R(A,AB)-R(A,B,AB)
AB	SS(AB|A,B)=R(A,B)-R(A,B,AB)	SS(AB|A,B)=R(A,B)-R(A,B,AB)	SS(AB|A,B)=R(A,B)-R(A,B,AB)

Slika 3. *Matematički prikaz računanja tri sume kvadrata*
Izvor: https://www.utstat.utoronto.ca/reid/sta442 f/2009/ typeSS.pdf

Kad se provodi dvosmjerna analiza varijance (*two way ANOVA*), razlikuju se sljedeće solucije:

- Nijedan faktor (A; B) nije statistički značajan; nema interakcijskog učinka A x B.
- Nijedan faktor (A; B) nije statistički značajan; postoji interakcijski učinak A x B.
- Faktor A statistički je značajan – nema interakcijskog učinka A x B.
- Faktor B statistički je značajan – nema interakcijskog učinka A x B.
- Faktor A i faktor B statistički su značajni – nema interakcijskog učinka A x B.
- Faktor A i faktor B statistički su značajni – postoji interakcijski učinak A x B.

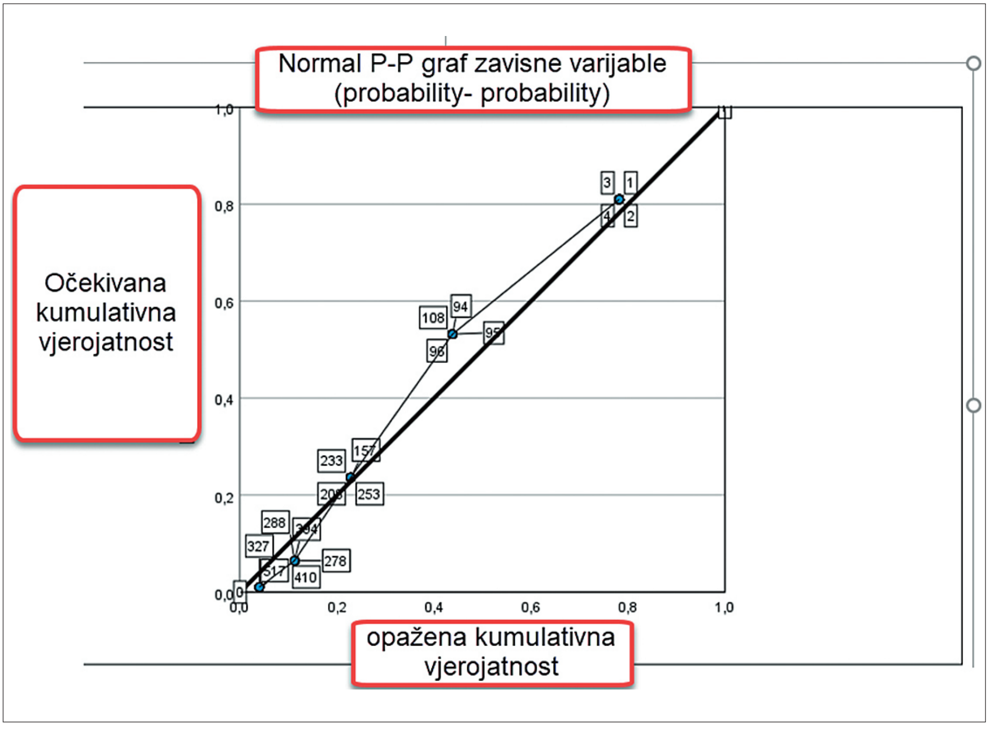
Navedene sume kvadrata u faktorijalnoj analizi varijance (*ANOVA*) povezuju se u pogreške prvoga i drugog reda. Pogreška prvog reda (*type one error*) je odbacivanje H0 kad je ona istinita (*false positive*), dok je pogreška drugog reda (*type two error*) obrnuta, tj. ne odbacuje se H0 kad je ona lažna *(false negative)*. No, pogreška trećeg reda (*type three error*) može se promatrati s nekoliko područja:

1. Statistički značajna razlika između subuzoraka je točna, no netočna je interpretacija smjera razlika.
2. Davanje točnog odgovora na netočno pitanje.
3. Pripisivanje „loših rezultata" slabosti provedenog tretmana, a zapravo je riječ o tretmanu koji nije dobro proveden.
4. Točno je odbacivanje H0, ali zbog neopravdanog razloga (prilagođeno prema Vogt i Johnson, 2016.).

Empirijski primjer

Kao primjer korištena je matrica s dvije nezavisne varijable (n = 767). Prva nezavisna varijabla je vrsta nastave; kategorijalna dihotomna varijabla je učitelj razredne i učitelj predmetne nastave, dok je druga nezavisna varijabla politomna kategorijalna – usavršavanje (četiri kategorije). Cilj je testirati hipotezu o tome razlikuju li se učitelji razredne i predmetne nastave u procjeni discipline; zavisna varijabla (glavni učinak 1; *main effect* 1) te razlikuju li se ispitanici s obzirom na napredovanje (zvanje – bez napredovanja, mentor, savjetnik i viši savjetnik) u istoj zavisnoj varijabli (glavni učinak 2; *main effect* 2) te konačno u sklopu faktorijalne dvosmjerne analize varijance (*two way ANOVA*) postojanje interakcijskog učinka u obje nezavisne varijable na zavisnu. Dakle, riječ je o 2*4 dvosmjernoj analizi varijance (*two way ANOVA)*.

Budući da je *ANOVA* parametrijski test, trebaju biti ispunjeni preduvjeti za njihovo korištenje. Glavni preduvjet je normalna distribucija. Vrijednosti KS-testa zavisne varijable (KS = 0,245; df = 767; p < 0,001) te *Shapiro Wilks test* (SW = 0,804; df = 767; p < 0,001) upućuju da nije ispunjen preduvjet o normalnosti distribucije. Stoga treba biti oprezan u zaključivanju jer je riječ o velikom uzorku, a veličina uzorka djeluje na statističku značajnost (npr., Opić i Rijavec, 2022.). Dodatnim uvidom u specifičnosti distribucije uzorka (Skew = -0,982 ; Kurtosis = -0,060) te M = 3,90; SD = 1,253; CV = 26,3 %) vrijednosti upućuju na to da distribucija znatno ne odstupa od normalne (blaga lijeva asimetrija, iznimno blaga platikurtična distribucija, homogeno obilježje). Osim statističkih podataka dobro je prikazati i grafički uvid, npr., P-P plot i/ili Q-Q plot (Slika 4.).



Slika 4. *Grafički prikaz normalnosti distribucije zavisne varijable; Probability – Probability plot (P-P)*

Homogenost varijanci subuzoraka duž zavisne varijable upućuje da je ispunjen preduvjet o homoscedascitetu. Točnije, testirana je nul-hipoteza da su varijance pogreške (*error*) zavisne varijable jednake duž uspoređivanih grupa, što je utemeljeno na AS (L = 1,728, df1 = 759, df2 = 759, p = 0,099), na medijanu (L = 1,868, df1 = 7, df2 =759, p = 0,072)) na medijanu s prilagođenom df (L = 1,868, df1 = 7, df2 = 734,190, p = 0,072) te na *trimmed AS* (L = 1,567, df1 = 7, df2 = 759, p = 0,142). U literaturi postoje tumačenja da je *ANOVA* robusna na odsutnost homogenosti varijanci kad su subuzorci približno iste veličine (Glass, Peckham, Sanders, 1972.; Field, 2009.). No, kad grupa s većim brojem ispitanika ima veću standardnu devijaciju nego grupa s manjim brojem ispitanika, tada F omjer nastoji biti konzervativan, odnosno pokazuje statističku beznačajnost, iako razlike između subuzoraka na zavisnoj varijabli postoje u populaciji. Obrnuto, kad grupa s većim brojem ispitanika ima manju standardnu devijaciju u odnosu na grupu s manjim brojem ispitanika, rezultat je da F-test ne kontrolira pogrešku prvog reda (*type one error*).

No, suprotno je promišljanje o tome, npr., Coombsa, Altina i Oltmana (1996., prema Pituch i Stevens, 2016.) koji su na primjeru prikazali sljedeće: iako su subuzorci iste veličine, *ANOVA* nije robusna na nejednake varijance. Budescu (1982.) ističe sljedeće: kad su varijance proporcionalne aritmetičkim sredinama, snaga F-testa (omjera) nije pod utjecajem heterogenih varijanci (neispunjeni preduvjeti o homogenosti) te načini stabiliziranja i kontrole varijance ne povećavaju snagu F-testa.

Osim Levene testa za testiranje homogenosti varijanci vrlo je koristan White-test kojim se testira jesu li varijance pogrešaka (*error variance*) konstantne (*homoskedasticity*). Testira se nul-hipoteza kojom se pretpostavlja da varijance pogrešaka ne zavise o vrijednostima nezavisnih varijabli (vrlo sličan je i za računanje jednostavniji modificirani (*modified*) *Breusch-Pagan Test for Heteroskedasticity*.

Glavni učinci V3, V5 te interakcijski učinak V3*V5 prikazani su u Tablici 1.

Tablica 1. Testovi učinaka između varijabli *(Tests of Between-Subjects Effects)*
Zavisna varijabla; v10.10 PR (disciplina)

Izvor	Tip 3 sume kvadrata (Type III SS)	Broj stupnjeva slobode (df)	Prosjek kvadrata (MS)	F	Statistčka značajnost	Parcijalna eta na kvadrat
Korigirani model	23,498 [a]	7	3,357	2,160	,036	,020
Intercept	2009,023	1	2009,023	1292,494	,000	,630
V3	2,458	1	2,458	1,581	,209	,002
V5	,309	3	,103	,066	,978	,000
V3 * V5	16,284	3	5,428	3,492	,015	,014
Pogreška	1179,772	759	1,554			
Total	12867,000	767				
Korigirana suma	1203,270	766				

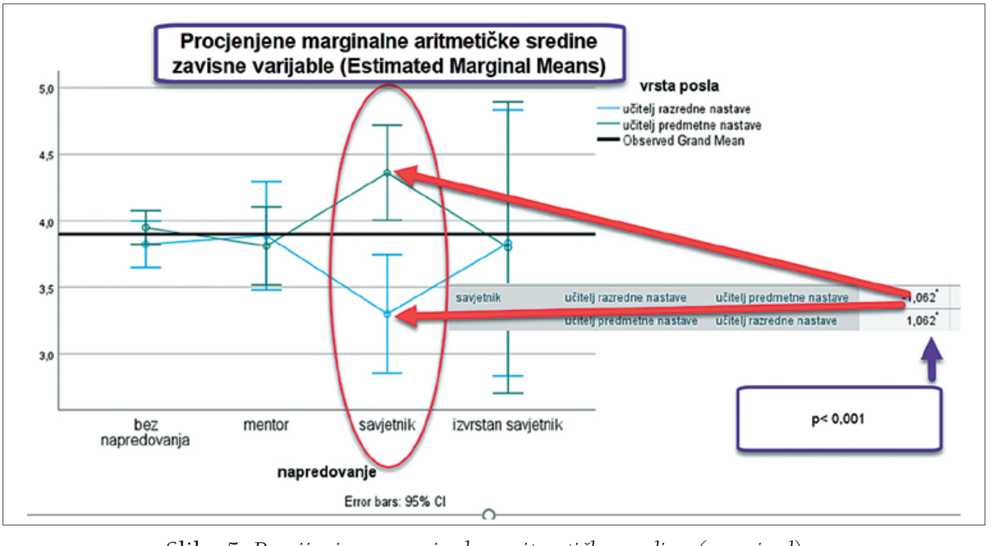
a. R^2 = ,020 (prilagođena R^2 = ,010)

Dakle, dvosmjerna analiza varijance (*two way* ANOVA), Tip 3 sume kvadrata (*type three SS*) nebalansiranog dizajna pokazuje odsutnost glavnih učinaka varijabli V3 i V5, ali postoji njihov interakcijski učinak na zavisnu varijablu. Nove procijenjene deskriptivne vrijednosti (*estimated values*) interakcijskog učinka 2*4 dvosmjerne analize varijance s Tipom 3 sume kvadrata (*type three SS*) prikazane su u Tablici 2.

Tablica 2. Vrsta posla * napredovanje
Zavisna varijabla: v10.10 PR

VRSTA POSLA (V3)	NAPREDOVANJE (V5)	ARITMETIČKA SREDINA	STANDARDNA GREŠKA	95 % INTERVAL POUZDANOSTI	
				Niža granica	Viša granica
Učitelj razredne nastave	bez napredovanja	3,824	,088	3,651	3,998
	Mentor	3,889	,208	3,481	4,297
	Savjetnik	3,300	,228	2,853	3,747
	izvrstan savjetnik	3,833	,509	2,834	4,833
Učitelj predmetne nastave	bez napredovanja	3,949	,064	3,823	4,076
	Mentor	3,812	,150	3,517	4,106
	Savjetnik	4,362	,182	4,005	4,719
	izvrstan savjetnik	3,800	,558	2,705	4,895

Usporedba parova (*Pairwise comparison*) utemeljena na procijenjenim marginalnim aritmetičkim sredinama (*estimated marginal means*) s prilagodbom za mutiplu usporedbu parova LSD (*Least significant difference*) napravljena je u *syntaxu* i prikazuje da postoji razlika samo kod učitelja (RN i PN) koji su savjetnici (AS_{diff} = -1,062 (I – J); standardna pogreška (*error*) = 0,291; p = 0,000287; 95 % interval pouzdanosti razlika; niža granica (*lower bound*) = -1,634; te viša granica (*upper bound*) = -0,490). Grafički prikaz navedenog prikazan je na Slici 5.



Slika 5. *Procijenjene marginalne aritmetičke sredine (marginal) – prikaz razlika usporedbe parova (pairwise comparison)*

No, ovdje je riječ o nebalansiranom dizajnu i zbog principa marginalnosti treba biti oprezan u interpretaciji glavnih učinaka. U ovome primjeru s pomoću Tipa 3 sume kvadrata (*type three SS)* u dvosmjernoj analizi varijance (*two way ANOVA*) nisu determinirani glavni učinci, iako oni zapravo postoje. Nakon što je promijenjeno u Tip 1 sumu kvadrata (*type one SS)* i Tip 2 sumu kvadrata (*type two SS)* potvrđen je glavni učinak (*main effect*) varijable V3.

V3; Tip 1 suma kvadrata; *type one SS* (6,517); df = 1; MS = 6,517; F = 4,192; p = 0,041
V5; Tip 1 suma kvadrata; *type one SS* (0,697); df = 3; MS = 0,232; F = 0,192; p = 0,930
V3; Tip 2 suma kvadrata; *type two SS* (6,540); df = 1; MS = 6,540; F = 4,207; p = 0,041
V5; Tip 2 suma kvadrata; *type two* SS (0,697); df = 3; MS = 0,232; F = 0,149; p = 0,930.

Dodatno, u jednosmjernoj analizi varijance (*One way ANOVA*) za V3 također se potvrđuje učinak (F(1,765) = 4,166, p = 0,042. Dakle, s pomoću Tipa 3 sume kvadrata (*type three* SS) u nebalansiranom dizajnu dvosmjerne analize varijance (*two way ANOVA*) 2*4 potvrđen je interakcijski učinak varijabli v3* v5, no izostao je glavni učinak. Iz primjera se vidi da glavni učinak varijable V3 postoji, ali ne i učinak V5.

Zaključak

Faktorijalna analiza varijance (*ANOVA*) vrlo je koristan test za testiranje interakcijskih učinaka dviju ili više nezavisnih varijabli na zavisnu. Proširenjem faktorijalnih dizajna na multivarijatno područje inferencijalne statistike ulazi se u još složenije interakcijske odnose (*MANOVA, MANCOVA*). Budući da je riječ o složenim statističkim postupcima, treba biti oprezan u zaključivanju. U ovome radu, u teorijskom i empirijskom dijelu, prikazane su specifičnosti Tipa 1 sume kvadrata (*type one SS*), Tipa 2 sume kvadrata (*type two SS*) i Tipa 3 sume kvadrata (*type three SS*) u nebalansiranom dizajnu, i to s obzirom na princip marginalnosti. Riječ je o mogućem krivom zaključivanju o glavnim učincima u području testiranja interakcijskih učinaka u faktorijalnom dizajnu analize varijance (*ANOVA*). Uzrok je u specifičnosti izračuna sume kvadrata (Tip 3 – *type three SS*) koja je u većini statističkih programa postavljena prema *defaultu.* U simulaciji je prikazan nebalansirani faktorijalni dvosmjerni (*two way)* model u kojemu je izostao glavni učinak, iako on postoji. Riječ je o principu marginalnosti i nameće se potreba, pri testiranju interakcijskih učinaka, da se koristi Tip 3 sumu kvadrata (*three way SS*), ali ne i za testiranje glavnih učinaka. Preporučuje se da se koristi usporedo Tip 1 i Tip 2 sumu kvadrata (*type one SS, type* *two SS*), odnosno da se dodatno provjeri s pomoću jednosmjerne analize varijance (*one way ANOVA*) zasebne učinke (glavne) svake nezavisne varijable na zavisnu.

Literatura

1. Aho, K. A. (2013), *Foundational and Applied Statistics for Biologists using R.* Taylor and Francis Group.
2. Budescu, D. V. (1982), The Power of the F test in normal Populations with Heterogeneous variances. *Educational and Psychological Measurement, 42,* 609-616.
3. Field, A. (2009), *Discovering Statistics using SPSS* (third edition). Sage publication.
4. Foster, J. J., Barkus, E., & Yavorsky, C. (2006), *Understanding and Using Advanced Statistics: A Practical Guide for Students.* SAGE. https://doi.org/10.4135/9780857020154
5. Glass, G. V., Peckham, P. D., & Sanders. J. R. (1972), Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research, 42*(3), str. 237–288.
6. Hector, A., von Felten, S., & Schmid, B. (2010), Analysis of variance with unbalanced data: an update forecology & evolution. *Journal of Animal Ecology*, 79, str. 308-316.
7. Herr, D. G. (1986), On the History of ANOVA in Unbalanced, Factorial Designs: The First 30 Years. *The American Statistician, 40,* str. 265-270.
8. Kempthorne, O. (1975), Fixed and Mixed Models in the analysis of Variance. *Biometrics*, *38,* str. 613-621.
9. Landsheer, J. A., & van den Wittenboer, G. (2015), *Unbalanced 2 x 2 Factorial Designs and the Interaction Effect: A Troublesome Combination*. PLoS ONE 10(3): e0121412. https://doi.org/10.1371/journal.pone.0121412.
10. Lewsey, J. D., Gardiner, W. P., &Gettinby, G. (1997), A study of simple unbalanced factorial designs that use type II and type III sums of squares. *Commun. Stat. Simulat. Comput*. 26, str. 1315–1328. doi: 10.1080/03610919708813442
11. Lewsey, J. D., Gardiner, W. P., & Gettinby, G. (2001), A study of type II and type III power for testing hypotheses from unbalanced factorial designs. *Commun. Stat. Simulat. Comput*. 30, str. 597–609. doi: 10.1081/SAC-100105081
12. Langsrud, Ø. (2003), ANOVA for unbalanced data: Use Type II instead of Type III
13. sums of squares. *Statistics and Computing*, *13*(2), str. 163-167.
14. Milliken, G. A., &Johnson, D. E. (2009), *Analysis of Messy Data: Vol. I. Designed Experiments* (2n d edition). CRC Press.
15. Nelder, J. A., & Lane, P. W. (1995), The Computer Analysis of Factorial Experiments: In Memoriam – Frank Yates. *The American Statistician,* 49, str. 382-385.
16. Opić, S., & Rijavec, M. (2022), Misconceptions of the p-value – let us use new approaches and procedures. U D. Velički, M. Dumančić, (ur.), *Suvremene teme u odgoju i obrazovanju (STOO 2) – in memoriam prof. emer. dr. sc. Milanu Matijeviću*, 2022. Učiteljski fakultet Sveučilišta u Zagrebu i Zavod za znanstvenoistraživački rad u Bjelovaru Hrvatske akademije znanosti i umjetnosti.

17. Pituch, K. A., & Stevens, J. P. (2016), *Applied Multivariate Statistics for the Social Science (Analysis with SAS and IBM´S SPSS)* (sixt edition). Routledge.
18. Stewart-Oaten, A. (1995), Rules and Judgments in Statistics: Three Examples.
19. *Ecology,* 76(6), 2001–2009. Retrieved from http://www.jstor.org/stable/1940736
20. Smith, C. E., & Cribbie, R. (2014), Factorial ANOVA with unbalanced data: A fresh look at the types of sums of squares. *Journal of Data Science,* 12, str. 385-404.
21. Vogt, P. W., & Johnson, B. R.(2016), *The Sage Dictionary of Statistics & Methodology – A Nontechnical Guide for the Social Sciences* (fifth edition). Sage publication.
22. Yates, F. (1934), The analysis of multiple classifications with unequal numbers in
23. the different classes. *Journal of the American Statistical Association*, 29(185), str. 51.
24. https://www.utstat.utoronto.ca/reid/sta442 f/2009/typeSS.pdf; preuzeto 11.2.2023.
25. https://www.flutterbys.com.au/stats/tut/tut7.6a.html; poreuzeto 02.1.2023.

Factorial designs Analysis of variance (ANOVA) in types sum of squares 1, 2, 3 (type one SS, type two SS, type three SS) in unbalanced designs – principle of marginality

Summary

The aim of this paper is to show the specifics of use, but also the limitations of factorial designs in quantitative differential designs in the field of social sciences. The specifics of two-way and multi-way (n-way) factorial designs in the Analysis of Variance (ANOVA) are presented with regard to the specifics of the types of sum of squares (type one SS, type two SS, type three sum of squares). An unbalanced 2*4 two-way ANOVA model was tested using Type three sum of squares (type three SS) in which the main effect was absent. Due to the principle of marginality, the main effects cannot be interpreted using the Type 3 sum of squares (type three SS), but for the main effects the need for a particular approach, one-way ANOVA and two-way analysis of variance (factorial design) is imposed, but not with using Type 3 sum of squares.

Keywords: ANOVA; interaction effects; marginality; sum of squares of type 1, type 2, type 3

Prof. dr. sc. Siniša Opić
Učiteljski fakultet Sveučilišta u Zagrebu
Savska cesta 77, HR – 10000 Zagreb
sinisa.opic@ufzg.unizg.hr