UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI
FAKULTET
DEPARTMAN ZA MATEMATIKU
I INFORMATIKU

Vladan Todić

Primena faktorijalnih ogleda

-Master rad-

Mentor: prof. dr Zagorka Lozanov-Crvenković

Novi Sad, 2014. godine

Sadržaj

Predgovor

Tema ovog master rada su faktorijalni ogledi i njihova primena u raznim naučnim oblastima. Faktorijalni ogledi su eksperimenti u kojima se ispituju uticaji dva ili više faktora, od kojih svaki ima bar dva nivoa, na jednu ili više promenljivih. Eksperimentalne jedinice se sastoje od svih mogućih kombinacija nivoa faktora. U ovakvom eksperimentu se pored uticaja pojedinih faktora, može ispitivati i uticaj interakcije faktora na promenljive. Na primer, ispituje se uticaj tri sorte pšenice i četiri vrste đubriva na prinos i veličinu zrna.

Rad se sastoji iz četiri celine, i svaka od njih pored teorijskog dela sadrži i primere na kojima su objašnjena primena i način izvođenja statističke analize.

U prvom delu rada kroz primer je data motivacija za upotrebu faktorijalnih ogleda, osnovni principi statističkog planiranja eksperimenta – ponavljanje, randomiziranje i pravljenje blokova. Takođe su dati osnovni pojmovi jednofaktorske i dvofaktorske analize varijanse.

U drugom delu rada date su matematičke osnove faktorijalnih ogleda, specijalno kroz oglede sa dva faktora. Takođe su prikazana višestruka poređenje, testiranje adekvatnosi modela i ocena parametara modela.

U trećem delu, obrađeni su 2^k faktorijalni ogledi - ogledi sa k faktora, od kojih svaki ima samo dva nivoa. Specijalno, pokazani su 2^2 – ogledi sa dva faktora, 2^3 – ogledi sa tri faktora i na kraju opšti slučaj 2^k ogled sa *k* faktora i sa jednim ponavljanjem.

U posledenjem, četvrtom delu prikazana je primena faktorijalnih ogleda u poljopriredi, konkretno u proizvodnji šećerne repe.

Statistička obrada podataka je urađena pomoću statističkog softvera *Statistica 7.0.*

****

Veliku zahvalnost dugujem svom mentoru, prof. dr Zagorki Lozanov-Crvenković, na ukazanom poverenju i pomoći prilikom izrade ovog master rada.

U Novom Sadu, 2014. godine. Vladan Todić

Glava 1

1. Uvod u faktorsku analizu

Metodi matematičke statistike posvećeni su rešavanju dva osnovna problema. Prvi problema odnosi se na planiranje eksperimenta, odnosno na izbor odgovarajućih faktora te na ocenu njihovog uticaja. Drugi problem, koji sledi posle prvog odnosi se na analizu rezlutata dobijenih eksperimentima.

Pri proučavanju uticaja faktora na vrednost nekog obeležja obično se pretpostavlja da je zavisnost linearna i da se taj uticaj odvija u uslovima koji dovode do normalne raspodele. Linearni modeli su zato i najzastupljiniji u sprovođenju statističkih ogleda. Specijalan slučaj linearnih modela gde su faktori fiksirani jesu i faktorijalni ogledi. Ovde će biti predstavljeni neki važni faktorijalni ogledi: faktorijalni ogledi sa tačno dva nivoa ili uticaja, faktorijalni ogledi sa više faktora i samo dva nivoa, te disperziona analiza, ili analiza varijanse. Svaki statistički model ilustrovan je kroz odgovarajuće primere.

1.1. Osnovne definicije i principi faktorijalnih ogleda, motivacioni primer

Mnogi eksperimenti, bilo da se radi o prirodnim ili društvenim naukama, podrazumevaju ispitivanje uticaja ili efekata jednog, dva ili više faktora na vrednost nekog obeležja. Generalno, faktorijalni ogledi su najefiksaniji za ovaj tip eksperimenata. Faktorijalnim ogledom se u svakom izvođenju eksperimenta, ili njegovom ponavljanju ispituju sve kombinacije nivoa ili tretmana svih faktora. Na primer, ako je a nivoa faktora A i b nivoa faktora B onda svako ponavljanje eskperimenta sadrži ab kombinacija nivoa, odnosno sve moguće. [1]

Osnovne pojmove vezane za faktorijalne oglede uvešćemo pomoću jednostavnog primera. Uticaj ili efekat faktora definiše se kao promena vrednosti posmatranog obeležja nastalog promenom nivoa faktora. Ovo se često naziva ***glavni efekat.***

Na primer, posmatrajmo podatke u Tabeli 1.

Tabela 1.1.

		Faktor B	
		B_1	B_2
Faktor A	A_1	20	30
	A_2	40	52

Glavni efekat faktora A možemo da posmatramo kao razliku između prosečne vrednosti obeležja na prvom nivou faktora A i prosečne vrednosti na drugom nivou faktora A. Tako se dobije da je glavni efekat faktora A:

$$A = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

Ovo znači da promenom nivoa faktora A sa nivoa 1 na novo 2 dolazi do promene prosečne vrednosti obeležja za 21. Analogno, glavni efekat faktora B je:

$$B = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

Ako faktori imaju više od dva nivoa, prethodni postupak se mora modifikovati, jer onda postoji više načina da se izrazi razlika između prosečnih vrednosti posmatranog obeležja. O tome će biti reči kasnije.

U nekim eksperimentima moguće je uočiti da razlika vrednosti obeležja između nivoa jednog faktora nije ista za sve nivoe drugog faktora. Kada se ovo desi, onda se kaže da postoji ***interakcija*** između faktora.. Sad posmatrajmo nove podatke u Tabeli 2 i odredimo efekte faktora A i B.

Tabela 1.2.

		Faktor B	
		B_1	B_2
Faktor A	A_1	20	40
	A_2	50	12

Na prvom nivou faktora B, efekat faktora A je:

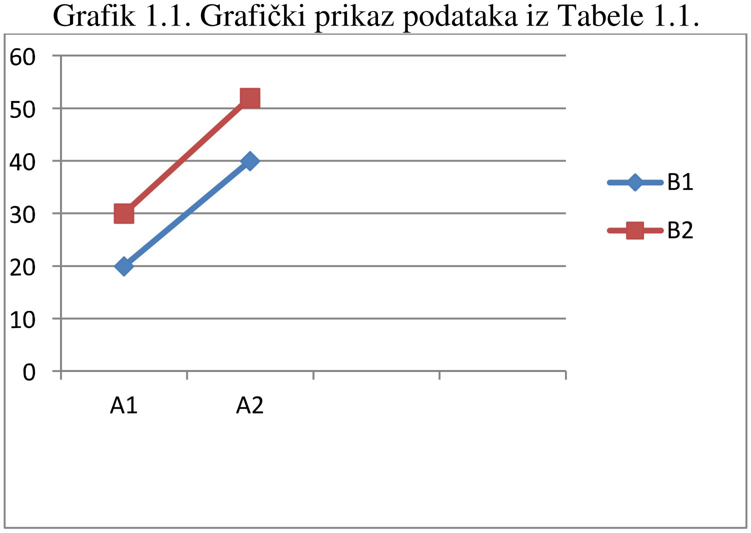
$$A = 50 - 20 = 30,$$

a na drugom nivou faktora B, efekat faktora A je:

$$A = 12 - 40 = -28.$$

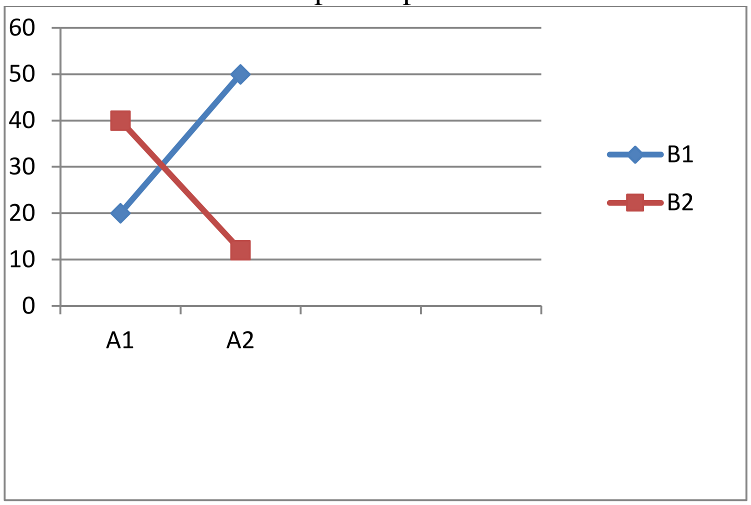
Kako efekat faktora A zavisi od izbora nivoa faktora B, vidimo da postoji interakcija između faktora A i B. Ovu ideju možemo ilustrovati i grafički. Posmatrajmo Grafik 1 i Grafik 2 kojima odgovaraju podaci iz Tabele 1 i Tabele 2.

Grafik 1.1. Grafički prikaz podataka iz Tabele 1.1.



Primetimo da su duži B_1i B_2 približno paralelne što upućuje na manjak interakcije između faktora A i B. Međutim, ako pogledamo drugu tabelu, Tabelu 2, videćemo da iste duži nisu paralelne. To govori o značajnoj interakciji između pomenutih faktora.

Grafik 1.2. Grafički prikaz podataka iz Tabele 1.2.



Ovakvi grafici su često jako korisni da se uoči značajna ineterakcija između faktora, ali ne treba da bude glavna tehnika analize podataka, jer je subjektivna i može da dovede do pogrešnih zaključaka. Primetimo i to da kada je interakcija velika, odgovarajući glavni efekti imaju mali praktični značaj. Za drugi primer, tj. za podatke iz Tabela 2 možemo izračunati glavni efekat faktora A i dobiti da je:

$$A = \frac{50 + 12}{2} - \frac{20 + 40}{2} = 1$$

što je veoma malo i iz ovoga bismo zaključili da nivoi faktora A ne utiču na vrednosti obeležja. Međutim, kada izračunamo efekte faktora A za različite nivoe faktora B vidimo da to nije tačno. Faktor A utiče na vrednost obeležja, ali je to uslovljeno nivoima faktora B. S toga, informacija da postoji interakcija AB je mnogo korisnija od informacije vezane za glavni efekat faktora A. Značajna interakcija će obično prikriti značajni glavni efekat faktora. Kada je to slučaj, ispitivač mora da testira sve nivoe jednog faktora, recimo faktora A, kombinujući ih sa svim nivoima drugog faktora, recimo faktora B, i tako doneti zaključak o glavnom efektu faktora A. [1]

1.2. Prednosti faktorijalnih ogleda

Prednosti faktorijalnih ogleda mogu se vrlo jednostavno ilustrovati. Pretpostavimo da imamo dva faktora A i B, svaki sa po dva nivoa. Definišimo nivoe faktora kao A_1, A_2, B_1 i B_2. U Tabeli 3 prikazane su kombinacije nivoa faktora:

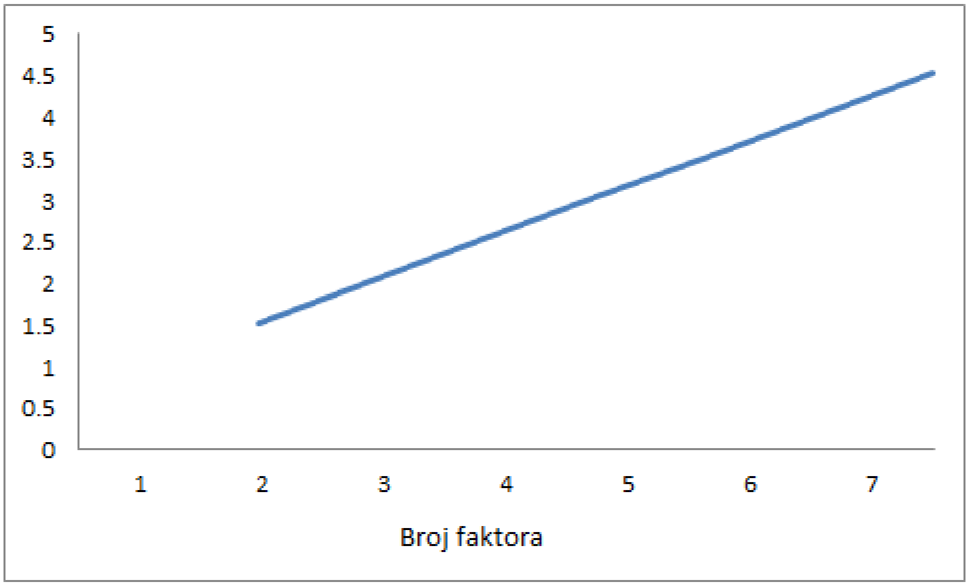
Tabela 1.3.

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Faktor B</td></tr>
<tr><td>B_1</td><td>B_2</td></tr>
<tr><td rowspan="2">Faktor A</td><td>A_1</td><td>A_1B_1</td><td>A_1B_2</td></tr>
<tr><td>A_2</td><td>A_2B_1</td><td></td></tr>
</table>

Efekat promene nivoa faktora A je dato sa $A_2B_1 - A_1B_1$. Kako je eksperimentalna greška prisutna, poželjno je ponoviti eksperiment dva puta za svaku kombinaciju i onda izračunati efekat faktora koristeći srednju vrednost obeležja. To znači ukupno šest posmatranja.

Ako bismo koristili faktorijalni eksperiment, onda bismo morali dodati i četvrtu kombinaciju nivoa faktora, tj. A_2B_2. Sada, koristeći samo četiri opservacije, moguće je izračunati dva efekta faktora A: $A_2B_1 - A_1B_1$ i $A_2B_2 - A_1B_2$. Slično je moguće izračunati i dva efekta za drugi faktor B. Ove dve ocene efekata za oba faktora mogu se iskoristiti za računanje odgovarajućih glavnih efekata koji će dati iste rezultate kao i t-test, ali su bile potrebne samo četiri opservacije, a ne šest. Možemo reći da je relativna efiksanost faktorijalnog eksperimenta u odnosu na t-test $\frac{6}{4} = 1{,}5$. Uopšteno, relativna efikasnost se povećava kako se broj faktora povećava kao što je prikazano na Grafiku 3.

Grafik 1.3. Relativna efikasnost faktorijalnih ogleda



Dalje, pretpostavimo da postoji interakcija između faktora. Ako bismo radili t-test i dobili da su vrednosti A_1B_2 i A_2B_1 bolje nego vrednost koju je dala kombinacija nivoa A_1B_1, logično bi bilo da zaključimo da će kombinacija nivoa A_2B_2 dati još bolje rezultate. Međutim, ako postoji interakcija, ovaj zaključak bi mogao biti pogrešan. Takvu situaciju smo imali u prethodnom primeru. [1]

Na kraju, primetimo da faktorijalni ogled ima nekoliko prednosti. Kao prvo efikasniji je, odnosno zahteva manji broj posmatranja obeležja. Dalje, faktorijalni ogled je neophodan kada postoji interakcija kako bi se izbegle greške nastale pogrešnim zaključivanjem. Na kraju, faktorijalni ogled omogućava izračunavanje efekata jednog faktora na više nivoa drugih faktora, što doprinosi relevantom zaključivanju.

1.3. Osnovni principi statističkog planiranja eksperimenata

Pod statitističkim modelovanjem eksperimenata podrazumeva se proces planiranja eksperimenta tako da se na odgovarajuće podatke može izvršiti anliza nekim statističkim metodom, te dobiti validni i objektivni zaključci. Statistički pristup dizajniranju eksperimenta je neophodan ako se želi doći do značajnih zaključaka iz podataka. Često neki problem koji se posmatra uključuje podatke koji dovode do subjektivnih eksperimentalnih grešaka, tada je statistička metodologija jedini pravi, objektivni pristup analizi. Inače, postoje dva aspekta kod svakog eksperimentalnog problema: dizajniranje eksperimenta i statistička analiza dobijenih podataka. Ova dva aspekta su u tesnoj vezi, jer metod statističke analize zavisi direktno od toga koji model eksperiment je urađen.

Tri osnovna principa planiranja eksperimenta su ponavljanje, randomiziranje i pravljenje blokova. Ovi principi su važan deo svakog eksperimenta.

Pod ponavljanjem se podrazumeva izvođenje osnovnog eksperimenta više puta. Ponavljanje ima dve važne karakteristike. Prva je ta da se omogućuje onome ko izvodi eksperiment da dobije procenu eksperimentalne greške. Procena greške postaje osnovna jedinica mere za određivanje da li su uočene razlike u podacima zaista statistički značajne. Drugo, ako se očekivanje uzorka ili njegova srednja vrednost koristi u eksperimentu za procenu uticaja nekog faktora na obeležje, onda ponavljanje omogućava dobijanje bolje, tj. preciznije procene ovog uticaja ili efekta. Na primer, ako je σ^2 disperzija podataka i eksperiment je ponovljen n puta, onda je odstupanje od srednje vrednosti uzorka:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

To praktično znači da ako se recimo eksperiment izvodi samo jednom ($n = 1$), iz dobijenih podataka nije moguće doneti zadovoljavajući zaključak o uticaju nekog faktora, jer uočene razlike mogu biti posledica eksperimentalne greške. Sa druge strane, ako se eksperiment ponavljao više puta, tačnije n puta (n je dovoljno veliko) i ako je eksperimentalna greška bila dovoljno mala, dobiće se i bolja procena uticaja određenog faktora.

Randomiziranje je osnova upotrebe statističkih metoda u dizajniranju eksperimenata. Pod randomiziranjem se misli na uzimanje prostog slučajnog uzorka. Statističke metode zahtevaju da posmatranja (ili greške) budu nezavisne slučajne promenljive. Randomizacija obično ove pretpostavke učini tačnim. Randomizacijom se takođe omogućava izračunavanje srednje vrednosti uticaja drugih faktora koji se u eksperimentu mogu pojaviti. Randomizacijom, tj. nasumičnim izvođenjem eksperimenta na minimum se svodi uticaj nekontrolisanih faktora.

Pravljenje blokova je tehnika koja se koristi kako bi se povećala preciznost eksperimenata. Blok predstavlja deo uzorka koji je na neki način homogeniji u odnosu na celi uzorak. Drugim rečima pravi se grupisanje. Pravljenje blokova uključuje pravljenje poređenja između stanja koja se ispituju eksperimentom unutar svakog bloka. [2]

1.4. Analiza varijanse

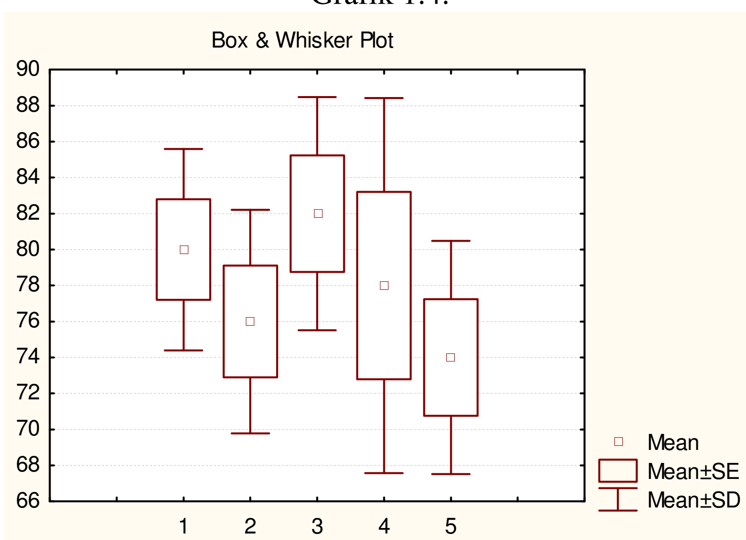
Analiza varijanse (ANOVA) je analitički model za testiranje značajnosti razlika. Prednost ove metode ogleda se u tome što u model ulaze u obzir svi varijabiliteti kao i njihov međusobni uticaj, što je nemoguće proceniti na drugi način. Prilikom izbora modela treba voditi računa o prirodi posmatranog obeležja, samim jedinicama posmatranja kao i karakteristikama samog modela, da bi se na najbolji način zadovoljili postavljeni ciljevi i omogućilo da pomoću prikupljenih podataka dođemo do validnih rezultata. Analiza varijanse u suštini predstavlja specijalan matematičko statistički postupak koji omogućuje testiranje značajnosti razlike između aritmetičkih sredina iz tri i više uzoraka, a u okviru toga i testiranja uticaja jednog ili više faktora na varijabilitet nekog testiranog numeričkog obeležja. Začetnik tehnika analize varijanse bio je statističar i genetičar Ronald Fišer dvadesetih i tridesetih godina prošlog veka. Samo ime analiza varijanse potiče od toga što se pre svega koriste statistike koje su zbirovi kvadrata nekih odstupanja. Moguće je posmatrati uticaj jednog, dva ili više faktora, pa razlikujemo jednofaktorsku i višefaktorsku analizu varijanse. [3]

Primer 1: Pretpostavimo da su studenti statistike podeljeni u četiri grupe i da je svaka grupa dobila drugačiji uđžbenik iz statistike. Dalje pretpostavimo da svi studenti idu na ista predavanja, rade iste domaće zadatke i polažu isti ispit na kojem dobijaju bodove od 0 do 100. Podaci o postignutim bodovima dati su u Tabeli 3:

Tabela 1.4.

	Studenti				
Udžbenik	1	2	3	4	Prosečan broj bodova
1	82	75	87	76	80
2	67	79	77	81	76
3	91	82	76	79	82
4	66	73	89	84	84
5	82	71	67	76	76

Grafik 1.4.



Razlike među ostvarenim bodovima postoje, ali se postavlja pitanje da li su one statistički značajne. Jedna od mogućnosti je da se izvede t-test za svaki od 10 parova prosečnih vrednosti. Sem velikog broja testova, koji se drastično povećava sa brojem nivoa na kojima su uzeti uzorci, postoji i drugi, ozbiljniji problem koji odbacuje ovu mogućnost. Ovakvo testiranje dovodi do drastičnog povećanja greške prve vrste. U našem primeru uzorci su izvučeni slučajno, pa su testovi nezavisni. Greška prve vrste svakog od 10 testova je 0.05, odnosno verovatnoća tačnog prihvatanja hipoteze ako je ona tačna je 0.95. Tada je verovatnoća tačnog prihvatanja hipoteze za svih 10 testova $(0.95)^{10} = 0.6$ što znači da je greška prve vrste za ukupan test 0.4.

Prvi način za testiranje hipoteze o jednakosti više očekivanja je analiza varijanse. Važno je uočiti da se testiraju hipoteze o matematičkim očekivanjima, a ne o varijansama. [3]

1.4.1 Jednofaktorska analiza varijanse

Posmatra se dejstvo nekog faktora *A* na vrednost obeleđja *X*. Faktor *A* deluje preko $k \geq 2$ svojih nivoa ili stanja koji se obeležavaju sa $A_1, A_2, ..., A_k$. U *primeru 1* faktor je udžbenik iz statistike, njegovi nivoi su 1, 2, 3, 4 i 5 a obeležje koje se meri je broj osvojenih poena. [3] Prilikom izvođenja eksperimenta cela populacija ε na kojoj se posmatra obeležje *X* se deli na podpopulacije $\varepsilon_1, \varepsilon_2, ..., \varepsilon_k$ i na element svake podpopulacije deluje samo jedan od nivoa faktora *A*. Neka je *m* očekivana vrednost obeležja *X* na celoj populaciji $m = E(X)$, a m_i očekivana vrednost obeležja *X* na podpopulaciji ε_i ; $i = 1,2,...,k$. Razlike $\gamma_i = m_i - m$ zovemo efekat $i - tog$ nivoa faktora *A*. Kako je svaki element populacije izložen samo dejstvu jednog nivoa faktora *A*, imamo da je:

$$\sum_{i=1}^{k} \gamma_i = 0$$

Testira se hipoteza

$$H_0(m_1 = m_2 = ... = m_k = m)$$

protiv alternativne

$$H_1(m_i \neq m_j \text{ } za \text{ } bar \text{ } jedan \text{ } par \text{ } i, j)$$

Ovaj par hipoteza se može zapisati na sledeći način

$$H_0(\gamma_1 = \gamma_2 = ... = \gamma_k = 0)$$

(Svi efekti nivoa su jednaki nuli, odnosno nema razlike među nivoima faktora *A*),
protiv alternativne

$$H_1(\gamma_i \neq 0 \text{ } za \text{ } bar \text{ } jedno \text{ } i = 1,2,...,k).$$

(postoji razlika između bar dva nivoa).

Testiranje se vrši na osnovu prostog slučajnog uzorka. Iz svake populacije se izvlači prost slučajni uzorak $(X_{i1}, X_{i2,}..., X_{in_i})$, $i = 1,2,...,k$. Obimi uzorka n_i ne moraju biti jednaki, a obim celog uzorka je N, $N = n_1 + n_2 + ... + n_k$. Obeležje X_{ij} je vrednost osnovnog obeležja *X* na $j - tom$ elementu uzorka iz $i - te$ populacije ε_i.
Rezultate možemo prikazati sledećom Tabelom 1.5.:

Tabela 1.5.

Nivoi faktora	Merenja							
	1	2	$\cdots$	j	$\cdots$	n_i	Zbir	Sredina
1	X_{11}	X_{12}	$\cdots$	X_{1j}	$\cdots$	X_{1n_1}	S_1	$\bar{X}_1$
2	X_{21}	X_{22}	$\cdots$	X_{2j}	$\cdots$	X_{2n_2}	S_2	$\bar{X}_2$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
i	X_{i1}	X_{i2}	$\cdots$	X_{ij}	$\cdots$	X_{in_i}	S_i	$\bar{X}_i$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
k	X_{k1}	X_{k2}	$\cdots$	X_{kj}	$\cdots$	X_{kn_k}	S_k	$\bar{X}_k$

Slučajna promenljiva X_{ij} slučajno odstupa od svog matematičkog očekivanja m_i, te je model analize varijanse:

$$X_{ij} = m_i + \varepsilon_{ij} = m + \gamma_i + \varepsilon_{ij}, \qquad i = 1,2,...,k \quad j = 1,2,...,n_i ,$$

gde ε_{ij} zovemo slučajne greške, nastale zbog dejstva velikog broja slučajnih faktora.

Osnovne predpostavke analize varijanse su:

- Greške ε_{ij} su nezavisne.
- Greške ε_{ij} imaju normalnu raspodelu $N(0,\sigma^2)$

Pretpostavka da sve greške imaju istu disperziju $D(\varepsilon_{ij}) = \sigma^2$, $i = 1,2,...,k$, $j = 1,2,...,n_i$ naziva se homoskedastičnost. Ova pretpostavka znači da različiti nivoi faktora A nemaju uticaj na rasejavanje, odnosno disperziju vrednosti obeležja oko njegove srednje vrednosti. Ovu pretpostavku je uvek potrebno proveriti pre provere hipoteze.

Iz pretpostavke da greške imaju normalnu raspodelu sledi i da slučajne promenljive X_{ij} imaju normalnu raspodelu. Za testiranje hipoteze potrebno je oceniti nepoznate parametre populacije na osnovu uzorka.

Kako se matematičko očekivanje obeležja ocenjuje pomoću sredine uzorka, uvode se sledeće statistike:

- Sredina celog uzorka

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}$$

- Sredina poduzorka izvučenog iz populacije ε_i:

$$\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} .$$

Suma kvadrata odstupanja pojedinačnih vrednosti od aritmetičke sredine koju zovemo zbir kvadrata totalna ili totalna varijacija:

$$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 = \sum_{i=1}^{k} X_{ij}^2 - N\bar{X}^2$$

Ova suma kvadrata totala se može podeliti na dva dela:

$$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i + \bar{X}_i - \bar{X})^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i}[(X_{ij} - \bar{X}_i)^2 + 2(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) + (\bar{X}_i - \bar{X})^2]$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 + 2\sum_{i=1}^{k}(\bar{X}_i - \bar{X})\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i) + \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$

S obzirom da je:

$$\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i) = 0$$

dobijmo

$$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2 = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 ,$$

odnosno

$$SS_T = SS_A + SS_E \qquad (*)$$

gde je:

$$SS_A = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2 , \qquad SS_E = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$$

Izraz SS_A predstavlja sumu kvadrata odstupanja grupnih sredina od opšte sredine i naziva se varijacija između grupa. Zbir kvadrata SS_P predstavlja odstupanje pojedinačnih vrednosti od njegovih grupnih sredina, i to je varijacija unutar grupa. Jednačina (*) znači da se totalni varijabilitet podataka, u smislu sume kvadrata, može podeliti na sume kvadrata razlike između sredina nivoa i ukupne sredine i na zbir kvadrata razilika između merenja unutar nivoa od sredina nivoa faktora *A*. Razlike između sredina nivoa faktora i ukupne sredine je mera razlike među sredinama nivoa, dok razlika između merenja unutar nivoa i sredine nivoa nastaje zbog slučajne greške.

Kada se totalna varijacija podeli brojem stepeni slobode koji iznosi $\sum_{i=1}^{k} n_i - 1 = N - 1$ dobija se varijansa. Jedan stepen slobode je izgubljen s obzirom na $\bar{X}$. Znači:

$$MS_T = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2}{N - 1}$$

Kod varijacije između grupa ima $k - 1$ stepeni slobode, jer se izračunava na osnovu opšte sredine tako da se gubi jedan stepen slobode. Njena varijansa prema tome je:

$$MS_A = \frac{\sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2}{k \ -1}.$$

Na kraju, varijansa varijacije unutar grupa je:

$$MS_E = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2}{N-k}.$$

Prema tome, totalna varijacija ima aditivni karakter i sastoji se iz varijacije između grupa i varijacije unutar grupa. Isto tako, stepeni slobode ove dve varijacije imaju aditivno svojstvo jer je:

$$(k-1)+(N-k)=N-1$$

Međutim, varijanse između grupa i unutar grupa nemaju tu osobinu, tj. njihov zbir nije jednak totalnoj varijansi. Praktično, to znači da se izračunavaju tri varijanse, koje se mogu smatrati kao tri nezavisne ocene varijanse osnovnog skupa. Treba napomenuti da se varijansa između grupa SS_A u praktičnom radu češće naziva varijansom tretmana, a varijansa unutar gurpa SS_P varijansom greške.

Rezultat računskog postupka se obično prikazuju na način dat u Tabeli 1.6.

Tabela 1.6. Šema analize varijanse

Izvori varijacije	*Zbir kvadrata*	*Stepeni slobode*	*Sredina zbira kvadrata (varijansa)*
Između gurpa	$SS_A = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$	$k-1$	$MS_A = \frac{Q_A}{k-1}$
Unutar grupa	$SS_E = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$	$N-k$	$MS_E = \frac{Q_P}{N-k}$
Ukupno	$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2$	$N-1$	

Izračunavanja za analizu varijanse u praksi se ne izvode na opisani način. Obično se izračunavaju totalna varijacija i varijacija između grupa, a varijacija unutar grupa se dobija kao razlika između njih. Međutim, izračunavanje i prve dve varijacije se izvodi na drugačiji način.

$$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij})^2}{N} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - C,$$

gde je

$$C = \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij})^2}{N},$$

i naziva se korektivni faktor.

Varijacija između grupa je tada

$$SS_A = \sum_{i=1}^{k} n_i (\bar{X}_i - \bar{X})^2 = \sum_{i=1}^{k} \frac{(\sum_{j=1}^{n_i} X_{ij})^2}{n_i} - C = \sum_{i=1}^{k} \frac{S_i^2}{n_i} - C.$$

Varijacija unutar gurpa dobiće se kao

$$SS_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 = SS_T - SS_A$$

Ako imamo jednak broj jedinica n u svakom poduzorku, izračunavanje je nešto jednostavnije i praktično se izvodi na sledeći način:

$$C = \frac{\left(\sum_{i=1}^{k} \sum_{j=1}^{n} X_{ij} \right)^2}{N},$$

$$SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n} X_{ij}^2 - C,$$

$$SS_A = \frac{\sum_{i=1}^{k} \left(\sum_{j=1}^{n} X_{ij} \right)^2}{n} - C.$$

Ovo objašnjenje analize varijanse ima više teorijski značaj jer se u praksi ispituje veći broj tretmana, tako da svaki uzorak u ogledu predstavlja drugi tretman. Na taj način, nije reč o jednom osnovnom skupu iz koga se uzimaju uzorci već bi pre moglo da se smatra da ima onoliko osnovnih skupova koliko i uzoraka, odnosno tretmana. S teorijske tačke gledišta smatra se da ima više osnovnih skupova koji imaju različite aritmetičke sredine, ali iste varijanse σ^2.

Najveća prednost analize varijanse je u tome što ona omogućava da se iz totalne varijacije izdvoje varijacije između grupa, rezultat dejstva tretmana, i varijacija unutar grupa, rezultat slučajnih kolebanja unutar svakog tretmana. Drugim rečima, da nema dejstva tretmana razlike između totalne, međugrupne i unutargrupne varijacije imale bi slučajan karakter i bile bi rezultat varijabilnosti jedinica. Međutim, dejstvo tretmana deluje na jedinice uzorka što se odražava na njihovu sredinu, pa tako i na varijaciju između grupa. To ne utiče na varijaciju unutar grupa, jer ona treba da je ista bez obzira da li se tretmani primenjuju ili ne.

Sada se mogu odrediti raspodele statistika SS_A, SS_E, SS_T.

Neka je $\bar{S}_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$ disperzija $i-tog$ uzorka, $i = 1,2,...,k.$

Tada je

$$SS_E = \sum_{i=1}^{k} n_i \bar{S}_i^2.$$

Poznato je da $\frac{n_i \bar{S}_i^2}{\sigma^2} : \chi^2_{n_i - 1}$, pa sledi

$$\frac{SS_E}{\sigma^2}=\sum_{i=1}^{k}\frac{n_i\bar{S}_i^2}{\sigma^2}:\chi^2_{\sum_{i=1}^{k}(n_i-1)}=\chi^2_{N-k}.$$

odatle je

$$E(\frac{SS_E}{\sigma^2})=(N-k),\ \text{odnosno}\ E(SS_E)=\sigma^2(N-k),$$

što znači da je

$$MS_E=\frac{SS_E}{N-k}$$

centrirana ocena za σ^2. S druge strane, SS_A ima necentriranu χ^2 raspodelu sa očekivanjem

$$E(SS_A)=(k-1)\sigma^2+\sum_{i=1}^{k}n_i\gamma_i^2.$$

Ako je hipoteza $H_0(\gamma_1=\gamma_2=...=\gamma_k=0)$ tačna, tada sledi da $\frac{SS_A}{\sigma^2}:\chi^2_{k-1}$, i da je $MS_A=\frac{SS_A}{k-1}$ centrirana ocena za σ^2. Takođe se može pokazati da su SS_A, SS_P nezavisne. Odatle sledi da ako je H_0 tačna, vrednosti MS_A i MS_E su približno iste. Za testiranje hipoteze koristićemo statistiku

$$F_{k-1,N-k}=\frac{MS_A}{MS_E}$$

Nulta raspodela ove statistike, (ako je H_0 tačna) je Fišerova sa $k-1$ i $N-k$ stepeni slobode.

Ako postoje razlike između tretmana, ukupna razlika između obeležja će nastati prvenstveno zbog dejstva tretmana, odnosno varijacija između grupa SS_A će biti velika u odnosu na varijaciju unutar grupa SS_E. Zbog toga hipotezu H_0 odbacujemo ako su vrednosti statistike $F_{k-1,N-k}$ suviše velike:

$$P\{F_{k-1,N-k}\geq c\}=\alpha,$$

gde je $c=F_{k-1,N-k,\alpha}$ kvantil reda $1-\alpha$ za Fišerovu slučajnu promenljivu $F_{k-1,N-k}$.
Označimo sa $f_{k-1,N-k}$ realizovanu vrednost statistike $F_{k-1,N-k}$. Tada verovatnoću

$$P\{F_{k-1,N-k}\geq f_{k-1,N-k}\}=p$$

nazivamo p vrednost. Hipotezu H_0 odbacujemo ako je

$$f_{k-1,N-k}\geq F_{k-1,N-k}\quad \text{odnosno ako je}\quad p\leq\alpha.$$

Konačno izračunavanje se prikazuje u Tabeli 1.7.

Tabela 1.7.

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	Realizovana vrednost	Teorijska vrednost	p -vrednost
Faktor A	SS_A	$k-1$	MS_A	$f_{k-1,N-k}$	$F_{k-1,N-k,\alpha}$	$P\{F_{k-1,N-k}\geq f_{k-1,N-k}\}=p$
Greška	SS_E	$N-k$	MS_E			
Ukupno	SS_T	$N-1$				

1.4.1.1. Post hoc analiza, odeđivanje značajnih razlika između sredina pojedinih tretmana

F-test daje odgovor o homogenosti eksperimenata u celini. On ništa ne govori o razlikama između sredina pojedinih tretmana. Normalno je očekivati da u slučaju tačne hipoteze razlike između pojedinih tretmana neće biti značajne. Često se dešava da rezultati *F*-testa ne dovede do odbacivanja nulte hipoteze, a neke razlike između sredina, obično ekstremne, pokažu se značajne. To može da bude kako rezultat slučaja tako i postojenja stvarno značajne razlike. U konačnom zaključku treba voditi računa i o onom na šta ukazuje *F*-test. Naime, ako se nulta hipoteza o homogenosti prihvata, tada svakako treba primiti sa više opreznosti eventualne značajne razlike između sredina do kojih dolazimo individualnim pogađanjem.

Na kraju celog ogleda treba da se uporede sredine pojedinih tretmana. Testiranje razlika između sredina se izvodi na više načina. Ovde ćemo objasniti test koji nije bez prigovora, iako se do sada u praksi najviše upotrebljavao.

U ovom testu se izračunava standardna greška razlike sa kojom se deli razlika između sredina dva tretmana. Broj ovih razlika koje mogu da se testiraju je $\frac{k(k-1)}{2}$ gde je k broj tretmana. Varijacija unutar grupe se smatra za najbolju ocenu slučajnih varijacija ogleda. Standardna greška je standardna devijacija distribucije sredine uzorka i izračunava se po obrascu:

$$S_{\bar{X}} = \sqrt{\frac{SS_E/(N-k)}{n}} = \sqrt{\frac{MS_E}{n}}$$

gde je *n* broj jedinica u uzorku.
Za testiranje razlike između dve sredine potrebna je standardna greška razlike i ona se izračunava po obrascu:

$$S_{(\bar{X}_i-\bar{X}_j)} = \sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

Gde su n_i i n_j brojevi u jednom i drugom uzorku. Ako je broj jedinica u oba uzorka jednak, standardna greška razlike je tada:

$$S_{(\bar{X}_i-\bar{X}_j)} = \sqrt{\frac{2MS_E}{n}}.$$

Zatim se izračunava:

$$t = \frac{\bar{X}_i - \bar{X}_j}{S_{(\bar{X}_i-\bar{X}_j)}}$$

Ova veličina upoređuje sa kritičnom vrednosti u tabelama *t*-distribucije za *N-k* stepeni slobode. To je broj stepeni slobode varijacije unutar grupa u analizi varijanse. Hipoteza o jednakosti dve sredine se prihvata uz α rizik ako je izračunato *t* manje od kritične vrednosti u tabelama *t*-distirbucije.

Za upoređivanje više razlika između sredina tretmana u praksi se obično izračunava najmanja značajna razlika, NZR.

$$NZR = t_{\alpha(N-k)} S_{(\bar{X}_i - \bar{X}_j)}\ .$$

Kada je razlika između sredine dva tretmana veća on NZR, onda se izvodi zaključak da je značajna, u protivnom da nije značajna.

Interval poverenja za svaku sredinu tretmana određuje se pomoću nejednačine:

$$\bar{X}_i - t_{\alpha(N-k)} S_{\bar{X}} < \mu < \bar{X}_i + t_{\alpha(N-k)} S_{\bar{X}}$$

Jasno je da značajna razlika u velikoj meri zavisi od veličine $S_{(\bar{X}_i - \bar{X}_j)}$. Ukoliko je $S_{(\bar{X}_i - \bar{X}_j)}$ manje, onda i manje razlike između sredina mogu biti značajne. Manja standardna greška utiče na osetljivost ogleda. Zbog toga treba pridati veliki značaj merama za smanjenje standardne greške. Praktično, to će se postići povećanjem broja ponavljanja i isključenjem iz standardne greške svih onih varijacija koje nemju slučajan karakter.

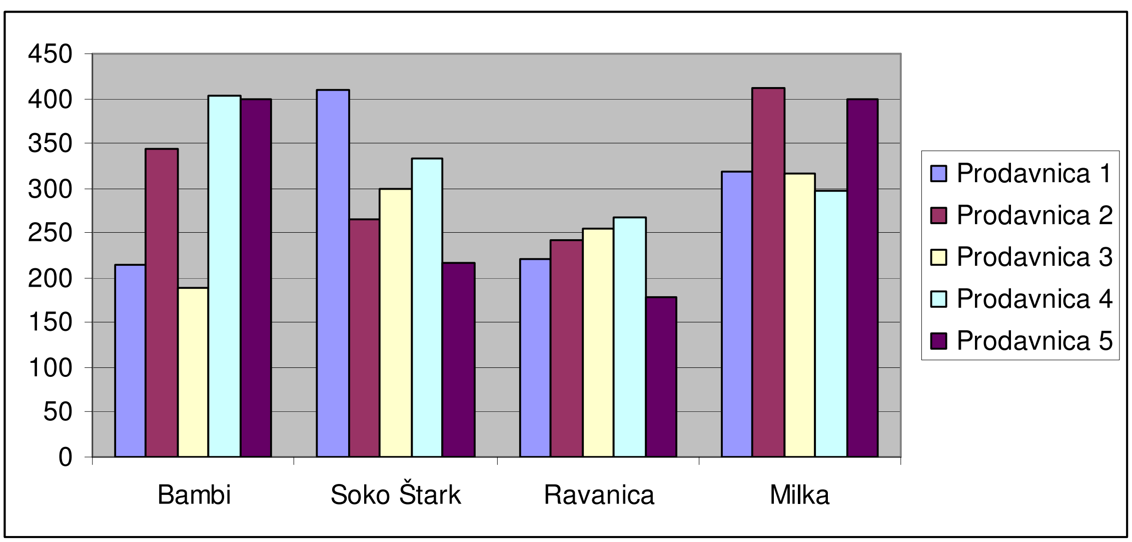
Primeri za jednofaktorsku analizu varijanse

Primer 2: Za ilustraciju ćemo uzeti primer sa ispitivanjem uticaja vrste porizvođača na obim prodaje čokolada u 5 prodavnica tokom mesec dana. [4] Dobijeni su sledeći rezultati:

Tabela 1.8. Obim prodaje čokolada

Proizvođač	*Prodavnica1*	*Prodavnica 2*	*Prodavnica3*	*Prodavnica4*	*Prodavnica5*	*Ukupno*
Bambi	215	344	189	403	399	1550
Soko Štark	410	266	300	333	217	1526
Ravanica	221	241	255	267	178	1162
Milka	319	411	316	298	400	1744

Grafik 1.5.



Izračunavanja za analizu varijanse su sledeća:

$$\sum X^2 = (215)^2 + (410)^2 + ... + (400)^2 = 1{,}902.988,$$

$$(\sum X^2) = (5{,}982)^2 = 35{,}784.324,$$

$$C = \frac{(\sum X)^2}{N} = \frac{35{,}784.324}{20} = 1{,}789.216{,}2,$$

$$SS_T = \sum X^2 - C = 1{,}902.988 - 1{,}789.216{,}2 = 113.771{,}2,$$

$$SS_A = \frac{\sum_{i=1}^{4}(\sum_{j=1}^{5} X_{ij})^2}{n} - C = \frac{(1.550)^2 + (1.526)^2 + (1.162)^2 + (1.744)^2}{5} - 1{,}789.216{,}2 = 35.375,$$

$$SS_P = SS_T - SS_A = 113.771{,}2 - 35.375 = 78.396{,}2$$

Tabela 1.9.

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	Realizovana vrednost	Teorijska vrednost	*p*-vrednost
Faktor A	35.375	3	21916.4	4.55	3,24	0,105254
Greška	78.396,2	16	4811.6			
Ukupno	113.771,2	19				

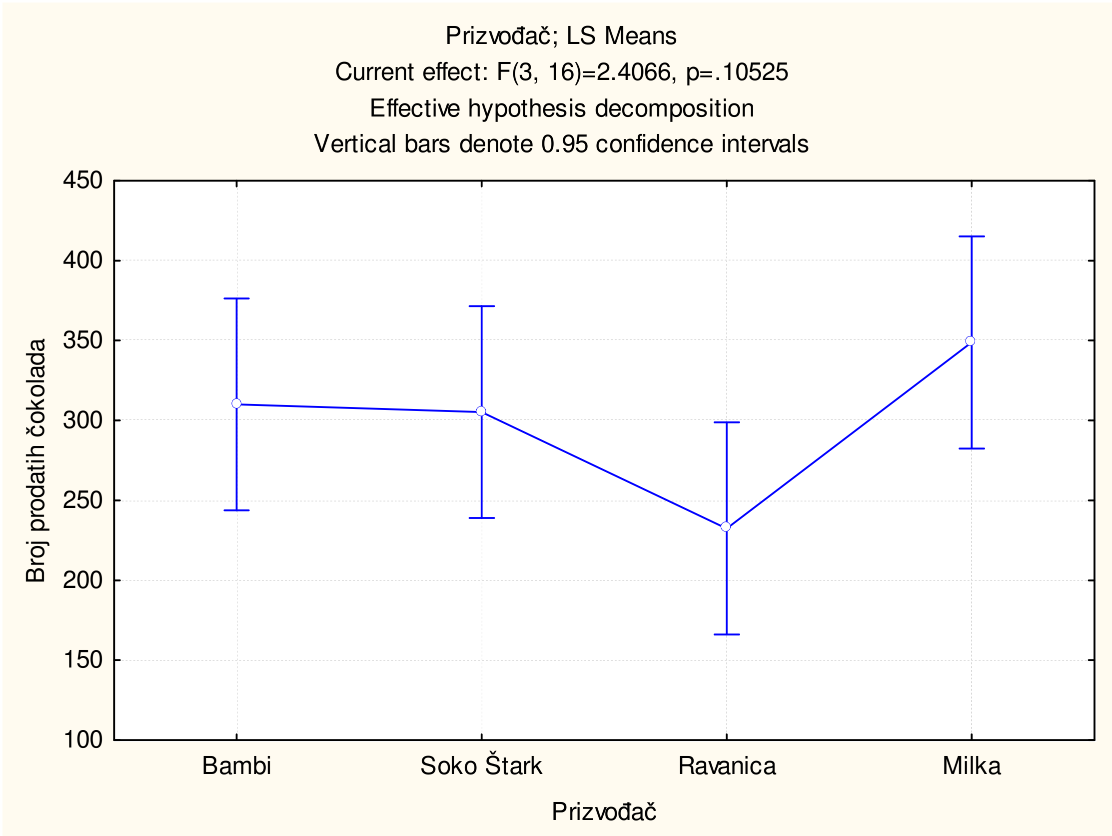
U statističkom paketu 7.0 izračunavanja izgledaju ovako:
Za izradu zadatka koristimo jednofaktorsku ANOVU, faktor proizvođač ima četiri nivoa sa po pet uzoraka na svakom nivou. Naš osnovni cilj je da utvrdimo da li vrsta proizvođača utiče na obim prodaje, odnosno da li postoje značajne razlike prema efeku proizvođač. *P-vrednost* iznosi 0.105254 što je veće od 0.05 te dolazimo do zaključka da kupci ne vode računa o vrsti proizvođača.

Tabela 1.10.

Effect	Univariate Tests of Significance for Broj prodatih èokolada (primer 1) Sigma-restricted parameterization Effective hypothesis decomposition SS	Degr. of Freedom	MS	F	p
Intercept	1789216	1	1789216	365,1611	0,000000
Proizvođač	35375	3	11792	2,4066	0,105254
Error	78397	16	4900		

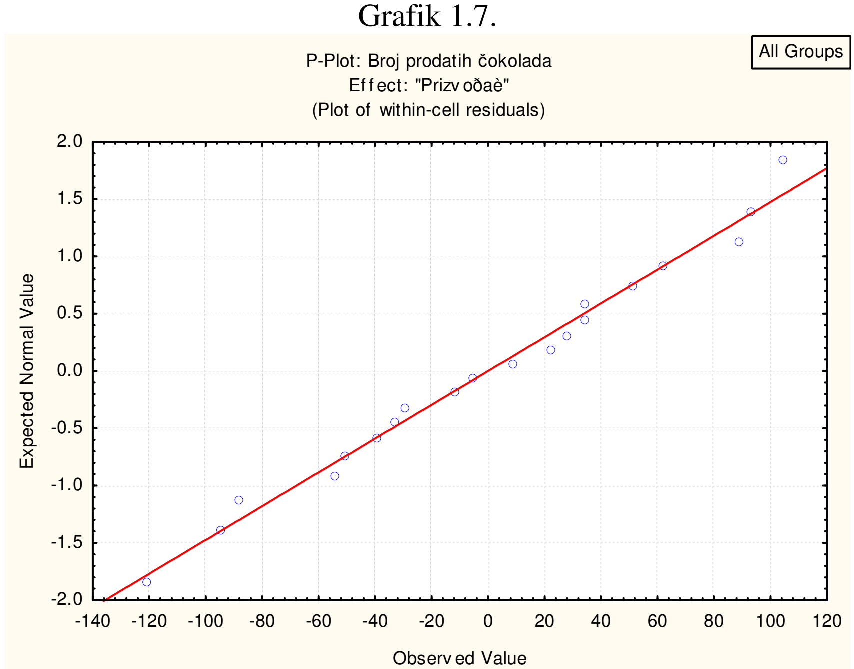
Takođe pomoću programskog paketa Statistika možemo uporediti i grafički prikazati srednje vrednosti svih efekata na sledeći način:

Grafik 1.6.



Sledećim Grafikom 1.7. proveravamo da li rezultati imaju normalnu raspodelu:

Grafik 1.7.



Post hoc analizom možemo prikazati značajne razlike i odrediti koji nivoi datih faktora odskaču od ostalih. Primenjujemo Tukeyev test i dolazimo do sledećih rezultata:

Tabela 1.11.

	Tukey HSD test; variable Broj prodatih čokolada (Primer 1) Approximate Probabilities for Post Hoc Tests Error: Between MS = 4899.8, df = 16.000				
Cell No.	Prizvođač	{1} 310.00	{2} 305.20	{3} 232.40	{4} 348.80
1	Bambi		0.999566	0.330652	0.816970
2	Soko Štark	0.999566		0.383449	0.760006
3	Ravanica	0.330652	0.383449		0.077512
4	Milka	0.816970	0.760006	0.077512	

Primer 3: Želimo da uporedimo „dužinu života'', četiri marke automobilskih guma: Pirelli, Goodyear, Bridgestone, Michelin. U tom cilju uzimamo po tri gume prve tri marke ($n_1 = n_2 = n_3 = 3$) i pet guma četvrte marke ($n_4 = 5$), ispitujemo njihovu trajnost i dobijamo sledeće vrednosti „dužine života'', (u hiljadama pređenih kilometara) [3]

Tabela 1.12.

Pirelli	**Goodyear**	**Bridgestone**	**Michelin**
35	**30**	**40**	**35**
11	**21**	**20**	**20**
20	**30**	**24**	**15**
			25
			30

Konstatujemo izvesne razlike u „dužini života'', ali se postavlja pitanje da li su te razlike značajne ili su rezultat samo slučajnih fluktuacija oko srednje vrednosti koje su jednake za sve četiri marke. Testiramo sa pragom značajnosti $\alpha = 0.05$ hipotezu o podjednakoj izdržljivosti guma svih marki, tj. hipotezu $H_0 = (\mu_1 = \mu_2 = \mu_3 = \mu_4 = 0)$ protiv alternativne hipoteze $H_1 = (\mu_j \neq 0 \text{ bar za jedno } j = 1,2,3,4)$.

Tabela 1.13.

Nivoi faktora	Merenja						
	1	2	3	4	5	Zbir	Sredina
Pirelli	35	11	20			66	22
Goodyear	30	21	30			81	27
Bridgestone	40	20	24			84	28
Michelin	35	20	15	25	30	125	25
						S=356	$\bar{X}$=25,5

$$SS_T = SS_A + SS_E, \quad SS_A = \sum_{i=1}^{k} n_i (\bar{X}_i - \bar{X})^2, \quad SS_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

SS_A=63,43 SS_E=822 SS_T=885,43

Tabela 1.14.

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	Realizovana vrednost	Teorijska vrednost	*p*-vrednost
Tretman	63,43	3	21,14	$f_{3,10}$=0,26	$F_{3,10,0.05}$=3,71	0,854536
Greška	822	10	82,2			
Ukupno	885,43	13				

Kako je $F_{3,10,0.05} = 3{,}71 > 0{,}26 f_{3,10}$ nema razloga da na osnovu prezentiranih podataka odbacimo pretpostavku o podjednakoj izdržljivosti guma sve četiri marke.
Do istog rezultata dolazimo koristeći programski paket Statistica 7.0:
Za izradu zadatka koristimo jednofaktorsku ANOVU, faktor proizvođač ima četiri nivoa sa po 3 uzoraka na prva 3 nivoa i 5 uzoraka na 4. nivou. *P-vrednost* iznosi 0,854536 što je veće od 0,05 te zaključujemo da se srednje vrednosti izdržljivosti guma ne razlikuju značajno.

Tabela 1.15.

Effect	Univariate Tests of Significance for Kilometraža (GumeJednofaktorsk Sigma-restricted parameterization Effective hypothesis decomposition SS	Degr. of Freedom	MS	F	p
Intercept	8,670000E+09	1	8,670000E+09	105,4745	0,000001
Proizvođač	6,342857E+07	3	2,114286E+07	0,2572	0,854536
Error	8,220000E+08	10	8,220000E+07		

Grafik 1.8.

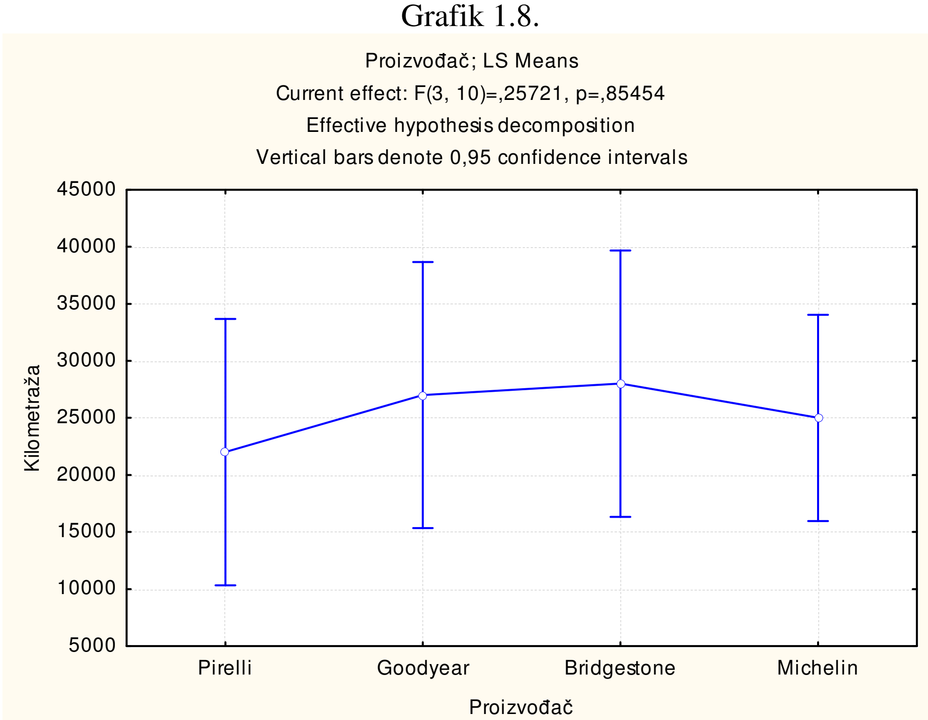
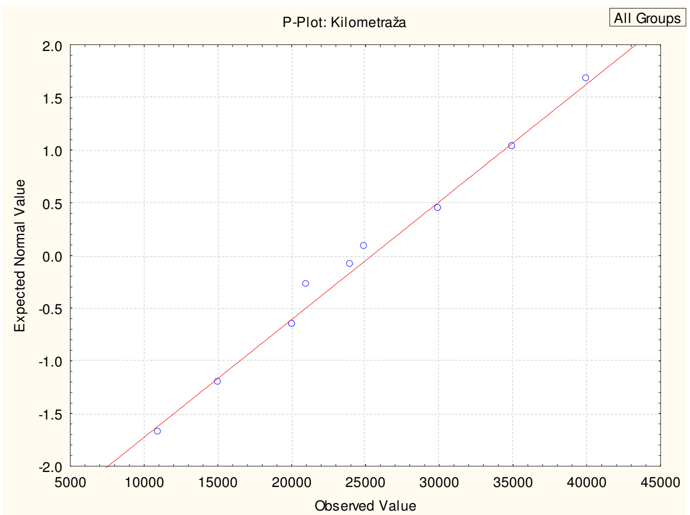


Tabela 1.16.

Cell No.	Tukey HSD test; variable Kilometraža (GumeJednofaktorsk Approximate Probabilities for Post Hoc Tests Error: Between MS = 8220E4, df = 10,000 Proizvođač	{1} 22000,	{2} 27000,	{3} 28000,	{4} 25000,
1	Pirelli		0,904119	0,848228	0,967576
2	Goodyear	0,904119		0,999114	0,989906
3	Bridgestone	0,848228	0,999114		0,967576
4	Michelin	0,967576	0,989906	0,967576	

Grafik 1.9.



1.4.2. Dvofaktorska analiza varijanse

Tehnika analize varijanse objašnjena je na slučajevima gde je postojao samo jedan kriterijum klasifikacije jedinica. Kod takvih ogleda totalna varijacija se deli na dve komponente, i to varijacije između i unutar grupe. Prva varijacija proizilazi iz primene različitih tretmana, dok je druga rezultat slučajnih kolebanja unutar svakog uzorka. Ovakvi ogledi su podesni zato što u njih može da se uključi veliki broj tretmana, a broj ponavljanja nije ograničen. Pored toga, njihova statistička analiza je vrlo jednostavna. U ogledima sa potpuno slučajnim rasporedom gubitak podatka o jednoj eksperimentalnoj jedinici nema nekog značajnijeg uticaja na samu vrednost ogleda.

Analiza varijanse se uspešno primenjuje i u ogledima s klasifikacijom jedinica po dva ili više kriterijuma. Ovde ćemo razmotriti njenu primenu kada je u pitanju klasifikacija po dve osnove.

1.4.2.1. Dvofaktorski problem na prostom uzorku

Dvofaktorski problem nastaje kada treba da se ispita uticaj dva faktora recimo A i B na obeležje X. Neka se uticaj faktora A ispituje na k, $k \geq 2$, nivoa, a faktor B na l, $l \geq 2$, nivoa. Prost slučajni uzorak obima $k \times l$ predstavlja se dvodimenzionalno, tabelom:

Tabela 1.17.

$A \downarrow / B \rightarrow$	b_1	b_2	...	b_j	...	b_k	Sredina vrste
a_1	X_{11}	X_{12}	...	X_{1j}	...	X_{1k}	$\bar{X}_{1\bullet}$
a_2	X_{21}	X_{22}	...	X_{2j}	...	X_{2k}	$\bar{X}_{2\bullet}$
.	.	.	...	.	...	.	.
.	.	.	...	.	...	.	.
.	.	.	...	.	...	.	.
a_i	X_{i1}	X_{i2}	...	X_{ij}	...	X_{ik}	$\bar{X}_{i\bullet}$
.	.		...	.	...	.	.
.	.	.	...	.	...	.	.
.	.		...	.	...	.	.
a_l	X_{l1}	X_{l2}	...	X_{lj}	...	X_{lk}	$\bar{X}_{l\bullet}$
Sredina kolone	$\bar{X}_{\bullet 1}$	$\bar{X}_{\bullet 2}$	...	$\bar{X}_{\bullet j}$	...	$\bar{X}_{\bullet k}$	

gde je X_{ij} vrednost obeležja X na elementu populacije izloženom $i-tom$ nivou faktora A i $j-om$ nivou faktora B:

$$\bar{X} = \frac{1}{kl}\sum_{i=1}^{k}\sum_{j=1}^{l} X_{ij},$$

$$\bar{X}_{i\bullet} = \frac{1}{k}\sum_{j=1}^{l} X_{ij},$$

$$\bar{X}_{\bullet j} = \frac{1}{l}\sum_{i=1}^{k} X_{ij}.$$

Slično kao i kod jednofaktorskog problema, neka je $m = E(X)$, a $m_{i\bullet} = E(X_{ij})$, $i = 1,2,...,k;$ $m_{\bullet j} = E(X_{ij})$, $j = 1,2,...,l;$ tj. $m_{i\bullet}$ je matematičko očekivanje obeležja X u populaciji koja je kod faktora A izložena samo $i-tom$ nivou, a $m_{\bullet j}$ je matematičko očekivanje obeležja X u populaciji koja je od faktora B izložena samo $j-tom$ nivou. Sa $\mu_i = m_{i\bullet} - m$ označava se efekat $i-tog$ nivoa faktora A, a sa $\lambda_j = m_{\bullet j} - m$ efekat $j-tog$ nivao faktora B.
Matematički linearni model dvofaktorske analize varijanse sa prostim uzorkom je

$$X_{ij} = m + \mu_i + \lambda_j + \varepsilon_{ij,} \quad i = 1,2,...,k; \quad j = 1,2,...,l;$$

gde su ε_{ij} nezavisne slučajne promenljive sa istom $N(0,\sigma^2)$ raspodelom, pri čemu se podrazumeva da je σ^2 nepoznato. U okviru ovog modela testiraju se sledeće nulte hipoteze
-efekti nivoa faktora *A* na obeležje *X* su bez bitnih razlika;

$$H_{0A}:(\mu_1=\mu_2=...=\mu_k=0)$$

-efekti nivoa faktora *B* na obeležje *X* su bez bitnih razlika;

$$H_{0B}:(\lambda_1=\lambda_2=...=\lambda_l=0)$$

-efekti nivoa posmatranih faktora *A* i *B* su bez bitnih razlika.

$$H_{0AB}:(\mu_1=\mu_2=...=\mu_k=\lambda_1=\lambda_2=...=\lambda_l=0)$$

protiv alternativnih hipoteza redom:

$$H_{1A}:(\exists i; i\in\{1,2,...k\},\mu_i\neq 0)$$

$$H_{1B}:(\exists j; j\in\{1,2,...l\},\lambda_j\neq 0)$$

$$H_{1AB}:(\exists(i,j); takoda(i,j)\in\{1,2,...k\}\times\{1,2,...,l\},(\mu_i,\lambda_j)\neq(0,0))$$

Za sprovođenje ovih testova potrebne su sledeće sume:

- Ukupna suma kvadrata odstupanja od srednje vrednosti
$$SS_T=\sum_{i=1}^{k}\sum_{j=1}^{l}(X_{ij}-\bar{X})^2$$
- Suma kvadrata za faktor A:
$$SS_A=l\sum_{i=1}^{k}(\bar{X}_{i\bullet}-\bar{X})^2;$$
- Suma kvadrata za faktor B:
$$SS_B=k\sum_{j=1}^{l}(\bar{X}_{\bullet j}-\bar{X})^2;$$
- Rezidualna suma kvadrata, suma kvadrata greške:

$$SS_E=SS_T-SS_A-SS_B.$$

Pod uslovom da su nulte hipoteze tačne, statistike

$$F_{k-1,(k-1)(l-1)} = \frac{(k-1)(l-1)SS_A}{(k-1)SS_E} = \frac{MS_A}{MS_E} \ (1)$$

$$F_{l-1,(k-1)(l-1)} = \frac{(k-1)(l-1)SS_B}{(l-1)SS_E} = \frac{MS_B}{MS_E} \ (2)$$

imaju Fišerovu raspodelu sa naznačenim brojem stepeni slobode. Kao i kod jednofaktorske analize varijanse, ako je tačna hipoteza H_0, suma kvadrata odstupanja SS_A, SS_B nastale zbog dejstva faktora *A* i *B* neće biti velike u poređenju sa sumom kvadrata podgreške SS_E. Ako su nulte hipoteze netačne, količnici (1) i (2) će biti veliki. Da bi smo utvrdili da li su razlike značajne, označićemo sa $F_{n,m,1-\alpha}$ kvantil reda $1-\alpha$ Fišerove raspodele, a sa $f_{k-1,(k-1)(l-1)}$ i $f_{l-1,(k-1)(l-1)}$ realizovane vrednosti poslednjih statistika. Dalje, neka su *p*-vrednosti

$$p_1 = P\{F_{k-1,(k-1)(l-1)} \geq f_{k-1,(k-1)(l-1)}\}$$
$$p_2 = P\{F_{l-1,(k-1)(l-1)} \geq f_{l-1,(k-1)(l-1)}\}.$$

Tada je postupak odbacivanja, odnosno prihvatanja hipoteze sledeći:

Tabela 1.18.

Odbacujemo	Prihvatamo
$H_{0A}(\mu_1 = \mu_2 = ... = \mu_k = 0)$	$H_{1A}(\exists i : i \in \{1,2,...,k\}, \mu_i \neq 0)$
ako je $f_{k-1,(k-1)(l-1)} \geq F_{k-1,(k-1)(l-1),1-\alpha}$, odnosno $p_1 < \alpha$	
$H_{0B}(\lambda_1 = \lambda_2 = ... = \lambda_l = 0)$	$H_{1B}(\exists j : j \in \{1,2,...,k\}, \lambda_j \neq 0)$
ako je $f_{l-1,(k-1)(l-1)} \geq F_{l-1,(k-1)(l-1),1-\alpha}$, odnosno $p_2 < \alpha$	

Tabela 1.19.

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	Realizovana vrednost	Teorijska vrednost	*p*-vrednost
Faktor A	SS_A	$k-1$	MS_A	$f_{k-1,(k-1)(l-1)}$	$F_{k-1,(k-1)(l-1),1-\alpha}$	p_1
Faktor B	SS_B	$l-1$	MS_B	$f_{l-1,(k-1)(l-1)}$	$F_{k-1,(k-1)(l-1),1-\alpha}$	p_2
Greška	$SS_{E`}$	$(k-1)(l-1)$	MS_E			
Ukupno	SS_T	$N-1$				

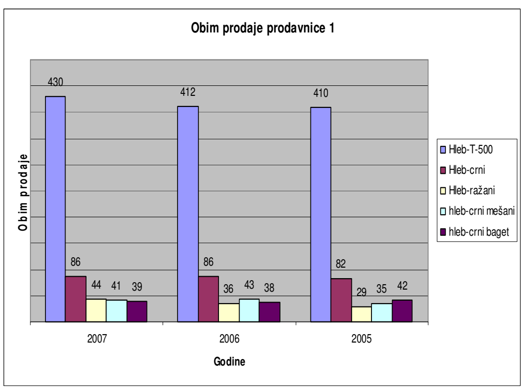
Primeri za dvofaktorsku analizu varijanse:

Primer 4: Podaci se odnose na obim prodaje hleba u dve pekare na dan 23.11. u tri različite godine (2005, 2006 i 2007.god)[5] i na osnovu metoda analize varijanse dva faktora varijabiliteta utvrdiće se da li postoje statistički značajne razlike u proizvodnji po godinama odnosno prema vrsti.

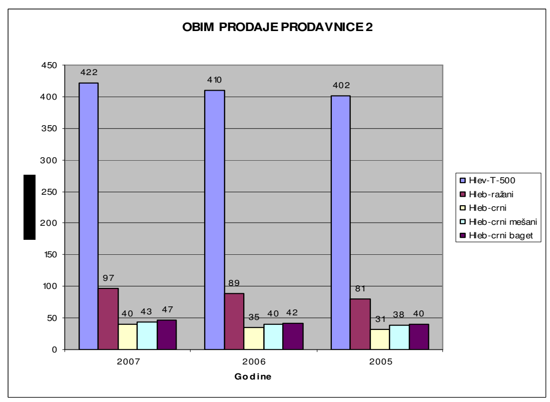
Tabela 1.20. Prikupljeni podaci

Naziv proizvoda	Proizvodnja na dan 23.11.2007. god.	Proizvodnja na dan 23.11.2006. god.	Proizvodnja na dan 23.11.2005. god.
Hleb T-500	422/430	410/412	402/410
Hleb Polje	97/86	89/86	81/82
Crni specijal hleb	40/44	35/36	31/29
Crni mešani hleb	43/41	40/43	38/35
Crni baget	47/39	42/38	40/42
Ukupno	649/640	616//615	592//598

Grafik 1.10.



Grafik 1.11.



Obrada podataka u statističkom paketu Statistica 7.0 izgleda ovako:

Tabela 1.21.

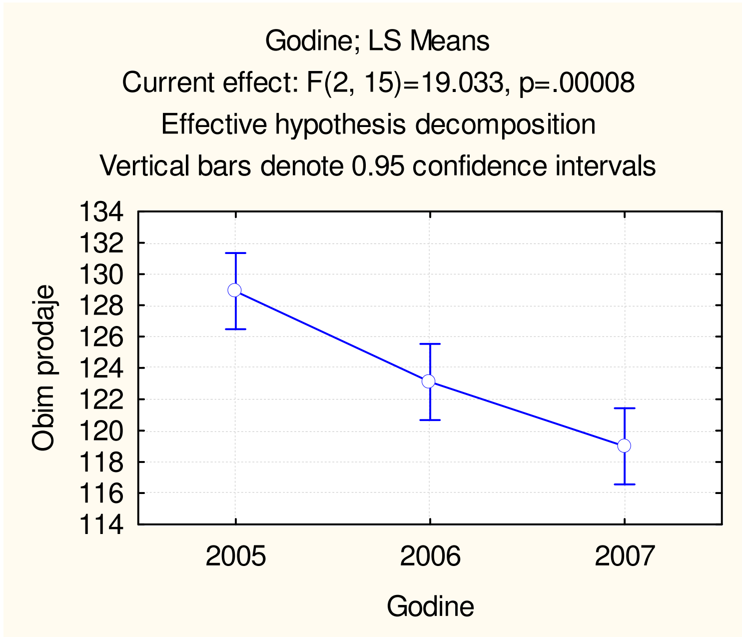
Effect	Univariate Tests of Significance for Obim prodaje (Spreadsheet14) Sigma-restricted parameterization Effective hypothesis decomposition SS	Degr. of Freedom	MS	F	p
Intercept	458803.3	1	458803.3	35292.56	0.000000
Godine	494.9	2	247.4	19.03	0.000077
Vrsta hleba	644024.3	4	161006.1	12385.08	0.000000
Godine*Vrsta hleba	230.5	8	28.8	2.22	0.087593
Error	195.0	15	13.0		

U Tabeli 1.21. možemo primetiti da godine i vrsta hleba utiču na obim prodaje jer su tu p-vrednosti manje od 0.05 dok njihovo interaktivno dejstvo ne utiče na obim prodaje

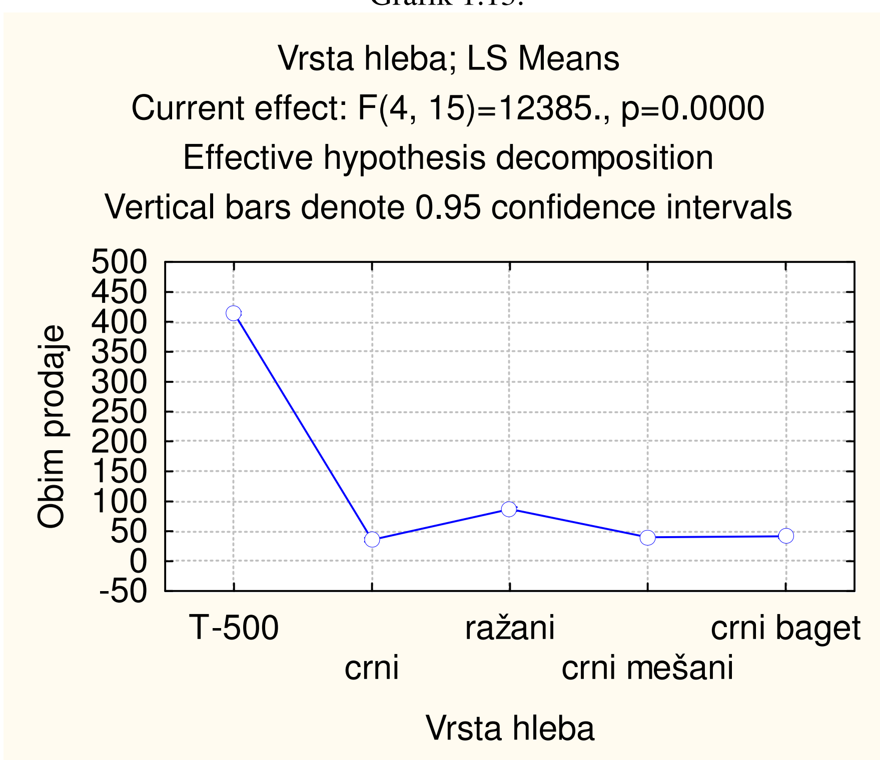
$p < 0.05$.

I ovde kao i kod jednofaktorske analize možemo uporediti i grafički prikazati srednje vrednosti

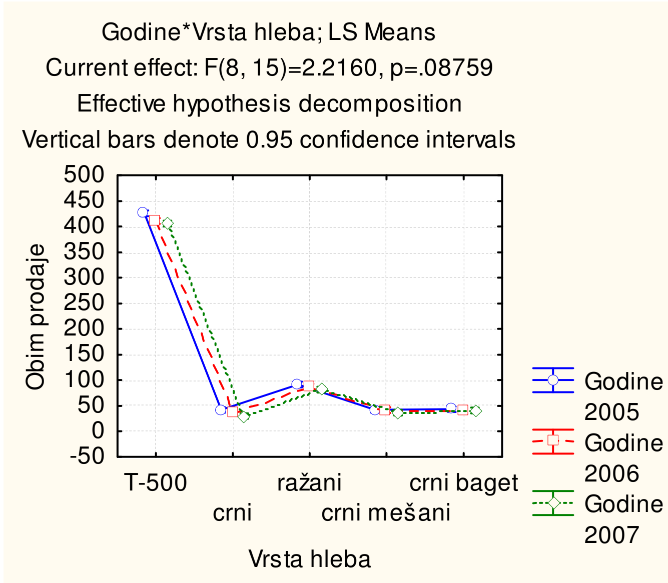
Grafik 1.12.



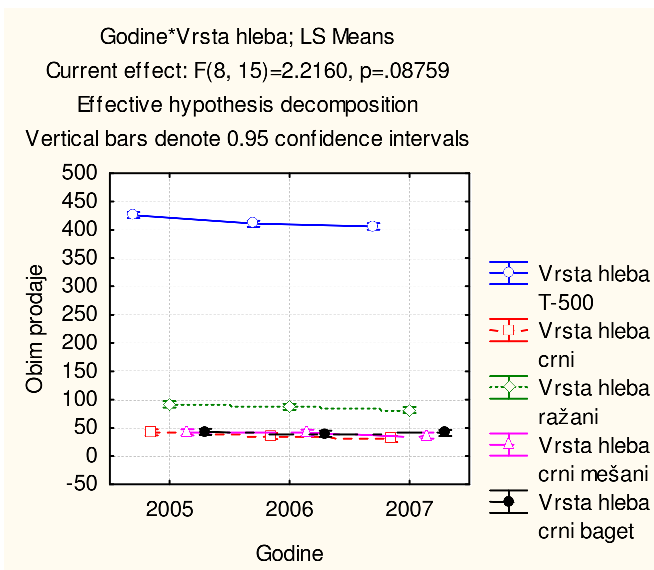
Grafik 1.13.



Grafik 1.14.

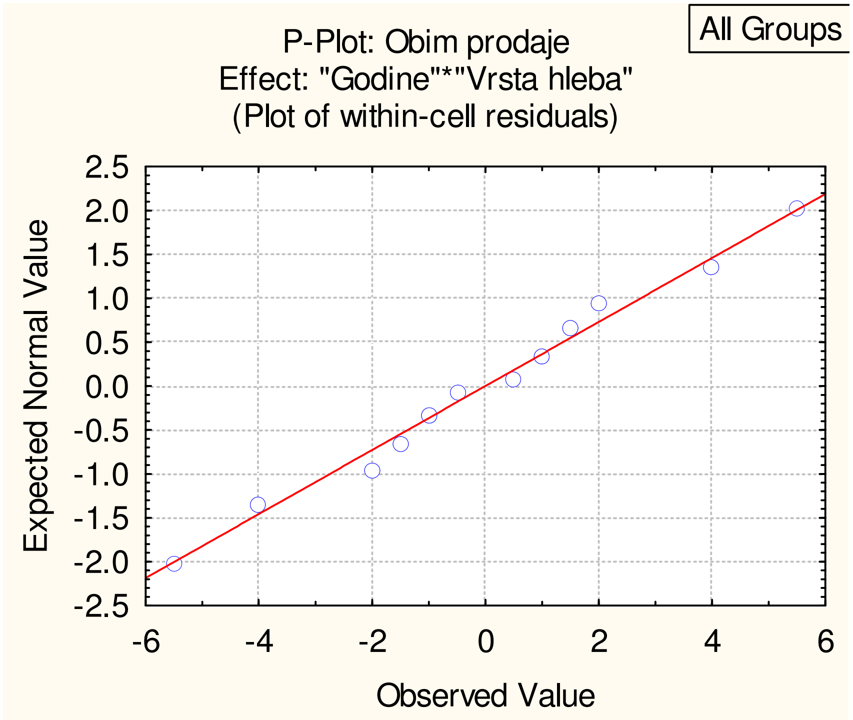


Grafik 1.15.



Možemo grafički proveriti da li rezultati imaju normalnu raspodelu:

Grafik 1.16.


All Groups
P-Plot: Obim prodaje
Effect: "Godine"*"Vrsta hleba"
(Plot of within-cell residuals)
Expected Normal Value
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-6
-4
-2
0
2
4
6
Observed Value

1.4.2.2. Dvofaktorski problemi na uzorku sa ponavljanjem

Uzorak može da ima i više elemenata koji odgovaraju svakom uređenom paru nivoa (i, j) postmatranih faktora, u tom slučaju radi se o uzorku sa ponavljanjem pri čemu broj elemenata (ponavljanja) ne mora biti jednak u svakoj ćeliji. Tada je prikaz u tabeli (za slučaj jednakog broja elemenata, r) trodimenzionalni i omogućava ispitivanje međuzavisnosti dva posmatrana faktora:

Tabela 1.22.

A/B	1	2	...	l
1	X_{111} X_{112} ... X_{11r}	X_{121} X_{122} ... X_{12r}	... ...	X_{1l1} X_{1l2} ... X_{1lr}
2	X_{211} X_{212} ... X_{21r}	X_{221} X_{222} ... X_{22r}	... ...	X_{2l1} X_{2l2} ... X_{2lr}
...	...	...	...	...
k	X_{k11} X_{k12} ... X_{k1r}	X_{k21} X_{k22} ... X_{k2r}	... ...	X_{kl1} X_{kl2} ... X_{klr}

Tada je $E(X_{ij\alpha}) = m_{ij}$, $\alpha = 1,2,...,r$, a međusobni efekat faktora *A* na nivou i i faktora *B* na nivou j predstavlja se sa:

$$\eta_{ij} = m_{ij} - (m + \mu_i + \lambda_j)$$

i važi:

$$\sum_{i=1}^{k} \mu_i = \sum_{j=1}^{l} \lambda_j = \sum_{i=1}^{k} \sum_{j=1}^{l} \eta_{ij} = 0$$

Matematički linearni model dvofaktorske analize varijense na uzorku sa ponavjlanjem je:

$$X_{ij\alpha} = m + \mu_i + \lambda_j + \eta_{ij} + \varepsilon_{ij\alpha}$$

gde su $\varepsilon_{ij\alpha}$ nezavisne jednako raspodeljene slučajne promenljive sa $N(0, \sigma^2)$ raspodelom, gde se podrazumeva da je σ^2 nepoznato.

U okviru ovog modela testiraju se sledeće nulte hipoteze:

- Efekti nivoa faktora *A* na obeležje *X* su bez bitnih razlika;

$$H_{0A}:(\mu_1=\mu_2=...=\mu_k=0),$$

- Efekti nivoa faktora *B* na obeležje *X* su bez bitnih razlika;

$$H_{0B}:(\lambda_1=\lambda_2=...=\lambda_l=0),$$

- Nema interaktivnog djestva faktora *A* i *B* na obeležje *X*;

$$H_{0AB}:(\eta_{ij}=0,\forall(i,j)i=1,2,...,k;\ j=1,2,...,l),$$

Protiv alternativnih redom:

$$H_{1A}:(\exists i:i\in\{1,2,...,k\},\mu_i\neq 0)$$
$$H_{1B}:(\exists j:\ j\in\{1,2,...,l\},\lambda_j\neq 0)$$
$$H_{1AB}:(\exists(i,j):i\in\{1,2,...,k\},\ j\in\{1,2,...,l\},\eta_{ij}\neq 0).$$

Za sprovođenje ovih testova definišu se sledeće statistike:

- Uzoračka sredina celog uzorka

$$\bar{X}=\frac{1}{klr}\sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}X_{ij\alpha};$$

- Uzoračka sredina ćelije ij

$$\bar{X}_{ij}=\frac{1}{r}\sum_{\alpha=1}^{r}X_{ij\alpha};$$

- Uzoračka sredina na nivou i faktora *A*

$$\bar{X}_{i\bullet}=\frac{1}{lr}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}X_{ij\alpha};$$

- Uzoračka sredina na nivou j faktora *B*

$$\bar{X}_{\bullet j}=\frac{1}{kr}\sum_{i=1}^{k}\sum_{\alpha=1}^{r}X_{ij\alpha};$$

- Ukupna suma kvadrata odstupanja od srednje vrednosti

$$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}(X_{ij\alpha} - \bar{X})^2;$$

- Suma kvadrata za faktor A

$$SS_A = \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}(\overline{X}_{i\bullet} - \overline{X})^2 = lr\sum_{i=1}^{k}(\overline{X}_{i\bullet} - \overline{X})^2 ;$$

- Suma kvadrata za faktor B

$$SS_B = \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}(\overline{X}_{\bullet j} - \overline{X})^2 = kr\sum_{j=1}^{l}(\overline{X}_{\bullet j} - \overline{X})^2 ;$$

- Suma kvadrata interaktivnog dejstva faktora A i B

$$SS_{AB} = \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}(\overline{X}_{ij} - \overline{X}_{i\bullet} - \overline{X}_{\bullet j} + \overline{X})^2 = r\sum_{i=1}^{k}\sum_{j=1}^{l}(\overline{X}_{ij} - \overline{X}_{i\bullet} - \overline{X}_{\bullet j} + \overline{X})^2;$$

- Suma kvadrata greške

$$SS_E = \sum_{i=1}^{k}\sum_{j=1}^{l}\sum_{\alpha=1}^{r}(X_{ij\alpha} - \bar{X}_{ij})^2.$$

Može se pokazati da je:

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E$$

Pod uslovom da su nulte hipoteze tačne, statistike

$$F_{k-1,kl(r-1)} = \frac{kl(r-1)SS_A}{(k-1)SS_E}$$

$$F_{l-1,kl(r-1)} = \frac{kl(r-1)SS_{AB}}{(l-1)SS_E}$$

$$F_{(k-1)(l-1),kl(r-1)} = \frac{kl(r-1)SS_{AB}}{(k-1)(l-1)SS_E}$$

imaju Fišerovu raspodelu sa naznačenim brojem stepeni slobode. Da bismo utvrdili da li su razlike značajne, označimo sa $F_{n,m,1-\alpha}$ kvantil reda $1-\alpha$ Fišerove rapodele, a sa $f_{k-1,kl(r-1)}$, $f_{l-1,kl(r-1)}$ i $f_{(l-1)(k-1),kl(r-1)}$ realizovane vrednosti poslednjih statistika. Dalje, neka su p-vrednosti

$$p_1 = P\{F_{k-1,kl(r-1)} \geq f_{k-1,kl(r-1)}\},$$
$$p_2 = P\{F_{l-1,kl(r-1)} \geq f_{l-1,kl(r-1)}\},$$
$$p_3 = P\{F_{(k-1)(l-1),kl(r-1)} \geq f_{(k-1)(l-1),kl(l-1)}\}$$

Tada je postupak odbacivanja, odnosno prihvatanja hipoteza sledeći:

Tabela 1.23.

Odbacujemo	Prihvatamo
$H_{0A}(\mu_1 = \mu_2 = ... = \mu_k = 0)$	$H_{1A}(\exists i : i \in \{1,2,...,k\}, \mu_i \neq 0)$
ako je $f_{k-1,kl(r-1)} \geq F_{k-1,kl(r-1),1-\alpha}$, odnosno $p_1 < \alpha$	
$H_{0B}(\lambda_1 = \lambda_2 = ... = \lambda_l = 0)$	$H_{1B}(\exists j : j \in \{1,2,...,k\}, \lambda_j \neq 0)$
ako je $f_{l-1,kl(r-1)} \geq F_{l-1,kl(r-1),1-\alpha}$, odnosno $p_2 < \alpha$	
$H_{0AB}(\eta_{ij} = 0, \forall (i,j) i = 1,2,...,k; j = 1,2,...,l)$	$H_{1AB}(\exists (i,j) i = 1,2,...,k; j = 1,2,...,l, \eta_{ij} \neq 0)$
ako je $f_{(l-1)(k-1),kl(r-1)} \geq F_{(l-1)(k-1),kl(r-1),1-\alpha}$, odnosno $p_3 < \alpha$	

Tabela 1.24. Analiza varijanse

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	Realizovana vrednost	Teorijska vrednost	p-vrednost
Faktor A	SS_A	$k-1$	MS_A	$f_{k-1,(k-1)(l-1)}$	$F_{k-1,(k-1)(l-1),1-\alpha}$	p_1
Faktor B	SS_B	$l-1$	MS_B	$f_{l-1,(k-1)(l-1)}$	$F_{k-1,(k-1)(l-1),1-\alpha}$	p_2
Interakcija faktora A i B	SS_{AB}	$(k-1)(l-1)$	MS_{AB}	$f_{(l-1)(k-1),kl(r-1)}$	$F_{(l-1)(k-1),kl(r-1),1-\alpha}$	p_3
Greška	SS_E	$kl(r-1)$	MS_E			
Ukupno	SS_T	$N-1$				

Primer 5: U primeru sa ogledom sa ispitivanjem produženog delovanja supersulfata (a_1), Tomasovog brašna (a_2) i njihovog unošenja pri dubini od 15cm (b_1) i 35cm (b_2) u drugoj godini po primeni đubriva na prinos suve smeše graška i raži, [5] izračunavanja izgledaju ovako:

Tabela 1.25.

Ponavljanje	a_1		a_2	
	b_1	b_2	b_1	b_2
1	5,1	4,8	5,9	4,5
2	4,8	5,3	5,5	4,6
3	5,0	5,7	5,2	4,8
4	5,3	5,3	3,5	3,4
5	4,5	4,5	4,5	4,6
Zbir	24,7	25,6	24,6	21,9
$\bar{X}$	4,94	5,12	4,92	4,38

Sredine nivoa faktora:

$$a_1 = 24{,}7 + 25{,}6 = 50{,}3$$
$$a_2 = 24{,}6 + 21{,}9 = 46{,}5$$
$$b_1 = 24{,}7 + 24{,}6 = 49{,}3$$
$$b_2 = 25{,}6 + 21{,}9 = 47{,}5$$

Sume kvadrata za analizu varijanse izračunavaju se na uobičajen način:

Suma kvadrata totala $(5{,}2)^2 + (4{,}8)^2 + \ldots + (4{,}6)^2 - \frac{(96{,}8)^2}{20} = 7{,}6080$

Suma kvadrata tretmana $\frac{(24{,}7)^2+(25{,}6)^2+(24{,}6)^2+(21{,}9)^2}{5} - \frac{(96{,}8)^2}{20} = 1{,}5320$

Suma kvadrata greške $7{,}6080 - 1{,}5320 = 6{,}0760$

Pored toga, suma kvadrata tretmana dalje se deli na sume kvadrata koje proiyilaze iz dejstva glavnih efekata *A* i *B* i njihove interakcije *AB*. Izračunavaju se na sledeći način:

Suma kvadarata glavnog efekta $A = \frac{(50{,}3)^2+(46{,}5)^2}{10} + \frac{(96{,}8)^2}{20} = 0{,}7220$

Suma kvadarata glavnog efekta $B = \frac{(49{,}3)^2+(47{,}5)^2}{10} + \frac{(96{,}8)^2}{20} = 0{,}1520$

Suma kvadrata interakcije $AB = 1{,}5320 - 0{,}7220 - 0{,}1520 = 0{,}6480.$

Na ovaj način dobili smo potrebne sume kvadrata za analizu varijanse prikazane u tabeli.
Pomoću F-testa testira se dejstvo glavnih efekata i interakcije i to:

$$F_A = \frac{0{,}722}{0{,}380} = 1{,}90,$$
$$F_B = \frac{0{,}162}{0{,}380} = 0{,}43$$
$$F_{AB} = \frac{0{,}648}{0{,}380} = 1{,}70$$

Tabela 1.26. Analize varijanse

Izvori varijacije	Stepeni slobode	Zbir kvadrata	Sredina zbira kvadarata	Realizovana vrednost	Teorijska vrednost	p-vrednost
A	1	0,7220	0,7220	1,90	4,94	0,187
B	1	0,1620	0,1620	0,43	4,94	0,52
AB	1	0,6480	0,6480	1,70	4,94	0,21
Greška	16	6,0760	0,380			
Ukupno	19	7,6080				

Prema tome, ni kod jednog efekta nema osnova za odbacivanje nulte hipoteze.

Sada ćemo se, radi ilustracije, zadržati na interpretaciji rezultata ogleda, iako treba imati u vidu da ni jedan efekat nije značajan. U tu svrhu poslužićemo se sledećom Tabelom 1.27.

Tabela 1.27.

Vrsta mineralnog đubriva	Dubina unošenja 15(b_1)	35(b_2)	Aritmetička sredina	$b_1 - b_2$
Superfosfat, a_1	4,94	5,12	5,03	-0,18
Tomasovo brašno, a_2	4,92	4,38	4,65	0,54
Aritmetička sredina	4,93	4,75	4,84	
$a_1 - a_2$	0,02	0,74		

Iz ove Tabele1.27. se vidi da se, kad je upotrebljen superfosfat pri dubini unošenja od 15cm u poređenju sa dubinom od 35cm, prosečan prinos smanjuje za 0,18t/ha, međutim, kad je upotrbljeno Tomasovo brašno, razlika u prosečnom prinosu kod dubine od 15cm je veća za 0,54 nego kod dubine od 35cm. Sa druge strane, upotreba supersulfata pri unošenju od 15cm povećala je prinos za 0,02t/ha, dok je pri dubini unošenja od 35cm ta razlika 0,74.

Rezultati ogleda se mogu inerpretirati i na drugi način. Pre svega može se postaviti pitanje: da li su rezultati u primeni pojedinih tretmana nezavisni? Da su nezavisni rezultati u pogledu dubine unošenja bez obzira da li je primenjen supersulfat ili Tomasovo brašno bili bi isti. Isto bi bilo i za vrstu đubriva nezavisno od dubine unošenja. U najboljem slučaju, kad su upoređenja nezavisna, imamo da je ili $a_1 - a_2 = 0$ ili $b_1 - b_2 = 0$. Ako je ova hipoteza tačna, kao što je slučaj u našem primeru, neslaganje je rezultat slučjanih varijacija. Kod tačne hipoteze vrednostima razlike (-0,18; 0,54; 0,02; 0,74) ne treba pridavati značaj. Sredina razlike od -0,18 i 0,54 je 0,18 a sredina razlike 0,02 i 0,74 je 0,38. Ove dve vrednosti predstavljaju glavne efekte dubine unošenja I vrste đubriva. Kod nezavisnosti ova dva faktora možemo da smatramo da dubina unošenja od 15cm daje veći prinos u proseku za 0,18t/ha, a da primena superfosfata daje u proseku veći prinos za 0,38t/ha.

Postoje međutim, mogućnosti da faktori u ogledu nisu nezavisni. To će biti onda kad dubina unošenja dovodi do promene u prinosu, ali ta promena je značajna uz primenu jednog od đubriva. Može se tako tražiti odgovor da li je dubina unošenja imala značajnijeg uticaja na veći prinos kad je bio upotrebljen superfosfat. Razlika između jednostavnih efekata za dubinu unošenja -0,18-0,54=-0,72 odnosno za vrstu đubriva 0,02-0,74=-0,72 predstavlja interakciju dubina unošenja i vrste đubriva. Ta se razlika ponekad deli sa dva, tj. -0,72/2=-0,36. Interakcija ovde pokazuje određen efekat na rezultate ogleda u ostvarenoj kombinaciji dva faktora. Raspored stepeni slobode takvog ogleda je sledeći:

Tabela 1.28.

Izvori varijacije	Stepeni slobode
Ponavljanje	4
Tretmani (A,B,AB)	3
Greška	12
Ukupno	19

Stepeni slobode ovde mogu da se podele na glavne efekte A i B i interakciji AB, svaki efekat sa po stepenom slobode 1. Na osnovu dobijenih rezultata u faktorijalnom ogledu mi možemo da zaključimo da li postoji ili ne interakcija među faktorima, tj. da li su zavisni ili nezavisni.

Za primer sa ogledom sa ispitivanjem produženog delovanja supersulfata (a_1), Tomasovog brašna (a_2) i njihovog unošenja pri dubini od 15cm (b_1) i 35cm (b_2) u drugoj godini po primeni đubriva na prinos suve smeše graška i raži, obrada podataka u statističkom paketu Statistica 7.0 izgleda ovako:

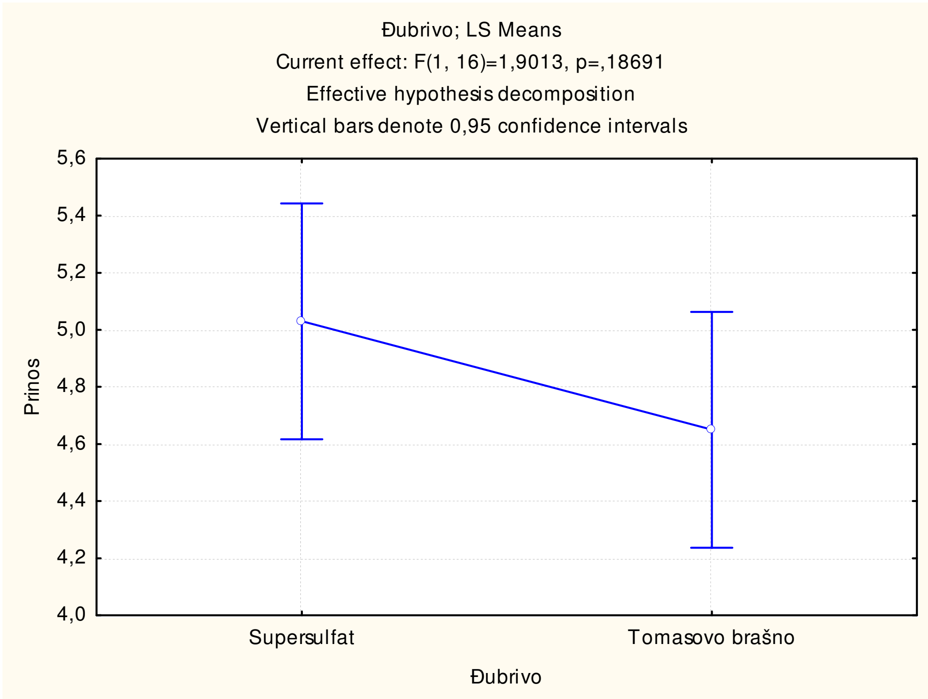
Tabela 1.29

Effect	Univariate Results for Each DV (Spreadsheet1) Sigma-restricted parameterization Effective hypothesis decomposition				
	Degr. of Freedom	Prinos SS	Prinos MS	Prinos F	Prinos p
Intercept	1	468,5120	468,5120	1233,738	0,000000
Đubrivo	1	0,7220	0,7220	1,901	0,186911
Dubina unošenja	1	0,1620	0,1620	0,427	0,522939
Đubrivo*Dubina unošenja	1	0,6480	0,6480	1,706	0,209923
Error	16	6,0760	0,3798		
Total	19	7,6080			

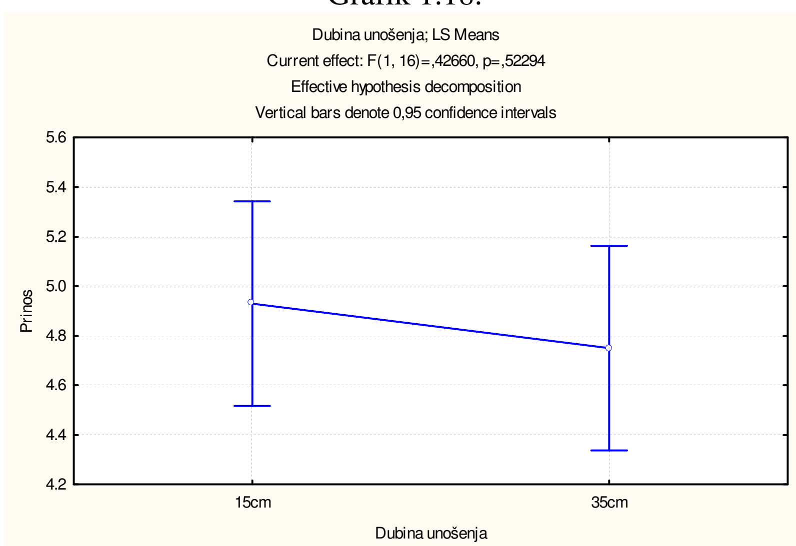
Tabela 1.30.

Effect	Descriptive Statistics (Spreadsheet1)							
	Level of Factor	Level of Factor	N	Prinos Mean	Prinos Std.Dev.	Prinos Std.Err	Prinos -95,00%	Prinos +95,00%
Total			16	4,962500	0,587509	0,146877	4,649439	5,275561
Đubrivo	Supersulfat		10	5,030000	0,386005	0,122066	4,753869	5,306131
Đubrivo	Tomasovo brašno		6	4,850000	0,861974	0,351900	3,945413	5,754587
Dubina unošenja	15cm		10	4,930000	0,661732	0,209258	4,456626	5,403374
Dubina unošenja	35cm		6	5,016667	0,491596	0,200693	4,500768	5,532565
Đubrivo*Dubina unošenja	Supersulfat	15cm	5	4,940000	0,304959	0,136382	4,561343	5,318657
Đubrivo*Dubina unošenja	Supersulfat	35cm	5	5,120000	0,471169	0,210713	4,534967	5,705033
Đubrivo*Dubina unošenja	Tomasovo brašno	15cm	5	4,920000	0,944458	0,422374	3,747301	6,092699
Đubrivo*Dubina unošenja	Tomasovo brašno	35cm	1	4,500000				

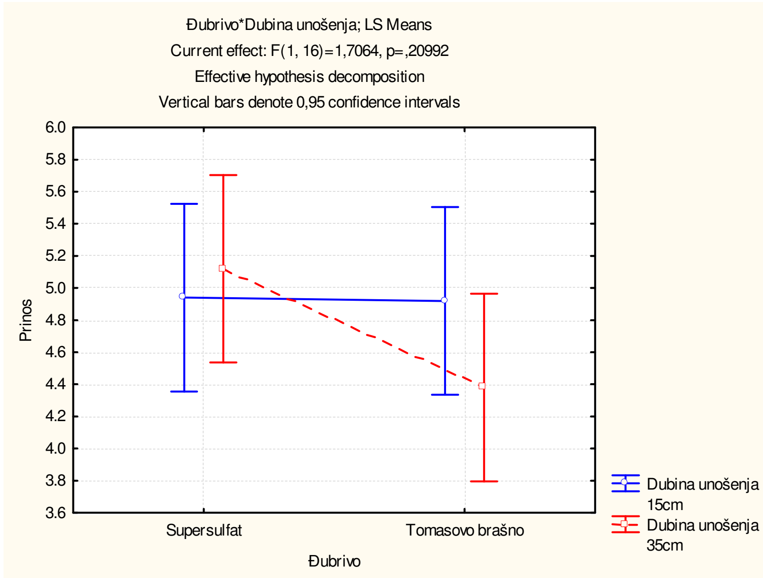
Grafik 1.17.



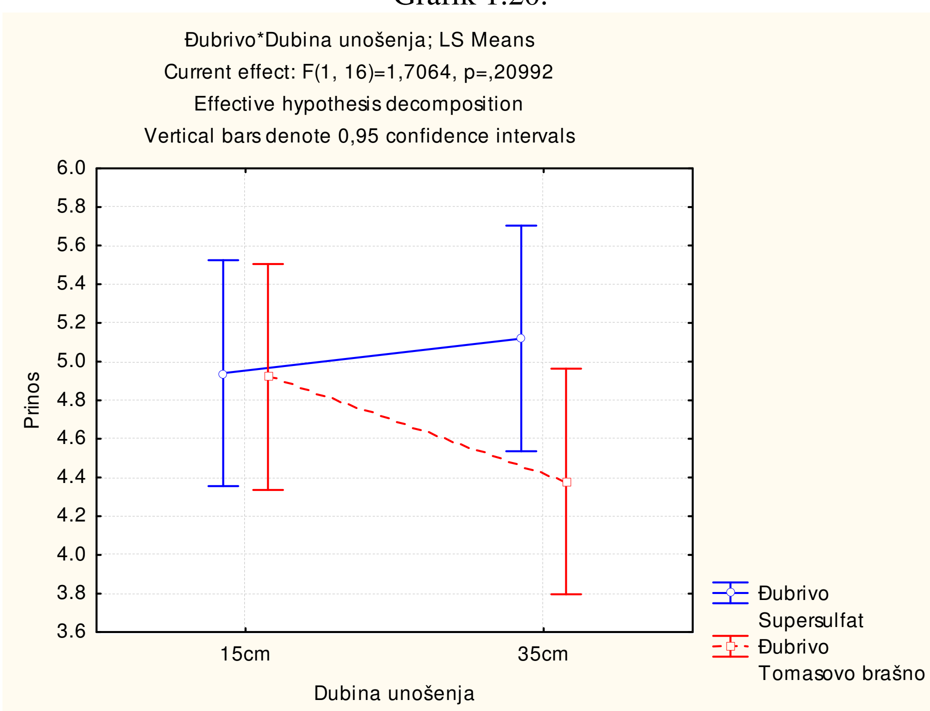
Grafik 1.18.



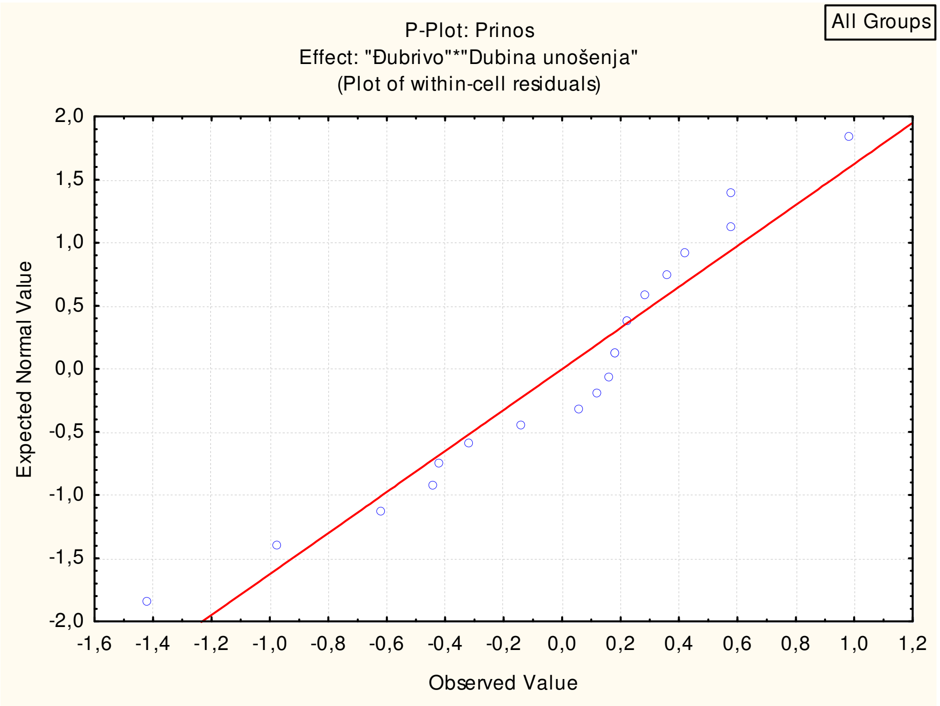
Grafik 1.19.



Grafik 1.20.



Grafik 1.21.


P-Plot: Prinos
Effect: "Đubrivo"*"Dubina unošenja"
(Plot of within-cell residuals)
All Groups
Expected Normal Value
Observed Value
2,0
1,5
1,0
0,5
0,0
-0,5
-1,0
-1,5
-2,0
-1,6 -1,4 -1,2 -1,0 -0,8 -0,6 -0,4 -0,2 0,0 0,2 0,4 0,6 0,8 1,0 1,2

Glava 2

2. Faktorijalni ogledi sa dva faktora

Najjednostavniji slučaj faktorijalnog ogleda jeste onaj sa samo dva faktora. Neka faktor *A* ima a nivoa ili tretmana i neka faktor B ima b nivoa. Tako svako ponavljanje eksperimenta, sadržaće sve moguće *ab* kombinacije nivoa. U opštem slučaju uzimaćemo da je broj ponavljanja ogleda *n.* [6]
Ovaj tip faktorijalnog eksperimenta razmotrićemo kroz primer. Naime, u pitanju je testiranje dužine trajanja određene baterije. Za dva faktora u ovom eksperimentu, čiji se uticaj želi ispitati uzeti su vrsta zaštitnog materijala baterije i različite temperature kojima je izložena baterija. Dakle, želi se proveriti da li ovi faktori pojedinačno ili zajednički utiču na trajanje baterije. I jedan i drugi faktor u ovom ogledu će imati po tri nivoa, tj. tri nivoa materijala (tip 1, tip 2, tip 3) i tri različite temperature (15°F , 70°F i 125°F), s obzirom na to da su to vrlo bitne temperature kad je u pitanju uobičajena upotreba baterije od strane korisnika. Nakon slučajnog testiranja četiri baterije u svim mogućim kombinacijama vrste materijala i spoljašnje temperature, tj. nakon 36 testova dobijeni su sledeći rezultati, odnosno uzorak:

Tabela 2.1. Trajanje baterije u časovima (h)

		Spoljašnja temperatura					
		15°F		**70°F**		**125°F**	
Vrsta zaštitnog materijala	**Tip 1**	130	155	34	40	20	70
		74	180	80	75	82	58
	Tip 2	150	188	126	122	25	70
		159	126	106	115	58	45
	Tip 3	138	110	174	120	96	104
		168	160	150	139	82	60

Odgovore koje želimo da dobijemo nakon ovakvog eksperimenta su:
1. Koji uticaj na dužinu trajanja baterije imaju ova dva faktora, odnosno vrsta zaštitnog maperijala i spoljašnja temperatura kojoj je izložena baterija?
2. Da li postoji zaštitini materijal koji će omogućiti relativno dugo trajanje baterije bez obzira kojoj temperaturi je ona sam izložena?

Ovo drugo pitanje je od posebne važnosti. Možda je moguće naći materijal na koji ne utiče mnogo temperatura? Ako je to moguće, onda inžinjeri mogu da načine ovu bateriju otpornom na temperaturna kolebanja.

2.1. Statistička analiza za model faktorijalnog ogleda sa dva faktora i fiksiranim brojem nivoa

Pre nego što analiziramo gore pomenuti primer, opisaćemo postupak statističke analize za faktorijalni ogled sa dva faktora. Posmatrajmo opšti slučaj faktorijalnog ogleda sa dva faktora. Neka je X_{ijk} vrednost obeležja X na elementu populacije izloženom i-tom nivou faktora A ($i =$

$1,2,\dots,a$), j-om nivou faktora B ($j = 1,2,\dots,b$), za k-to ponavljanje eksperimenta ($k = 1,2,\dots,n$). Testiranje se vrši na osnovu prostog slučajnog uzorka i rezultati uzorkovanja se mogu prikazati pomoću sledeće Tabele 2.2.

Tabela 2.2. Uzorak

		Faktor *B*			
		1	2	...	b
Faktor *A*	1	X_{111} X_{112} ... X_{11n}	X_{121} X_{122} ... X_{12n}	... ...	X_{1b1} X_{1b2} ... X_{1bn}
	2	X_{211} X_{212} ... X_{21n}	X_{221} X_{222} ... X_{22n}	... ...	X_{2b1} X_{2b2} ... X_{2bn}
	...	...	...	...	...
	a	X_{a11} X_{a12} ... X_{a1n}	X_{a21} X_{a22} ... X_{a2n}	... ...	X_{ab1} X_{ab2} ... X_{abn}

Posmatranja se mogu opisati pomoću linearnog modela

$$X_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

gde je $i = 1,2,\dots,a$; $j = 1,2,\dots,b$; $k = 1,2,\dots,n$; μ matematičko očekivanje obeležja X, τ_i označava efekat $i-tog$ nivoa faktora *A*, β_j efekat *j*-og nivao faktora *B*, $(\tau\beta)_{ij}$ je efekat interakcije između nivoa τ_i i β_j, i na kraju ε_{ijk} je slučajna greška. Oba faktora *A* i *B* su fiksirana, a uticaj nivoa faktora se definiše kao odstupanje od očekivanja obeležja, pa je tako

$$\sum_{i=1}^{a} \tau_i = 0$$

i

$$\sum_{j=1}^{b} \beta_i = 0$$

Slično, efekti interakcija nivoa su fiksirani pa i ovde važi

$$\sum_{i=1}^{a} \sum_{j=1}^{b} (\tau\beta)_{ij} = 0$$

S obizirom na to da je u pitanju *n* ponavljanja eksperimenta, biće ukupno *abn* vrednosti obeležja. Dalje ćemo testirati hipoteze:

- efekti nivoa faktora *A* na obeležje *X* su bez bitnih razlika

$$H_0 : (\tau_1 = \tau_2 = \dots = \tau_a = 0),$$

protiv alternativne hipoteze:

$$H_1 : (\exists i : i \in \{1,2,...,a\}, \tau_i \neq 0)$$

- efekti nivoa faktora *B* na obeležje *X* su bez bitnih razlika

$$H_0 : (\beta_1 = \beta_2 = ... = \beta_b = 0),$$

protiv alternativne hipoteze:

$$H_1 : (\exists j : j \in \{1,2,...,b\}, \beta_j \neq 0)$$

- nema interaktivnog dejstva faktora *A* i *B* na obeležje *X*;

$$H_0: (\tau\beta)_{ij} = 0$$

protiv alternativne hipoteze:

$$H_1 : (\exists (i, j) : i \in \{1,2,...,a\}, j \in \{1,2,...,b\}, (\tau\beta)_{ij} \neq 0).$$

. Nadalje ćemo koristiti dvofaktorsku analizu varijanse za testiranje ovih hipoteza. Neka je $X_{i..}$ suma svih vrednosti obeležja u uzorku za *i*-ti nivo faktora *A*, $X_{.j.}$ suma svih vrednosti obeležja u uzorku za *j*-i nivo faktora *B*, $X_{ij.}$ suma svih vrednosti obeležja u uzorku za *ij*-u ćeliju i $X_{...}$ ukupan total, tj. suma svih vrednosti u uzorku. Definišimo opet odgovarajuće uzoračke srednje vrednosti kao $\overline{X}_{i..}$, $\overline{X}_{.j.}$, $\overline{X}_{ij.}$ i $\overline{X}_{...}$ pa se dobije:

$$X_{i..} = \sum_{j=1}^{b}\sum_{k=1}^{n} X_{ijk} \quad , \quad \overline{X}_{i..} = \frac{X_{i..}}{bn} \quad i = 1,2, ... , a$$

$$X_{.j.} = \sum_{i=1}^{a}\sum_{k=1}^{n} X_{ijk} \quad , \quad \overline{X}_{.j.} = \frac{X_{.j.}}{an} \quad j = 1,2, ... , b$$

$$X_{ij.} = \sum_{k=1}^{n} X_{ijk} \quad , \quad \overline{X}_{ij.} = \frac{X_{ij.}}{n} \quad i = 1,2, ... , a; j = 1,2, ... , b$$

$$X_{...} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} X_{ijk} \quad , \quad \overline{X}_{...} = \frac{X_{...}}{abn}$$

Ukupna suma kvadrata odstupanja od srednje vrednosti može se zapisati na sledeći način

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} (X_{ijk} - \overline{X}_{...})^2 =$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} [\left(\overline{X}_{i..} - \overline{X}_{...}\right) + \left(\overline{X}_{.j.} - \overline{X}_{...}\right) + \left(\overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} + \overline{X}_{...}\right) + (X_{ijk} - \overline{X}_{ij.})]^2$$

$$= bn\sum_{i=1}^{a}\left(\overline{X}_{i..} - \overline{X}_{...}\right)^2 + an\sum_{j=1}^{b}\left(\overline{X}_{.j.} - \overline{X}_{...}\right)^2$$

$$+ n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} + \overline{X}_{...}\right)^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(X_{ijk} - \overline{X}_{ij.})^2$$

pri čemu je šest međusobnih proizvoda jednaki nuli. Na ovaj način suma kvadrata odstupanja od srednje vrednosti se može razbiti na zbir četiri sume. Prva predstavlja sumu kvadrata odstupanja za faktor A (SS_A), druga sumu kvadrata odstupanja za faktor B (SS_B), treća sumu kvadrata odstupanja za interakciju faktora A i B (SS_{AB}), i na kraju četvrta sumu kvadrata greške (SS_E). Iz poslednjeg sabirka sa desne strane jednačine vidimo da moraju biti najmanje dva ponavljanja eksperimenta kako bi se dobila suma kvadrata greške. Prethodna jednačina može se zapisati na standardan način

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E$$

Broj stepena slobode za sve sume je

Tabela 2.3.

Efekat	Stepeni slobode
Faktor A	$a-1$
Faktor B	$b-1$
Interakcija faktora A i B	$(a-1)(b-1)$
Greška	$ab(n-1)$
Total	$abn-1$

Faktori A i B imaju redom a i b nivoa, te s toga imaju i $a-1$ i $b-1$ stepen slobode. Stepen slobode je jednostavno broj stepen slobode za ćelije, što iznosi $ab-1$ umanjen za stepene slobode oba faktora, odnosno

$$ab-1-(a-1)-(b-1)=(a-1)(b-1)$$

Za svaku ab ćeliju u uzorku je $n-1$ stepen slobode između n ponavljanja eksperimenta. Zato je $ab(n-1)$ stepen slobode za grešku. Stepen slobode za total odgovara zbiru prethodnih stepeni slobode, baš kao što se i računa suma kvadrata odstupanja. Deljenjem zbira kvadrata odgovarajućim stepenom slobode dobiju se sredine zbira kvadrata (varijanse) čija su očekivanja

$$E(MS_A) = E\left(\frac{SS_A}{a-1}\right) = \sigma^2 + \frac{bn}{a-1}\sum_{i=1}^{a}\tau_i^2$$

$$E(MS_B) = E\left(\frac{SS_B}{b-1}\right) = \sigma^2 + \frac{an}{b-1}\sum_{j=1}^{b}\beta_j^2$$

$$E(MS_{AB}) = E\left(\frac{SS_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n}{(a-1)(b-1)}\left(\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2\right)$$

$$E(MS_E) = E\left(\frac{SS_E}{ab(n-1)}\right) = \sigma^2$$

Ako su testirane nula hipoteze netačne, onda je σ^2 ocena za sve sredine suma kvadrata: MS_A, MS_B, MS_{AB} i MS_E . Međutim, ako postoje razlike između nivoa, recimo nivoa faktora *A,* onda će MS_A biti mnogo veća od MS_E. Isto važi i za drugi faktor *B,* u tom slučaju će MS_B biti mnogo veća od MS_E . Da bismo testirali značajnost efekata faktora, kao i njihove interakcije, potrebno je podeliti njihove zbirove kvadrata sa zbirom kvadrata greške. Ukoliko se dobiju velike vrednosti onda to upućuje na značajne razlike, odnosno na odbijanje početne hipoteze. Ako se pretpostavi da je statistički linearni model sa početka adekvatan, i ako nezavisne slučajne promenljive koje predstavljaju greške ε_{ijk} imaju normalnu raspodelu sa konstantnom disperzijom σ^2, tada svi razlomci: $\frac{MS_A}{MS_E}$, $\frac{MS_B}{MS_E}$ i $\frac{MS_{AB}}{MS_E}$, imaju Fišerovu raspodelu sa odgovarajućim stepenima slobode redom:

$$F_{a-1,ab(n-1)} = \frac{MS_A}{MS_E}$$
$$F_{b-1,ab(n-1)} = \frac{MS_B}{MS_E}$$
$$F_{(a-1)(b-1),ab(n-1)} = \frac{MS_{AB}}{MS_E}$$

Procedura anlize varijanse ponovo je predstavljena u sledećoj Tabeli 2.4.

Tabela 2.4. Analiza varijanse

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	F_0
Faktor A	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
Faktor B	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
Interakcija faktora A i B	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
Greška	SS_E	$ab(n-1)$	$MS_E = \frac{SS_E}{ab(n-1)}$	
Ukupno	SS_T	$abn-1$		

gde se sume kvadrata razlika računaju na standardan način:

- suma kvadrata odstupanja od srednje vrednosti

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} X_{ijk}{}^2 - \frac{X_{...}{}^2}{abn}$$

- suma kvadrata odstupanja za faktor *A*

$$SS_A = \sum_{i=1}^{a} \frac{X_{i..}^2}{bn} - \frac{X_{...}^2}{abn}$$

- suma kvadrata odstupanja za faktor *B*

$$SS_B = \sum_{j=1}^{b} \frac{X_{.j.}^2}{an} - \frac{X_{...}^2}{abn}$$

Poželjno je da sumu kvadrata interakcije SS_{AB} dobijemo na drugačiji način. Prvo ćemo izračunati sumu kvadrata podtotala (Subtotals) odstupanje *ab* čelija od srednje vrednosti uzorka:

$$SS_{Subtotals} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{X_{ij.}^2}{n} - \frac{X_{...}^2}{abn}$$

koja sadrži sume kvadrata oba faktora SS_A i SS_B , pa kad njih oduzmemo dobija se tražena suma kvadrata interakcije

$$SS_{AB} = SS_{Subtotals} - SS_A - SS_B$$

I na kraju lako se dobije suma kvadrata greške pomoću prethodno izračunatih vrednosti:

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = SS_T - SS_{Subtotals}$$

2.2. Primer faktorijalnog ogleda sa dva faktora i fiksiranim brojem nivoa

U Tabeli 2.5. su ponovo prikazani rezultati testiranja baterija, odnosno njihovog trajanja u odnosu na vrstu materijala kojim su obložene i spoljašnjoj temperaturi kojoj su izložene. Takođe, u marginama tabele su prikazani odgovarajući totali vrsta i kolona, te boldovani brojevi koji predstavljaju totale odgovarajućih ćelija.

Tabela 2.5.

		Spoljašnja temperatura									
		15°F			**70**°F			**125**°F			$\boldsymbol{X_{i..}}$
Vrsta zaštitnog materijala	**Tip 1**	130	155	**539**	34	40	**229**	20	70	**230**	998
		74	180		80	75		82	58		
	Tip 2	150	188	**623**	126	122	**479**	25	70	**198**	1330
		159	126		106	115		58	45		
	Tip 3	138	110	**576**	174	120	**583**	96	104	**342**	1501
		168	160		150	139		82	60		
	$\boldsymbol{X_{.j.}}$	1738			1291			770			$\boldsymbol{X_{...} = 3799}$

Suma kvadrata se računa na sledeći način:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} X_{ijk}^2 - \frac{X_{...}^2}{abn} = (130)^2 + (155)^2 + (74)^2 + \cdots + (60)^2 - \frac{(3799)^2}{36}$$
$$= 77646{,}97$$

$$SS_{vrsta\ materijala} = \sum_{i=1}^{a} \frac{X_{i..}^2}{bn} - \frac{X_{...}^2}{abn} = \frac{(998)^2 + (1300)^2 + (1501)^2}{3 \cdot 4} - \frac{(3799)^2}{36} = 10683{,}72$$

$$SS_{temperatura} = \sum_{j=1}^{b} \frac{X_{.j.}^2}{an} - \frac{X_{...}^2}{abn} = \frac{(1738)^2 + (1291)^2 + (770)^2}{3 \cdot 4} - \frac{(3799)^2}{36} = 39118{,}72$$

$$SS_{interakcija} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{X_{ij.}^2}{n} - \frac{X_{...}^2}{abn} - SS_{vrsta\ materijala} - SS_{temperatura}$$
$$= \frac{(539)^2 + (229)^2 + \cdots + (342)^2}{4} - \frac{(3799)^2}{36} - 10683{,}72 - 39118{,}72$$
$$= 9613{,}78$$

i na kraju

$$SS_E = SS_T - SS_{vrsta\ materijala} - SS_{temperatura} - SS_{interakcija}$$
$$= 77646,97 - 10683,72 - 39118,72 - 9613,78 = 18230,75$$

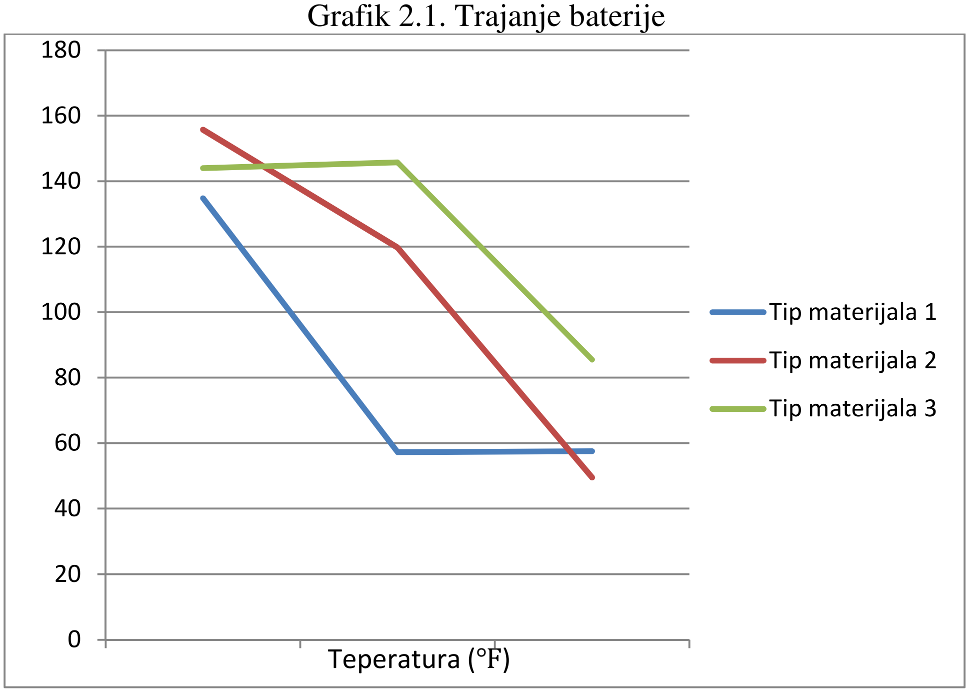
Rezultati analize varijanse su prikazani u Tabeli 2.6.

Tabela 2.6.

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	F_0
Tip materijala	10683,72	2	5341,86	7,91
Temperatura	39118,72	2	19558,36	28,97
Interakcija	9613,78	4	2403,44	3,56
Greška	18230,75	27	675,21	
Ukupno	77646,97	35		

Iz tablice za Fišerovu raspodelu dobije se da je $F_{0,05;4;27} = 2,73$ i iz toga sledi da postoji značajna interakcija između faktora, odnosno vrste materijala i temperature. Šta više, $F_{0,05;2;27} = 3,35$, pa su i uticaji samih faktora značajni. Da bi se bolje interpretirali rezultati analize, od pomoći je nacrtati grafik sa srednjim vrednostima za svaku kombinaciju nivoa.

Grafik 2.1. Trajanje baterije



Značajna interakcija imeđu faktora se vidi u nedostatku paralelnosti duži. Generalno gledajući, vidi se da postoji tendencija da baterija ima duži vek trajanja ukoliko je spoljašnja temperatura manja, bez obzira na vrstu materijala kojom je sama baterija zaštićena. Idući od najmanje tepmerature koja se ispitivala (15°F), ka višoj temperaturi (70°F) trajanje baterije se smanjuje za vrste materijala 1 i 2, a samo za treću vrstu materijala se blago povećava. Nastavljajući ka najvišoj temperaturi iz eksperimenta (125°F) trajanje baterije za vrste materijala 2 i 3 se drastično smanjuje, a za vrstu materijala 1 se uočava tek blago smanjenje. Iz svega viđenog možemo zaključiti da je najbolja opcija vrsta materijala 3 jer obezbeđuje najmanje gubitke efikasnosti trajanja baterije za različitu promenu temperature.

2.3. Višestruko poređenje

Kada analiza varijanse ukazuje da postoje razlike između srednjih vrednosti faktora, obično je poželjno napraviti poređenje između pojedinačnih nivoa da bi se otkrile značajne razlike. Da bismo ovo izveli, primenićemo Dankanov test za prethodni primer. Ne zaboravimo da je u ogledu interakcija bila značajna. Kada je interakcija značajna, poređenja srednjih vrednosti nivoa za jedan faktor (npr. faktor *A*) mogu biti sakrivena u interakciji *AB*.

Dankanov test višestrukog raspona. Ovaj test je razvio Dankan (Duncan) 1955. godine. Standardna greška sredine tretmana se računa kao

$$S_{\overline{X}} = \sqrt{\frac{MS_E^2}{n_h}}$$

gde je $$n_h = \frac{k}{\sum_{i=1}^{k} \frac{1}{n_i}}$$

Iz tabele sa Dankanovim značajnim rasponima izaberu se brojevi $r_\propto(j, n-k)$, za $j = 2,3, \ldots, k$, pri čemu je α nivo značajnosti, a $n-k$ broj stepeni slobode greške. Pomoću njih se određuje $k-1$ najmanje značajnih raspona

$$R_j = r_\propto(j, n-k) S_{\overline{X}} \quad \text{za } j = 2,3, \ldots, k \ .$$

Razlike među sredinama se testiraju tako što se prvo upoređuju najveća sredina sa najmanjom, odnosno njihova razlika sa R_k . Nakon toga se pravi razlika između najveće sredine nivoa i pretposlednje po veličini (prva veće od najmanje). Ova razlika se upoređuje sa R_{k-1} . Upoređivanje se nastavlja dok se sve sredine ne uporede sa najvećom. Onda se prave razlike između druge po veličini i najmanje, i upoređuje se sa R_{k-1} . Postupak se na dalje ponavlja sve dok se ne uporede sve moguće razlike sredina.

Vratimo se na primer. Kako je interakcija značajna, upoređivaćemo nivoe materijala sa samo jednim nivoa za temperaturu. Neka to bude nivo dva, odnosno 70°F . Pretpostavljamo da je najbolja ocena za varijansu greške MS_E , i da je greška ista za sve kombinacije nivao. Posmatramo tri srednje vrednosti materijala za odabranu temperaturu:

$$\overline{X}_{12.} = 57{,}25 \ \text{(tip materijala 1)}$$
$$\overline{X}_{22.} = 119{,}75 \ \text{(tip materijala 2)}$$
$$\overline{X}_{32.} = 145{,}75 \ \text{(tip materijala 3)}$$

Standardna greška sredine tretmana je

$$S_{\overline{X}_{i2.}} = \sqrt{\frac{MS_E}{n}} = \sqrt{\frac{675{,}21}{4}} = 12{,}99$$

gde je $n = 4$, jer je toliko ponavljanja bilo eksperimenta. Iz tabele sa Dankanovim značajnim rasponima dobiju se sledeće vrednosti: $r_{0.05}(2{,}27) \approx 2{,}91$ i $r_{0.05}(3{,}27) \approx 3{,}06$. Najmanji značajni rasponi su

$$R_2 = r_{0.05}(2{,}27)S_{\overline{X}_{i2.}} = 2{,}91 \cdot 12{,}99 = 37{,}80$$
$$R_3 = r_{0.05}(3{,}27)S_{\overline{X}_{i2.}} = 3{,}06 \cdot 12{,}99 = 39{,}75$$

Sada upoređujemo dobijene vrednosti:

$$\overline{X}_{32.} - \overline{X}_{12.} = 145{,}75 - 57{,}25 = 88{,}50 > 39{,}75 = R_3$$
$$\overline{X}_{32.} - \overline{X}_{22.} = 145{,}75 - 119{,}75 = 26{,}00 < 37{,}80 = R_2$$
$$\overline{X}_{22.} - \overline{X}_{12.} = 119{,}75 - 57{,}25 = 62{,}50 > 37{,}80 = R_2$$

Na osnovu prethodnog zaključujemo da na temperaturi od 70°F postoje značajne razlike u pogledu prosečnog trajanja baterije između trećeg i prvog tipa materijala, i između drugog i prvog, ali ne postoje značajne razlike između trećeg i drugog tipa materijala. Drugim rečima prvi tip materijala daje značajno lošije rezultate nego drugi i treći tip.

Ako je inerakcija značajna, moguće je sprovesti poređenje svih sredina *ab* ćelija kako bi se ustanovilo koja ima značajniju razliku. Za ovaj eksperiment to bi značilo da se upoređuju 36 parova, jer je broj ćelija 9.

2.4. Testiranje adekvatnosti modela

Prilikom formiranja linearnog modela, potrebno je izvršiti analizu reziduala, da bi se utvrdilo koliko je model odgovarajući, koliko dobro je prilagođen podacima. Obično se ispituje da li reziduali imaju normalnu raspodelu, upoređuju se reziduali sa fitovanim vrednostima, kao i sa svakom od nezavisnih promenljivih u modelu. Takođe, ako postoje promenljive koje nisu uključene u model, a mogu biti interesantne, i njih treba uporediti sa rezidualima. Upoređivanje se vrši tako što se crtaju dijagrami rasturanja, i ako se na njima može uočiti neka pravilnost, model je moguće još poboljšati. [7]

Pre nego što prihvatimo zaključke analize varijanse za primer dužine trajanja baterije, izvršićemo testiranje adekvatnosti modela, i u tu svrhu koristiti prethodno pomenutu analizu reziduala. Reziduali za faktorijalni model sa dva faktora je

$$e_{ijk} = X_{ijk} - \hat{X}_{ijk}$$

i kako je fitovana vrednost $\hat{X}_{ijk} = \overline{X}_{ijk}$ (prosečna vrednost uzorka u *ij*-toj ćeliji) prethodna jednačina se svodi na

$$e_{ijk} = X_{ijk} - \overline{X}_{ijk}$$

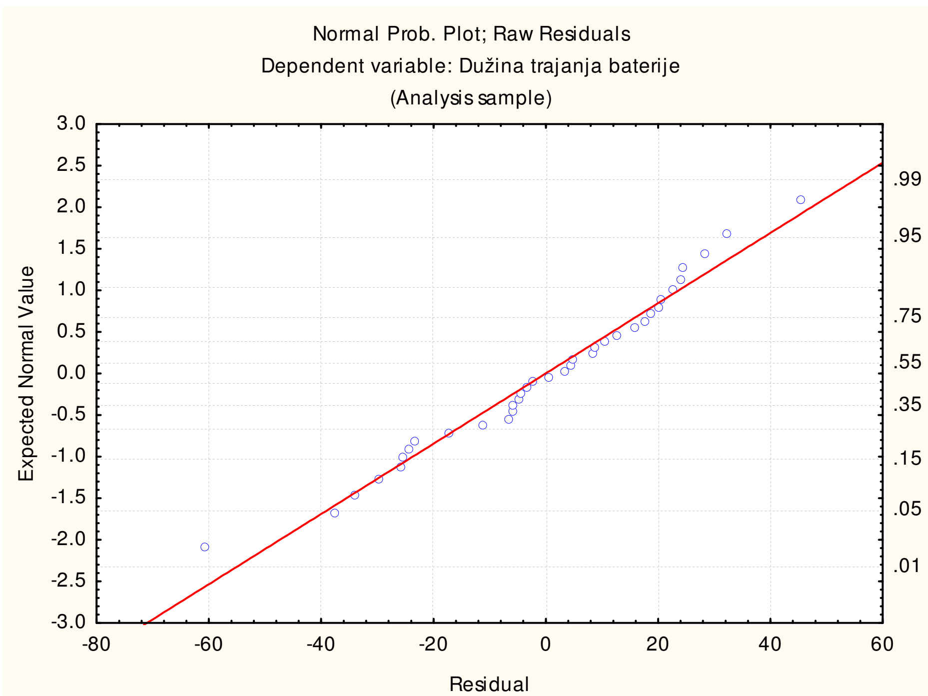
Reziduali za prethodni primer prikazani su u Tabeli 2.7.

Tabela 2.7. Reziduali

		Spoljašnja temperatura					
		15°F		70°F		125°F	
Vrsta zaštitnog materijala	Tip 1	-4.7500	20.2500	-23.2500	-17.2500	-37.5000	12.5000
		-60.7500	45.2500	22.7500	17.7500	24.5000	0.5000
	Tip 2	-5.7500	32.2500	16.2500	2.2500	-24.5000	20.5000
		3.2500	-29.7500	-13.7500	-4.7500	8.5000	-4.5000
	Tip 3	-6.0000	-34.0000	28.2500	-25.7500	10.5000	18.5000
		24.0000	16.0000	4.2500	-6.7500	-3.5000	-25.5000

kao i grafik normalne raspodele i dijagram reziduala

Grafik 2.2.

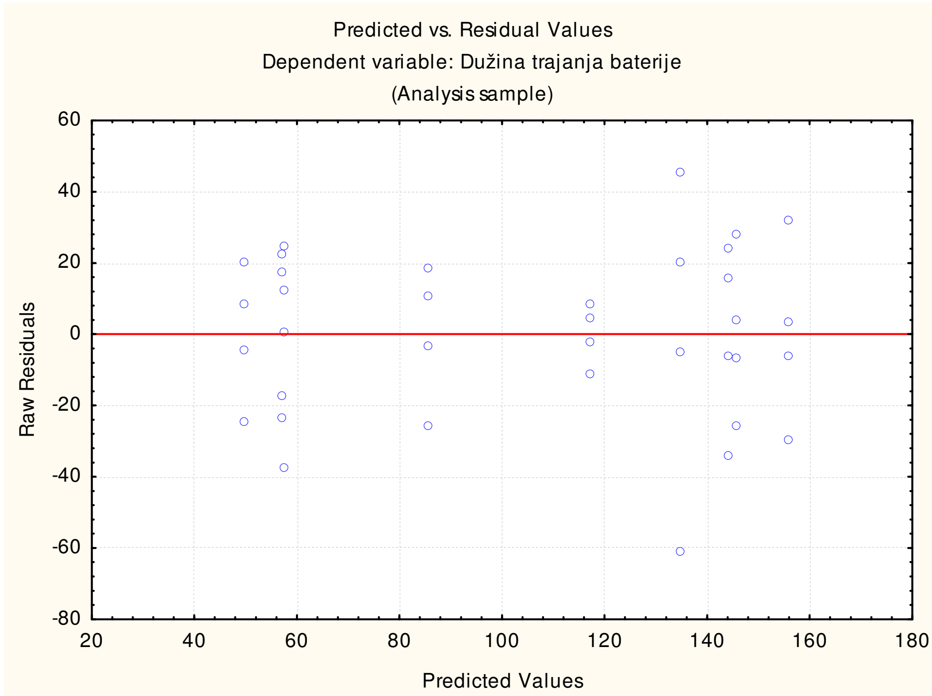


na kome se ne uočavaju nikakva odstupanja. Takođe, najveći negativni rezidual (-60,75 za 15°F i tip 1 materijala) izdvaja se u odnosu na druge. Standardizovana vrednost ovog reziduala je

$$\frac{-60,75}{\sqrt{MS_E}} = \frac{-60,75}{\sqrt{675,21}} = -2,34$$

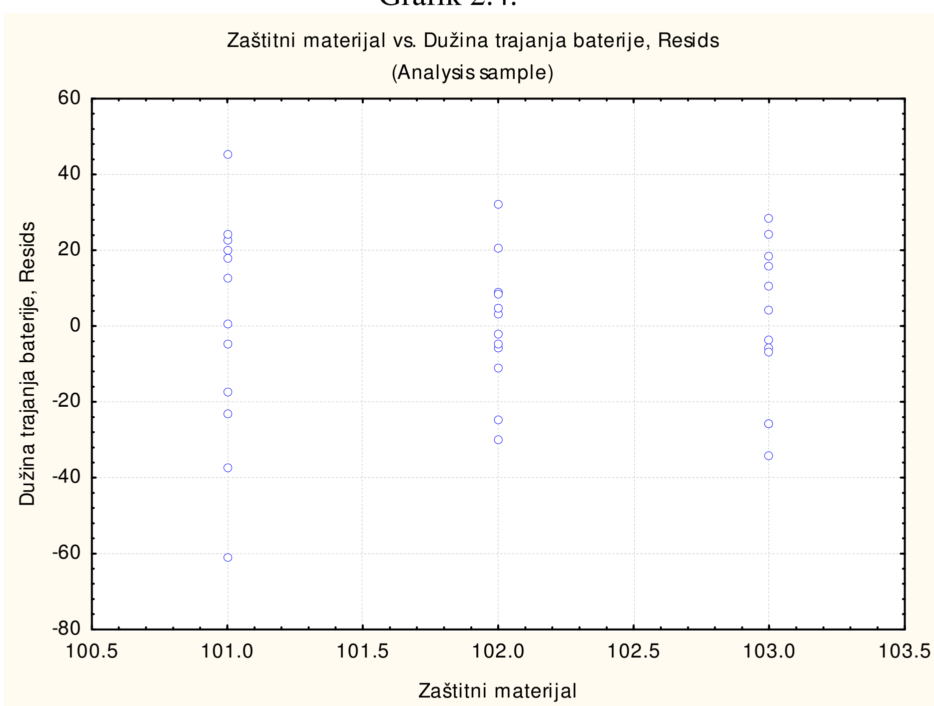
i ovo je jedini rezidual čija je apsolutna vrednost veća od dva. Sledeći Grafik 2.3.

Grafik 2.3.

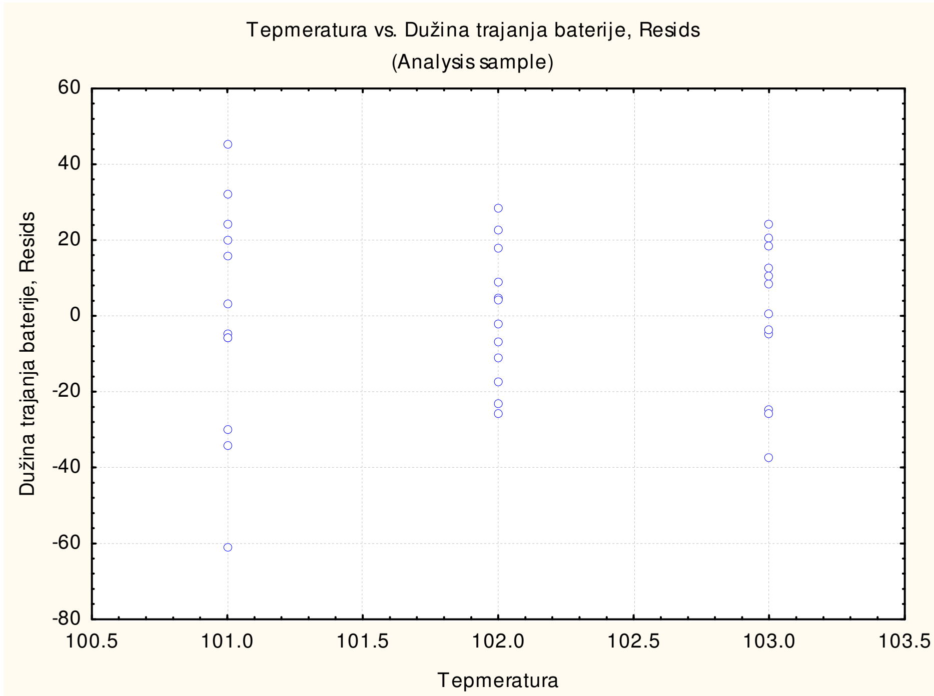


prikazuje reziduale u odnosu na fitovane vrednosti $\hat{X}_{ijk}$. Ovaj grafik upićuje na blago povećanje disperzije reziduala kako dužina trajanja baterije raste. Naredni Grafik 2.4. i Grafik 2.5. prikazuju reziduale u odnosu na tip materijala i temperaturu

Grafik 2.4.



Grafik 2.5.



Oba grafika ukazuju na blagu nejednakost disperzije. Tako kombinacija nivoa 15°F i tip 1 materijala ima veću disperziju u odnosu na druge kombinacije nivao. Takođe, treba primetiti da se u ovoj kombinaciji nivoa javlja ne samo najveći negativni rezisual (-60,75), već i najveći pozitivni (45,25), tj. oba ekstremna reziduala. Ova dva reziduala su odgovorna za nejednakost disperzije u prethodna tri grafika. Ponovno ispitivanje podataka nije otkrila nijedan očiti problem, kao što su greška u beleženju vrednosti, tako da se rezultati prihvataju kao tačni. Moguće je da ova kombinacija nivoa uzrokuje neznatno veći nepravilan vek trajanja baterije nego ostale kombinacije. Međutim, to nije dovoljno ozbiljan problem koji ima značajniji uticaj na analizu i zaključke eksperimenta.

2.5. Ocena parametara modela

Parametri u faktorskom ogledu sa dva faktora

$$X_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$$

mogu biti ocenjeni metodom najmanjih kvadrata. Kako model ima $1 + a + b + ab$ parametara koje treba oceniti, koristiće se sistem od isto toliko normalnih jednačina.

$$abn\hat{\mu} + bn\sum_{i=1}^{a}\hat{\tau}_i + an\sum_{j=1}^{b}\hat{\beta}_j + n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij} = X_{...}$$

$$bn\hat{\mu} + bn\hat{\tau}_i + n\sum_{\boldsymbol{j=1}}^{\boldsymbol{b}}\hat{\beta}_j + n\sum_{\boldsymbol{j=1}}^{\boldsymbol{b}}(\tau\beta)_{ij} = X_{i..}$$

$$an\hat{\mu} + n\sum_{i=1}^{a}\hat{\tau}_i + an\hat{\beta}_j + n\sum_{j=1}^{b}(\tau\beta)_{ij} = X_{.j.}$$
$$n\hat{\mu} + n\hat{\tau}_i + n\hat{\beta}_j + n(\tau\beta)_{ij} = X_{ij.}$$
Prvo primetimo da zbir a drugih jednačina jeste prva jednačina i da je zbir b trećih jednačina ponovo prva jednačina. Takođe, sabiranjem četvrte jednačine po *j*-u za svako i dobiće se druga jednaćina, a sabiranjem četvrte jendačine po i za svako j dobiće se opet treća jednačina. S toga, postoje $a + b + 1$ nezavisnih promenljivih u sistemu i dobiće se nejedinstveno rešenje. Kako bi se sistem rešio, uvode se ograničenja, odnosno uslovi
$$\sum_{i=1}^{a}\hat{\tau}_i = 0$$
$$\sum_{j=1}^{b}\hat{\beta}_i = 0$$
$$\sum_{i=1}^{a}(\widehat{\tau\beta})_{ij} = 0$$
$$\sum_{j=1}^{b}(\widehat{\tau\beta})_{ij} = 0$$
Prve dve jednačine uvode dva uslova, dok presostale dve jednačine uvode $a + b - 1$ nezavisnoh uslova. Tako se dobije $a + b + 1$ uslova, što je broj koji je potreban. Uvodeći ove uslove, sistem normalnih jednačina se značajno pojednostavljuje
$$\hat{\mu} = \overline{X}_{...}$$
$$\hat{\boldsymbol{\tau}}_{\boldsymbol{i}} = \overline{X}_{i..} - \overline{X}_{...}$$
$$\hat{\boldsymbol{\beta}}_{\boldsymbol{j}} = \overline{X}_{.j.} - \overline{X}_{...}$$
$$\left(\widehat{\tau\beta}\right)_{ij} = \overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} - \overline{X}_{...}$$
Uticaji nivoa vrsta se ocenjuju tako što se od srednje vrednosti vrste oduzme ukupna srednja vrednost celog uzorka, Uticaji nivoa kolona se ocenjuju tako što se od srednje vrednosti kolone oduzme ukupna srednja vrednost celog uzorka, i uticaji *ij* interakcije se ocenjuje tako što se od srednje vrednosti *ij* ćelije oduzme srednja vrednost *i*-te vrste, srednja vrednost *j*-te kolone i ukupna srednja vrednost celog uzorka. Na osnovu prethodnih jednačina dobija se fitovana vrednost za X_{ijk}
$$\hat{X}_{ijk} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j + (\tau\beta)_{ij} = \overline{X}_{...} + \left(\overline{X}_{i..} - \overline{X}_{...}\right) + \left(\overline{X}_{.j.} - \overline{X}_{...}\right) - (\overline{X}_{ij.} - \overline{X}_{i..} - \overline{X}_{.j.} - \overline{X}_{...}) = \overline{X}_{ij.}$$

Odnosno, *k*-to posmatranje u *ij*-oj ćeliji je ocenjeno sa srednjom vrednošću te ćelije u *n* ponavljanja. Ovaj rezultat je korišćen kod ocene reziduala u prethodnom poglavlju. Kako su korišteni uslovi za rešavanje sisitema normalnih jednačina, model nije jednistveno ocenjen. Međutim, određene važne funkcije parametara modela su ocenjive, tj. jedinstveno ocenjene bez obzira na izabrane uslove. Na primer, $\tau_i - \tau_u + \overline{(\tau\beta)}_{i.} - \overline{(\tau\beta)}_{u.}$, može da se posmatra kao prava razlika između *i*-tog i *u*-tog nivoa faktora *A*. Podsetimo se da prava razlika među nivoima bilo kojeg faktora uključuje srednju vrednost interakcije. U principu, svaka funkcija koja predstavlja model parametara, a koja je linearna kombinacija levih strana normalnih jednačina, je ocenjiva.

Glava 3

3. 2^k faktorijalni ogledi

Faktorijalni ogledi se često upotrebljavaju u eksperimentima koji podrazumevaju ispitivanje uticaja više faktora na vrednost nekog obeležja. Postoji nekoliko specijalnih slučajeva faktorijalnih ogleda koji su bitni, ne samo zbog svoje široke upotrebe u istraživačkim radovima već i zbog toga što čine osnovu za druge modele ogleda.
Najbitniji među njima je svakako ogled koji ima k faktora od kojih svaki ima tačno dva nivoa ili tretmana. Ovi nivoi mogu biti kvantitativne vrednosti, kao na primer dve vrednosti temperature, ili pritiska, ili vremena. Takođe, mogu biti kvalitativne vrednosti, na primer dve mašine, ili dva operatora, ili visok i nizak nivo faktora, ili recimo prisustvo i odsustvo nekog faktora.
Sva moguća ponavljanja ovakvog ogleda zahtevaju $2 \times 2 \times 2 \times \cdots \times 2 = 2^k$ posmatranja i zato se i naziva 2^k faktorijalni ogled.

U ovom delu rada biće reči upravo o ovakvoj vrsti ogleda gde će kroz primere biti prikazani prvo najjednostavniji slučajevi 2^2 i 2^3 ogledi, a zatim će biti reči i o opštenom slučaju 2^k. Pretpostavićemo da su faktori fiksirani, uzorci su slučajni, i uslovi normalnosti su zadovoljeni.

Faktorijalni ogledi 2^k su pogodni za rane faze eksperimenata gde često bude mnogo faktora čije uticaje treba ispitati. Ova analiza omogućava najmanji broj koraka potrebnih da se potpuno sprovede faktorijalni ogled.

3.1 2^2 faktorijalni ogledi

Najjednostaniji slučaj 2^k ogleda je svakako onaj koji sadrži samo dva faktora, recimo A i B, svaki sa po dva nivoa. Ovakav ogled se naziva 2^2 faktorijalni ogled. Razmotrićemo ovakvu situaciju kroz jedan primer. Posmatraćemo uticaj koncetracije reaktanta i količine katalizatora na dužinu reakcije hemijskog procesa. Neka je koncetracija reaktanta faktor A i neka su njegova dva nivoa: niži nivo, odnosno 15% i viši nivo, tj. 25%. Količina katalizatora će tako biti faktor B i njegov niži nivo će odgovarati upotrebi jedne vrećice, a viši nivo upotrebi dveju vrećica katalizatora. [1] Ogled je ponovljen 3 puta i dobijeni su sledeći rezultati:

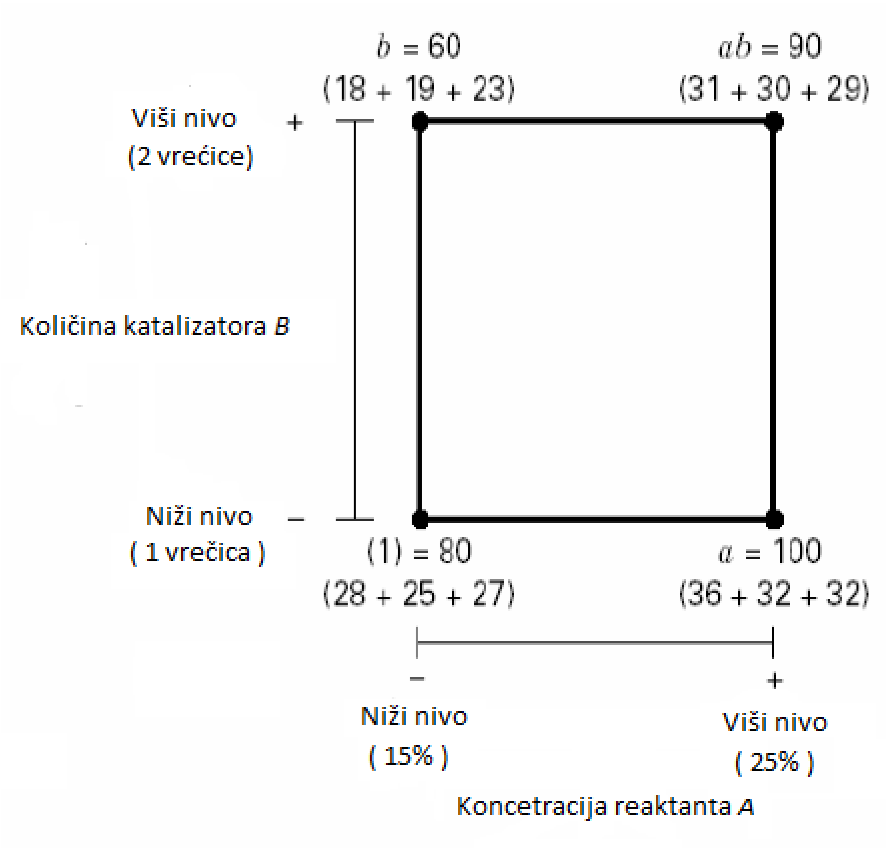
Tabela 3.1.

Kombinacija nivoa	**Ponavljanja**			
	1.	**2.**	**3.**	**Ukupno**
A nizak, B nizak	28	25	27	80
A visok, B nizak	36	32	32	100
A nizak, B visok	18	19	23	60
A visok, B visok	31	30	29	90

Obeležićemo uticaj faktora sa odgovarajućim velikim latiničnim slovima. Tako će „A“, označavati uticaj faktora A, „B“, će označavati uticaj faktora B, a „AB“, će označavati uticaj interakcije faktora A i B. U 2^2 ogledu niži i viši nivoi faktora A i B se obeležavaju sa „$-$“ i „$+$“ respektivno na A i B osama. Na taj način „$-$“ na A-osi predstavlja niži nivo koncetracije (15%) i „$+$“ viši nivo koncetracije (25%), te slično na B-osi „−“ predstavlja niži nivo katalizatora i „+“ viši nivo katalizatora.

Sve moguće kombinacije nivoa u ogledu se obično označavaju sa odgovarajućim malim latiničnim slovima. Kao što se vidi na Grafiku 3.1., viši nivo nekog faktora označen je njegovim malim slovom, dok je niži nivo faktora je označen odsustvom njegovog slova. Tako „a“ odgovara višem nivou faktora A i nižem nivou faktora B, „b“ odgovara nižem nivou faktora A i višem nivou faktora B, a oznaka „ab“ odgovara višem nivou faktora A i višem nivou faktora B. Preostala kombinacija nivoa, tj. kad su oba faktora na nižem nivou obeležava se sa (1). Ista notacija se koristi i kod drugih 2^k faktorijalnih ogleda.

Grafik 3.1.



Definišimo ***srednji efekat*** faktora kao prosečnu promenu vrednosti posmatranog obeležja uzrokovanu promenom u nivou jednog faktora u odnosu na nivoe drugog faktora. Takođe, notacije (1), $a, b,$ i ab sada predstavljaju totale, odnosno zbir vrednosti kombinacija tretmana u svim n ponavljanjima. Na osnovu toga, efekat faktora A kada je faktor B na nižem nivou je $\frac{a-(1)}{n}$, a efekat faktora A kada je faktor B na višem nivou je $\frac{ab-b}{n}$. Njihova srednja vrednost daće prosečan efekat faktora A:

$$A = \frac{1}{2n}[(ab-b)+(a-(1))] = \frac{1}{2n}[ab+a-b-(1)]$$

Analgno ćemo izraziti prosečan efekat za drugi faktor, i to kao srednju vrednost efekta faktora B kada je faktor A na nižem nivou ($\frac{b-(1)}{n}$) i efekta faktora B kada je faktor A na višem nivou ($\frac{ab-a}{n}$), tj. dobije se:

$$B=\frac{1}{2n}[(ab-a)+(b-(1))]=\frac{1}{2n}[ab+b-a-(1)]$$

Dalje, definišimo i efekat interakcije AB kao srednju vrednost razlike između efekta faktora A za viši nivo faktora B i efekta faktora A za niži nivo faktora B:

$$AB=\frac{1}{2n}[(ab-b)-(a-(1))]=\frac{1}{2n}[ab+(1)-a-b]$$

Do efekta interakcije AB može se doći i na drugi način, tako što se izrazi srednja vrednost razlike između efekta faktora B za viši nivo faktora A i efekta faktora B za niži nivo faktora A, i rezultat će biti isti kao u prethodnoj jednačini.

Formule koje izražavaju efekte A, B i interakciju AB moguće je izvesti i drugim metodom. Efekat faktora A se može naći kao razlika prosečne vrednosti dve kombinacije tretmana na desnoj stranici kvadrata prikazanog na prethodnom Grafiku 28 (obeležava se sa $\bar{X}_{A^+}$ jer predstavlja prosečnu vrednost kombinacija tretmana u kojima je nivo faktora A visok) i dve kombinacije tretmana na levoj stranici kvadrata (ili $\bar{X}_{A^-}$), odnosno dobija se:

$$A=\bar{X}_{A^+}-\bar{X}_{A^-}=\frac{ab+a}{2n}-\frac{b+(1)}{2n}=\frac{1}{2n}[ab+a-b-(1)]$$

Ovo je potpuno isti rezultat kao što smo već dobili za efekat faktora A. Slično se može naći i efekat faktora B kao razlika prosečne vrednosti dve kombinacije tretmana na gornjoj stranici kvadrata (ili $\bar{X}_{B^+}$) i dve kombinacije tretmana na donjoj stranici kvadrata (ili $\bar{X}_{B^-}$), pa se dobija:

$$B=\bar{X}_{B^+}-\bar{X}_{B^-}=\frac{ab+b}{2n}-\frac{a+(1)}{2n}=\frac{1}{2n}[ab+b-a-(1)]$$

Na kraju efekat interakcije AB je razlika prosečnih vrednosti kombinacija tretmana na dijagonalama kvadrata:

$$AB=\frac{ab+(1)}{2n}-\frac{a+b}{2n}=\frac{1}{2n}[ab+(1)-a-b]$$

Za konkretne vrednosti u prethodnom primeru to izgleda ovako:

$$A=\frac{1}{2\cdot 3}(90+100-60-80)=8{,}33$$

$$B=\frac{1}{2\cdot 3}(90+60-100-80)=-5{,}00$$

$$AB=\frac{1}{2\cdot 3}(90+80-100-60)=1{,}67$$

Efekat faktora A (koncetracija rektanta) je pozitivan što ukazuje da povećanje koncetracije reaktanta, odnosno prelazak sa nižeg nivoa faktora A od 15% na viši nivo od 25% utiče na povećanje vrednosti obeležja, odnosno na povećanje trajanja hemijske reakcije.

Efekat faktora B (količina katalizatora) je negativan, što opet upućuje da povećanje količine katalizatora, odnosno prelazak sa nižeg nivoa faktora B od jedne vrećice katalizatora na viši nivo faktora B od dve vrećice utiče na smanjenje vrednosti obeležja, odnosno na skraćenje vremena trajanja hemijske reakcije. Vrednost efekta interakcije AB je 1,67 i upućuje na relativno mali uticaj u odnosu na efekte faktora A i B.

U mnogim eksperimentima u kojima se koriste 2^k modeli, ispituju se jačina i pravac efekata kako bi se utvrdilo koji od njih ima važan uticaj. Kod 2^k modela primenićemo specijalnu metodu analize varijanse. Prvo ćemo uvesti pojam ***kontrasta*** koji se koristi pri izračunavanju efekata faktora A, B i interakcije AB i predstavlja njihove totalne efekte. Drugim rečima to znači:

$$Kontrast_A = ab + a - b - (1)$$
$$Kontrast_B = ab + b - a - (1)$$
$$Kontrast_{AB} = ab + (1) - a - b$$

i ova tri kontrasta su međusobno ortogonalni. Zbir kvadrata za sve varijable se može izračunati kada se kvadrati kontrasta podele sa brojem opservacija u svakom totalu. Na kraju dobije se:

suma kvadrata odstupanja za faktor A $\quad SS_A = \frac{[ab+a-b-(1)]^2}{4 \cdot n}$

suma kvadrata odstupanja za faktor B $\quad SS_B = \frac{[ab+b-a-(1)]^2}{4 \cdot n}$

suma kvadrata odstupanja za interakciju AB $\quad SS_{AB} = \frac{[ab+(1)-a-b]^2}{4 \cdot n}$

Sumu kvadrata totala ćemo zapisati na standardan način:

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} X_{ijk}^2 - \frac{X_{\dots}^2}{4n}$$

Generalno, SS_T ima $4n - 1$ stepen slobode dok suma kvadrata greške SS_E ima $4(n-1)$ stepen slobode i uobičajeno se računa preko ostalih suma kvadrata:

$$SS_E = SS_T - SS_A - SS_B - SS_{AB}$$

Koristeći podatke iz primera za prethodne jednačine dobijaju se sledeće vrednosti:

$$SS_A = \frac{(50)^2}{4 \cdot 3} = 208{,}33$$
$$SS_B = \frac{(-30)^2}{4 \cdot 3} = 75{,}00$$
$$SS_{AB} = \frac{(10)^2}{4 \cdot 3} = 8{,}33$$
$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} X_{ijk}^2 - \frac{X_{\dots}^2}{4 \cdot 3} = 9398{,}00 - 9075{,}00 = 323{,}00$$

i na kraju $\quad SS_E = SS_T - SS_A - SS_B - SS_{AB} = 323{,}00 - 208{,}33 - 75{,}00 - 8{,}33 = 31{,}34$

Kompletni rezultati analize varijanse predstavljeni su u sledećoj tabeli:

Tabela 38

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadarata	F_0
A	208.33	1	208.33	53.15
B	75.00	1	75.00	19.13
AB	8.33	1	8.33	2.13
Greška	31.34	8	3.92	
Ukupno	323.00	19		

Često je pogodno kombinacije tretmana zapisati redom $(1), a, b, ab$, tj. standardnim redosledom. Koristeći ovaj standardni redosled zapisaćemo koeficijente kontrasta korišćenih za izračunavanje efekata:

Tabela 3.2.

$Efekat$	(1)	a	b	ab
A	-1	+1	-1	+1
B	-1	-1	+1	+1
AB	+1	-1	-1	+1

Primetimo da se koeficijent interakcije AB dobije kao proizvod odgovarajućih koeficijenata faktora. Koeficijenti kontrasta su uvek ili +1 ili -1, i ova Tabela 39 se može koristiti za izračunavanje znaka određene kombinacije tretmana.

Sledećom Tabelom 40 predstavićemo i drugi način da se ovo uradi. Pored kombinacija tretmana u vrstama tabele i efekata A, B i interakcije AB u kolonama, pojavljuje se još jedna promenljiva u koloni, a to je I i predstavlja total ili prosečnu vrednost za celi eksperiment. Kolona koja odgovara I je uvek popunjena sa znakom plus.

Tabela 3.3.

Kombinacije tretmana	Efekti faktora			
	I	A	B	AB
(1)	+	-	-	+
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	+

Da bismo našli kontraste bilo kog faktora kako bismo izračunali njegov efekat, potrebno je jednostavno pomnožiti znakove iz njegove kolone sa odgovarajućim kombinacijama tretmana iz prve kolone tabele i sumirati. Recimo, kontrast faktora B se lako dobije da je $-(1) - a + b + ab$.

3.1.1. Analiza reziduala

Pomoću modela regresije lako je moguće izračunati reziduale, ili ostatke 2^k ogleda. Za hemijski proces iz primera, model regresije je:

$$X = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$$

gde je x_1 neslučajna promenljiva koja označava koncetraciju reaktanta, x_2 je neslučajna promenljiva koja označava količinu katalizatora i parametri $\beta_0, \beta_1, \beta_2$ su koeficijenti regresije. Promenljive x_1 i x_2 ćemo izraziti sledećim jednačinama:

$$x_1 = \frac{Konc - \dfrac{Konc_{nizak} + Konc_{visok}}{2}}{\dfrac{Konc_{visok} - Konc_{nizak}}{2}}$$

$$x_2 = \frac{Katal - \frac{Katal_{nizak} + Katal_{visok}}{2}}{\frac{Katal_{visok} - Katal_{nizak}}{2}}$$

Kada faktori imaju samo dva nivoa, ove promenljive će imati vrednosti ± 1. Za naš primer biće:

$$x_1 = \frac{Konc - \frac{15 + 25}{2}}{\frac{25 - 15}{2}} = \frac{Konc - 20}{5}$$

Ako je koncetracija rektanta na višem nivou ($Konc = 25\%$), onda je $x_1 = +1$, a ako je koncetracija reaktanta na niskom nivou ($Konc = 15\%$), onda je $x_1 = -1$. Slično za drugu promenljivu:

$$x_2 = \frac{Katal - \frac{1 + 2}{2}}{\frac{2 - 1}{2}} = \frac{Katal - 1{,}5}{0{,}5}$$

Ako je količina katalizatora na višem nivou ($Katal = 1\ vrećica$), onda je $x_2 = +1$, a ako je količina katalizatora na nižem nivou ($Katal = 2\ vrećica$), onda je $x_2 = -1$.

Fitovan model regresije je

$$\hat{X} = 27{,}5 + \frac{8{,}33}{2} x_1 + \frac{-5{,}00}{2} x_2$$

gde je parametar β_0 sredina uzorka, odnosno svih 12 posmatranja. Koeficijenti β_1 i β_2 su jedna polovina odgovarajućih ocena faktora. Razlog zbog čega su jedna polovina, jeste taj što koeficijenti regresije mere efekat promene jedinice, a to je u ovom slučaju od +1 do -1.

Ovaj model se može iskoristiti da se dobije očekivana vrednost obeležja X za četiri moguće kombinacije tretmana. Na primer, kada je koncetracija rekatanta na nižem nivou i kada je količina katalizatora na nižem nivou, onda je očekivana vrednost

$$\hat{X} = 27{,}5 + \frac{8{,}33}{2}(-1) + \frac{-5{,}00}{2}(-1) = 25{,}835$$

Postoje tri ponavljanja ove kombinacije nivoa, i reziduali su

$$e_1 = 28 - 25{,}835 = 2{,}165$$
$$e_2 = 25 - 25{,}835 = -0{,}835$$
$$e_3 = 27 - 25{,}835 = 1{,}165$$

Analogno se računaju i ostale očekivane vrednosti za ostale tri kombinace nivoa

- za koncetraciju raktanta na višem nivou i količine katalizatora na nižem je

$$\hat{X} = 27{,}5 + \frac{8{,}33}{2}(+1) + \frac{-5{,}00}{2}(-1) = 34{,}165$$
$$e_4 = 36 - 34{,}165 = 1{,}835$$
$$e_5 = 32 - 34{,}165 = -2{,}165$$
$$e_6 = 32 - 34{,}165 = -2{,}165$$

- za koncetraciju raktanta na nižem nivou i količine katalizatora na višem je

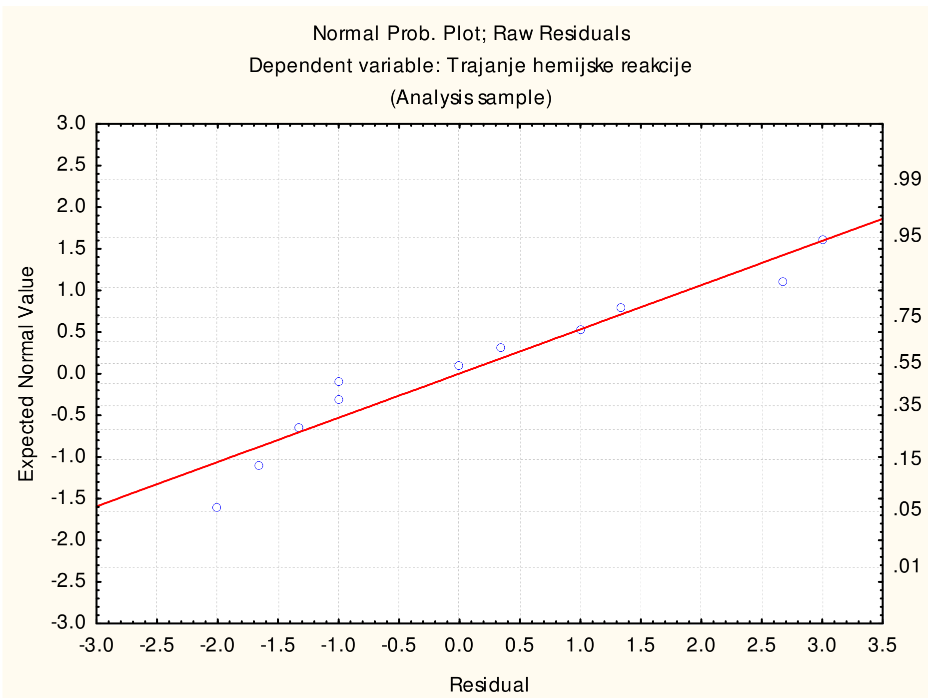
$$\hat{X} = 27{,}5 + \frac{8{,}33}{2}(-1) + \frac{-5{,}00}{2}(+1) = 20{,}835$$
$$e_7 = 18 - 20{,}835 = -2{,}835$$
$$e_8 = 19 - 20{,}835 = -1{,}835$$
$$e_9 = 23 - 20{,}835 = 2{,}165$$

- za koncetraciju raktanta na višem nivou i količine katalizatora na višem je

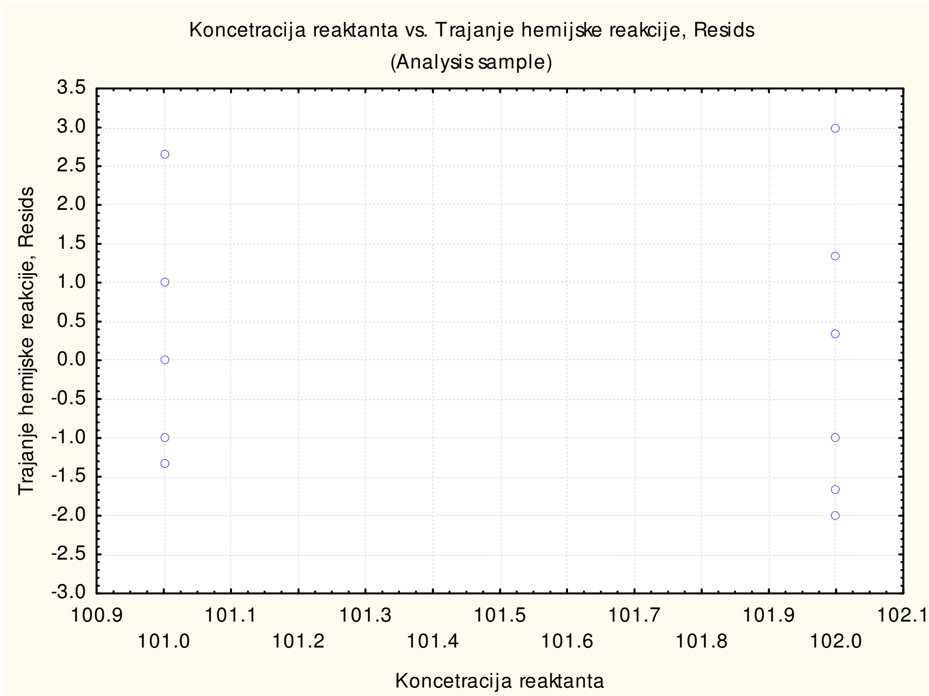
$$\hat{X} = 27{,}5 + \frac{8{,}33}{2}(+1) + \frac{-5{,}00}{2}(+1) = 29{,}165$$
$$e_{10} = 31 - 29{,}165 = 1{,}835$$
$$e_{11} = 30 - 29{,}165 = 0{,}835$$
$$e_{12}29 - 29{,}165 = 0{,}165$$

Grafik 3.2. prikazuju normalnu raspodelu za ove reziduale, te Grafik 3.3. prikazuje reziduale u odnosu na faktor *A*, tj koncetraciju reaktanta, i Grafik 3.4. prikazuje reziduale u odnosu na faktor *B*, tj. količine katalizatora. Ovi grafici su zadovoljavajući, pa nemamo razloga da sumnjamo u rezultate analize.

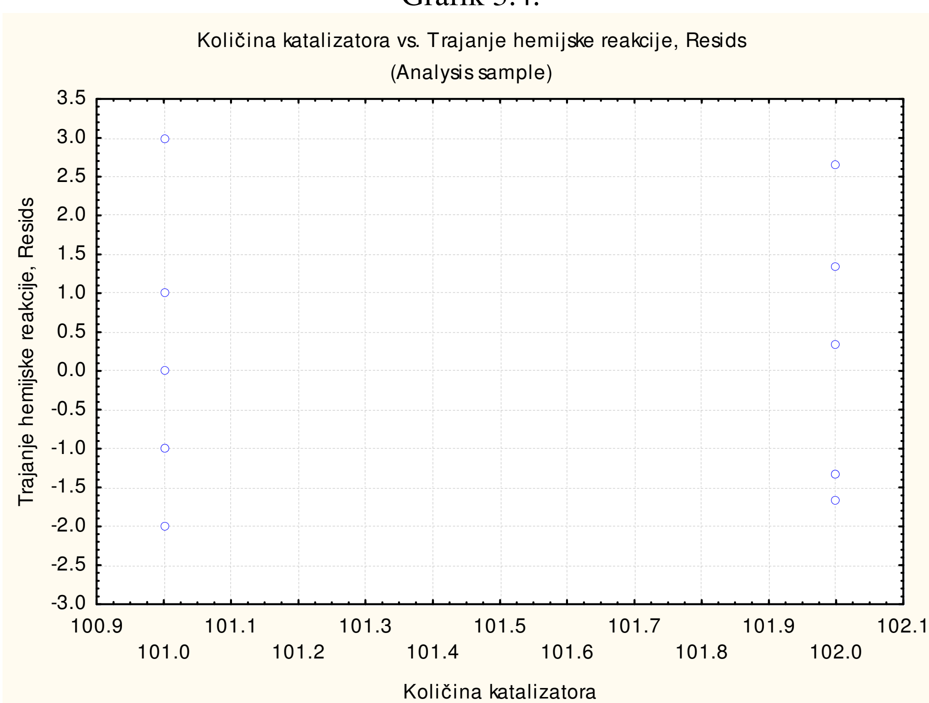
Grafik 3.2.



Grafik 3.3.



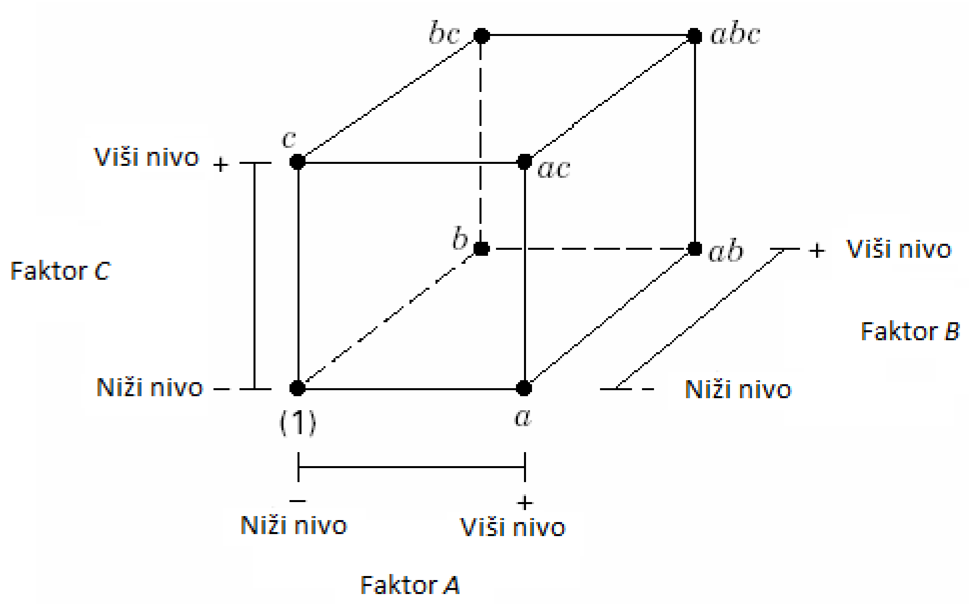
Grafik 3.4.



3.2. 2^3 faktorijalni ogledi

Posmatrajmo sada eksperiment sa tri faktora *A*, *B* i *C*, svaki sa po dva nivoa ili tretmana. Ogled se s toga naziva 2^3 faktorijalni ogled i ima ukupno osam mogućih kombinacija tretmana. Te kombinacije se mogu grafički prikazati sličnom notacijom kao i kod 2^2 ogleda samo što će ovde umesto kvadrata zbog trećeg faktora biti kocka. (Grafik 3.5.) Temena kocke su obeležena kombinacijama tretmana sa $(1), a, b, ab, c, ac, bc$ i abc, i osim toga predstavljaju i ukupne vrednosti svih *n* opservacija odgovarajućih kombinacija tretmana.

Grafik 3.5.



Postoje još dve različite nostacije koje se često upotrebljavaju u 2^k ogledima. Prva je „+ i −“ notacija, ili geometrijska notacija. Druga notacija je „1 i 0“ i označava viši i niži nivo faktora kao i što je to slučaj kod „+ i −“ notacije. Sve ove notacije prikazane su u Tabeli 3.3.

Tabela 3.3.

Mogućnost	*A*	*B*	*C*	Kombinacija tretmana	*A*	*B*	*C*
1	-	-	-	(1)	0	0	0
2	+	-	-	*a*	1	0	0
3	-	+	-	*b*	0	1	0
4	+	+	-	*ab*	1	1	0
5	-	-	+	*c*	0	0	1
6	+	-	+	*ac*	1	0	1
7	-	+	+	*bc*	0	1	1
8	+	+	+	*abc*	1	1	1

Postoji sedam stepeni slobode među osam mogućih kombinacija tretmana u 2^3 faktorijalinim ogledima, tj. po jedan stepen slobode za svaki fkator *A, B, C,* zatim po jedan stepen slobode za svaku interakciju *AB, AC, BC,* i na kraju jedan stepen slobode za interakciju *ABC.*
Sledeći korak jeste da se formulišu glavni efekti. Prvo ćemo izraziti glavni efekat faktora *A* kao srednju vrednost njegovih efakata za različite kombinacije nivoa faktora *B* i *C*: efekat faktora *A* kada su i *B* i *C* na niskom nivou ($\frac{a-(1)}{n}$), efakat faktora *A* kada je *B* na visokom nivou i *C* na niskom nivou ($\frac{ab-b}{n}$), efakat faktora *A* kada je *B* na niskom nivou i *C* na visokom nivou ($\frac{ac-c}{n}$) i

efekat faktora *A* kada su i *B* i *C* na visokm nivou ($\frac{abc-bc}{n}$). Dobijena srednja vrednost ovih četiri je:

$$A = \frac{1}{4n}(a - (1) + ab - b + ac - c + abc - bc)$$

Za faktore *B* i *C* postupak je isti i dobijeni glavni efekti su:

$$B = \frac{1}{4n}(b - (1) + ab - a + bc - c + abc - ac)$$

$$C = \frac{1}{4n}(c - (1) + ac - a + bc - b + abc - ab)$$

Drugi način da ovo dobijemo jeste da posmatramo razliku između srednjih vrednosti kad je odgovarajući faktor na višem nivou i srednjih vrednosti kad je odgovarajući faktor na nižem nivou. Na taj način se dobije:

$$A = \bar{X}_{A^+} - \bar{X}_{A^-} = \frac{a + ab + ac + abc}{4n} - \frac{(1) + b + c + bc}{4n}$$
$$= \frac{1}{4n}(a + ab + ac + abc - (1) - b - c - bc)$$
$$B = \bar{X}_{B^+} - \bar{X}_{B^-} = \frac{b + ab + bc + abc}{4n} - \frac{(1) + a + c + ac}{4n}$$
$$= \frac{1}{4n}(b + ab + bc + abc - (1) - a - c - ac)$$
$$C = \bar{X}_{C^+} - \bar{X}_{C^-} = \frac{c + ac + bc + abc}{4n} - \frac{(1) + a + b + ab}{4n}$$
$$= \frac{1}{4n}(c + ac + bc + abc - (1) - a - b - ab)$$

Dalje će nam trebati efekti dvofaktorske interakcije. Mera *AB* interakcije je razlika između prosečnih vrednosti efekta faktora *A* za dva nivoa faktora *B*.

Tabela 3.4.

B	***Srednja vrednost efekta faktora A***
Viši nivo (+)	$\frac{(abc - bc) + (ab - b)}{2n}$
Niži nivo (-)	$\frac{(ac - c) + (a - (1))}{2n}$
Razlika	$\frac{abc - bc + ac - b - ac + c - a + (1)}{2n}$

S obzirom da je *AB* interakcija jedna polovina ove razlike dobije se:

$$AB = \frac{abc - bc + ab - b - ac + c - a + (1)}{4n}$$

Na sličan način dobiju se i peostale dve interakcije *BC* i *AC*.

$$AC = \frac{(1) - a + b - ab - c + ac - bc + abc}{4n}$$

$$BC = \frac{(1) + a - b - ab - c - ac + bc + abc}{4n}$$

I na kraju trofaktorska interakcija se definiše kao prosečna razlika između recimo interakcija *AB* za različite nivoe faktora *C*. Tako se dobije:

$$ABC = \frac{1}{4n}[(abc - bc) - (ac - c) - (ab - b) + (a - (1))]$$
$$= \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)]$$

U prethonim jednačinama vrednosti u zagradama predstavljaju odgovarajuće kontraste kombinacija tretmana. Pomoću kontrasta moguće je napraviti tabelu sa svim mogućim kombinacijama tretmana. (Tabela 43) Prvo se unesu odgovarajući pluseve i minusi u kolone koje odgovaraju glavnim efektima faktora *A, B* i *C,* i predstavljaju viši i niži nivo faktora. Kada se to uradi preostale kolone se mogu dobiti prostim množenjem znakova u odgovarajućim vrstama. Na primer, znakovi za *AB* interakciju se dobiju tako što se pomnože znakovi iz prethodnih kolona za faktore *A* i *B.*

Tabela 3.5.

Kombinacije tretmana	Faktorijalni efekti							
	I	*A*	*B*	*AB*	*C*	*AC*	*BC*	*ABC*
(1)	+	-	-	+	-	+	+	-
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
ab	+	+	+	+	-	-	-	-
c	+	-	-	+	+	-	-	+
ac	+	+	-	-	+	+	-	-
bc	+	-	+	-	+	-	+	-
abc	+	+	+	+	+	+	+	+

Tabela 43 ima nekoliko interasntnih karakteristika:

- osim kolone *I,* sve ostale kolone imaju isti broj pluseva i minusa,
- zbir proizvoda znakova za svake dve kolone je nula,
- kolona *I* je neutralni element za množenje, jer kada se njeni znakovi pomnože sa odgovarajućim znakovima neke druge kolone, ta kolona ostaje nepromenjena,
- proizvod bilo koje dve kolone daje kao rezultat neku drugu kolonu iz tablice, npr. $A \times B = AB$ i $AB \times B = AB^2 = A$.

Suma kvadrata odstupanja efekata se lako računa. S obzirom da svaki efekat ima jedan stepen slobode u 2^3 ogledu suma kvadrata za bilo koji efekat je:

$$SS = \frac{(Kontrast)^2}{8n}$$

3.2.1. Primer 2^3 faktorijalnog ogleda

U ovom odeljku daćemo primer jednog 2^3 ogleda. Analiziraćemo eksperriment u kome se posmatra uticaj tri faktora na vrednost obeležja pri čemu naravno faktori imaju po dva nivoa delovanja. Naime, fabrika za proizvodnju i punjenje bezalkoholnih pića želi da se u procesu proizvdnje, punjenja flašica sa sokovima odradi što je moguće preciznije. Mašina koja puni, teoretski to radi tako što svaku flašicu puni do određene zadate visine, ali u praksi, postoje odstupanja od ove zadate visine i fabrika bi želela da otkrije izvor te varijacije kako bi poboljšala proces punjenja.

Inžinjeri koji rukovode procesom punjenja mogu da kontrolišu tri faktora: procenat karbonizacije (faktor *A*), pritisak pod kojim se vrši punjnje (faktor *B*) i brzina punjenja flišica u minuti ili brzina trake (faktor *C*). Pritisak i brzina se mogu lako kontrolisati i po potrebi menjati, ali procenat karbonizacije je teže kontrolisati jer zavisi od temperature pod kojom se punjenje vrši. Međutim, za potrebe eksperimenta posmatraće se dva nivoa pritiska (25 psi i 30 psi, gde je jedinica psi „The pound per square inch" i njena vrednost je 1 psi = 6894.757 Pa), dva nivoa brzine (200 bpm i 250 bpm, gde je bpm „**beats per minute**", tj. 200 i 250 flašica u minuti), i dva nivoa karbonizacije (10% i 12%) . Eksperiment je ponavljan dva puta i to za sve moguće kombinacije nivoa ovih faktora. U Tabeli 3.6. prikazani su dobijeni podaci nakon izvođenja eksperimenta.

Tabela 3.6.

		Pritisak			
		25 psi		**30 psi**	
		Brzina trake		**Brzina trake**	
		200	**250**	**200**	**250**
Procenat karbonizacije	**10%**	-3	-1	-1	1
		-1	0	0	1
		-4=(1)	-1=*c*	1=*b*	2=*bc*
	12%	0	2	2	6
		1	1	3	5
		1=*a*	3=*ac*	5=*ab*	11=*abc*

Koristeći totale naznačene u Tabeli 3.6. za svaku kombinaciju nivoa dobiju se sledeći efekti faktora

$$A = \frac{1}{4n}(a - (1) + ab - b + ac - c + abc - bc) = 3{,}00$$
$$B = \frac{1}{4n}(b - (1) + ab - a + bc - c + abc - ac) = 2{,}25$$
$$C = \frac{1}{4n}(c - (1) + ac - a + bc - b + abc - ab) = 1{,}75$$

zatim efekti interakcija dva efekta

$$AB = \frac{abc - bc + ab - b - ac + c - a + (1)}{4n} = 0{,}75$$
$$AC = \frac{(1) - a + b - ab - c + ac - bc + abc}{4n} = 0{,}25$$

$$BC = \frac{(1) + a - b - ab - c - ac + bc + abc}{4n} = 0{,}50$$

i na kraju efekat interakcije sva tri faktora

$$ABC = \frac{1}{4n}[(abc - bc) - (ac - c) - (ab - b) + (a - (1))]$$
$$= \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)] = 0{,}50$$

Iz dobijenih rezultata vidi se da su efekti pojedinčnih faktora veći u odnosu na efekte međusobnih interakcija faktora, što upućuje na njihov veći uticaj. Ovo će biti provereno analizom varijanse u nastavku.

Koristeći softverski program *Statistica 7.0* dobiju se podaci prikazani u narednoj Tabeli 3.7.

Tabela 3.7. Analiza varijanse

	Univariate Tests of Significance forOdstupanja od zadate visine Sigma-restricted parameterization Effective hypothesis decomposition				
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	16.00000	1	16.00000	25.60000	0.000977
Procenat karbonizacije	36.00000	1	36.00000	57.60000	0.000064
Pritisak	20.25000	1	20.25000	32.40000	0.000459
Brzina trake	12.25000	1	12.25000	19.60000	0.002205
Procenat karbonizacije*Pritisak	2.25000	1	2.25000	3.60000	0.094350
Procenat karbonizacije*Brzina trake	0.25000	1	0.25000	0.40000	0.544737
Pritisak*Brzina trake	1.00000	1	1.00000	1.60000	0.241504
Procenat karbonizacije*Pritisak*Brzina trake	1.00000	1	1.00000	1.60000	0.241504
Error	5.00000	8	0.62500		

Podaci iz Tabele 45 su rezultati do kojih se došlo analizom varijansi, i ti rezultati potvrđuju prethodno donešene zaključke. Drugim rečima, vidi se da su *p*- vrednosti za faktore male što upućuje na to da svi faktori pojedinačno bitno utiču na vrednost posmatranog obeležja. Od interakcija najmanju *p*- vrednost ima *AB* interakcija, tj. interakcija između procenta karbonizacije i pritiska., pa postoji manji uticaj ove interakcije na vrednost obeležja. Za ostale interakcije *p* - vrednost su značajno veće i oni ne utiču bitnije na preciznost punjenja flašica.

3.3. Opšti slučaj, 2^k faktorijalni ogledi

U prethodna dva odeljka predstavaljeni su najjednostavniji slučajevi 2^k faktorijalnih ogleda gde je k uzimalo najmanje vrednosti, odnosno 2 i 3. Sada posmatrajmo opšti slučaj za $k > 3$. Statistički model za 2^k faktorijalni ogled podrazumeva k glavnih efekata, $\binom{k}{2}$ dvofaktorskih interakcija, $\binom{k}{3}$ trofaktorskih interakcija,..., $\binom{k}{k-1}$ *k-1*-faktorskih interakcija i jednu k-faktorsku interakciju. Za 2^k faktorijalni ogled to znači ukupno $2^k - 1$ efekat. Koristićemo istu notaciju kao i do sada za obeležavanje kombinacija tretmana. Na primer u 2^5 ogledu oznaka *abd* znači da su faktori *A*, *B* i *D* na višem nivou i faktori *C* i *E* na nižem nivou. Kombinacije tretmana se mogu zapisivati u standardnom redosledu tako što se faktori uvode postepeno jedan po jedan i to na način da se kombinuju sa prethono uvedenim faktorima. Na primer, standardan redosled kombinacija tretmana za 2^4 ogled je $(1), a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, abcd$.

Da bismo definisali efekte faktora ili izračunali njihove sume kvadarata odstupanja moramo prvo definisati odgovarajuće kontraste. U tu svrhu uvek požemo koristiti tabelu sa plusevima i minusima kao što je bilo rađeno kod slučajeva sa dva i tri faktora. Međutim, za veliko k ovaj način može biti i previše komplikovan, pa ćemo pokazati alternativni način.
U opštem slučaju, recimo za efekat interakcije $AB \cdots K$ kontrast se može dobiti razvijajući desnu stranu sledeće jednačine:

$$Kontrast_{AB\cdots K} = (a \pm 1)(b \pm 1) \cdots (K \pm 1)$$

gde će na kraju kada se sve zagrade pomnože broj 1 biti zamenjen simbolom (1) koji predstavlja kombinaciju svih faktora na nižim nivoima. Znak ispred broja 1 u svakoj od zagrada je negativan ako je taj faktor u $AB \cdots K$ interakciji i pozitivan ako taj faktor nije u toj interakciji. Na primer, u 2^3 ogledu, kontrast za interakciju *AB* će biti:

$$Kontrast_{AB} = (a-1)(b-1)(c+1) = abc + ab + c + (1) - ac - bc - a - b$$

ili recimo u 2^5 ogledu kontrast za interakciju *ABCD* će biti:

$$\begin{aligned} Kontrast_{ABCD} &= (a-1)(b-1)(c-1)(d-1)(e+1) \\ &= abcde + cde + bde + ade + bcd + ace + abe + e + abcd + cd + bd + ad \\ &\quad + bc + ac + ab + (1) - a - b - c - abc - d - abd - acd - bcd - ae - be \\ &\quad - ce - abce - de - abde - acde - bcde \end{aligned}$$

Nakon definisanja kontrasta lako je defenisati efekte i sumu kvadrata odstupanja za interakcije, tj. biće redom:

efekat interakcije $\quad AB \cdots K = \frac{2}{n \cdot 2^k}(Kontrast_{AB\cdots K})$

i

suma kvadrata odstupanja interakcije $SS_{AB\cdots K} = \frac{1}{n \cdot 2^k}(Kontrast_{AB\cdots K})^2$

gde *n* predstavlja broj ponavljanja eksperimenta. Analiza varijanse za 2^k oglede je predstavljena u sledećoj tabeli:

Tabela 3.8. Analiza varijanse za 2^k oglede sa $k > 3$ nivoa

Izvor varijacije	**Suma kvadarata**	**Stepen slobode**
k glavnih efekata		
A	SS_A	1
B	SS_B	1
$\vdots$	$\vdots$	$\vdots$
K	SS_K	1
$\binom{k}{2}$ dvofaktorskih interakcija		
AB	SS_{AB}	1
AC	SS_{AC}	1
$\vdots$	$\vdots$	$\vdots$
JK	SS_{JK}	1
$\binom{k}{3}$ trofaktorskih interakcija		
ABC	SS_{ABC}	1
ABD	SS_{ABD}	1
$\vdots$	$\vdots$	$\vdots$
IJK	SS_{IJK}	1
$\vdots$	$\vdots$	$\vdots$
$\binom{k}{k}$=1 k-faktorijalna interakcija		
$ABC \cdots K$	$SS_{ABC\cdots K}$	1
Greška	SS_E	$2^k(n-1)$
Total	SS_T	$n \cdot 2^k - 1$

3.3.1. 2^k ogled sa jednim ponavljanjem

Čak i za umereno mali broj faktora, ukupan broj kombinacija tretmana ili nivoa u 2^k ogledima je veliki. Na primer, u 2^5 ogledu je ukupno 32 kombinacije tretmana, u 2^6 ogledu je 64 mombinacije tretamana i tako dalje. S obzirom na ograničene mogućnosti, istraživači pribegavaju smanjenju broja ponavljanja eksperimenta i to najčešće tako što sprovode ogled sa samo jednim ponavljanjem osim ako nisu voljni da izostave neke od faktora. Fsktorijalni ogledi 2^k bez ponavljanja se često nazivaju neponavljajući faktorijalni ogledi. Sa samo jednim izvođenjem, tj. bez ponavljanja nema izračunavanja greške. Jedan od pristupa ovakvim ogledima jeste da se pretpostavi da su efekti određenih višefaktorskih interakcija zanemarljivi i da se kombinovanjem njihovih srednjih vrednosti kvadarata izračuna greška. To podrazumeva da sistemom dominiraju glavni efekti i višefoktorske interakcije sa malim brojem faktora, dok su efekti višefaktorskih interakcija sa većim brojem faktora zanemarljivi.

Kada se analiziraju 2^k faktorijalni ogledi bez ponavljanja ponekad je moguć efekat višefaktorske interakcije većeg broja faktora. Upotreba greške srednje vrednosti kvadrata dobijene udruživanjem višefaktorskih interakcija je nemoguća. Metod analize koji je napravio Daniel (1959) omogućava jednostavno rešenje kojim se ovaj problem prevazilazi. On predlaže crtanje vrednosti efekata sa normalnom raspodelom. Tako, efekti koji su zanemarljivi imaju normalnu raspodelu sa očekivanjem nula i disperzijom σ^2 i imaju tendenciju pada oko prave na grafiku, a značajni efekti će imati normalnu raspodelu sa nenula očekivanjem i biće rasuti daleko od prave na grafiku. Ovaj metod ćemo ilustrovati na sledećem primeru.

3.3.1.1. Primer 2^k ogleda sa jednim ponavljanjem

Hemijska supstanca se proizvodi u posudama koje su pod pritiskom. Faktorijalnim ogledom želi se ustanoviti koji faktori utiču na količinu supstance koja se dobije u procesu, odnosno isfiltrira. Četiri faktora su u pitanju: temperatura (faktor A), pritisak (faktor B), koncetracija formaldehida (faktor C) i ubrzanje procesa (faktor D). Svaki faktor ima dva nivoa dejstva. Inžinjeri žele da se ovim procesima dobije što je veći mogući nivo filtracije, odnosno da se dobije što je više moguće supstance iz istog procesa. Trenutni laboratorijski uslovi omogućavaju dobijanje oko 75 gal/h (gde je gal/h jedinica mere – galona po satu, 1 galon = 3,78541 litara). Takođe, trenutno se koristi koncetracija formaldehida na višem nivou i namera inžinjera je upravo da redukuje tu koncetraciju, da je smanji što je moguće više. Dosadašnji pokušaji smanjenja koncetracije formaldehida nisu dali odgovarajuće rezultate, pa je tako zbog uticaja drugih faktora nivo filtracije uvek padao.
Rezultati sprovođenja 2^4 ogleda sa jednim ponavljanjem dati su u Tabeli 3.9. Analizu ćemo započeti crtanjem Grafika 3.6., odnosno upoređujući dobijene vrednosti efekata sa normalnom raspodelom.

Tabela 3.9.

Faktor				Kombinacija nivoa	Nivo filtracije (gal/h)
A	*B*	*C*	*D*		
-	-	-	-	*(1)*	45
+	-	-	-	*a*	71
-	+	-	-	*b*	48
+	+	-	-	*ab*	65
-	-	+	-	*c*	68
+	-	+	-	*ac*	60
-	+	+	-	*bc*	80
+	+	+	-	*abc*	65
-	-	-	+	*d*	43
+	-	-	+	*ad*	100
-	+	-	+	*bd*	45
+	+	-	+	*abd*	104
-	-	+	+	*cd*	75
+	-	+	+	*acd*	86
-	+	+	+	*bcd*	70
+	+	+	+	*abcd*	96

Nakon sprovođenja procedure opisane u odeljku 3.3. dobiju se sledeći rezultati efekata faktora i njihovih interakcija.

Tabela 3.10.

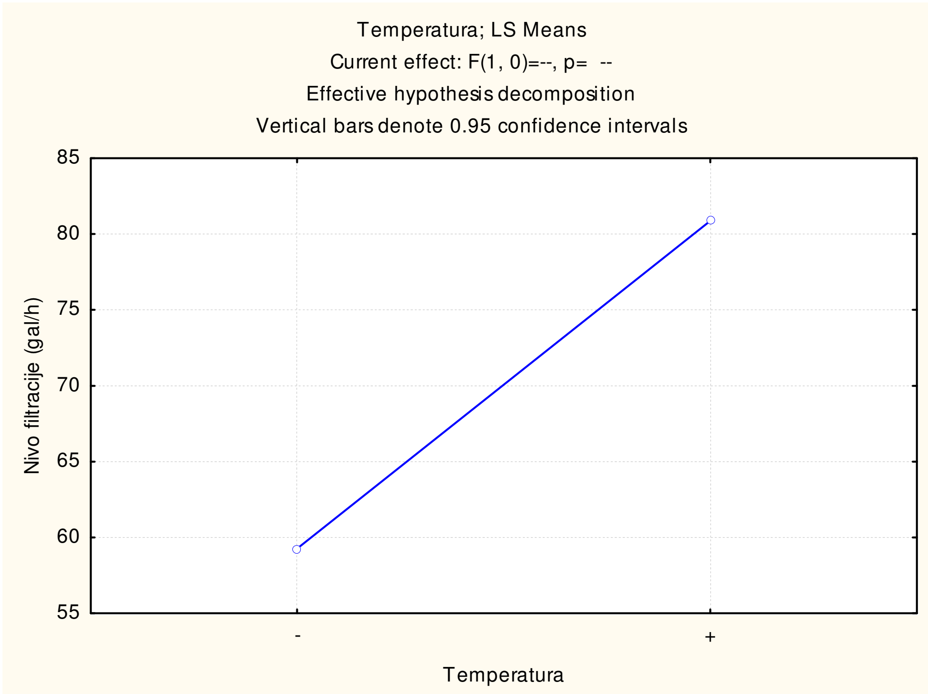
Redosled	Efekat	Vrednost	$(j-0{,}5)/15$
15	A	21.63	0.9667
14	AD	16.63	0.9000
13	D	14.63	0.8333
12	C	9.88	0.7667
11	ABD	4.13	0.7000
10	B	3.13	0.6333
9	BC	2.38	0.5667
8	ABC	1.88	0.5000
7	$ABCD$	1.38	0.4333
6	AB	0.13	0.3667
5	CD	-0.38	0.3000
4	BD	-1.13	0.2333
3	ACD	-1.63	0.1667
2	BCD	-2.63	0.1000
1	AC	-18.13	0.0333

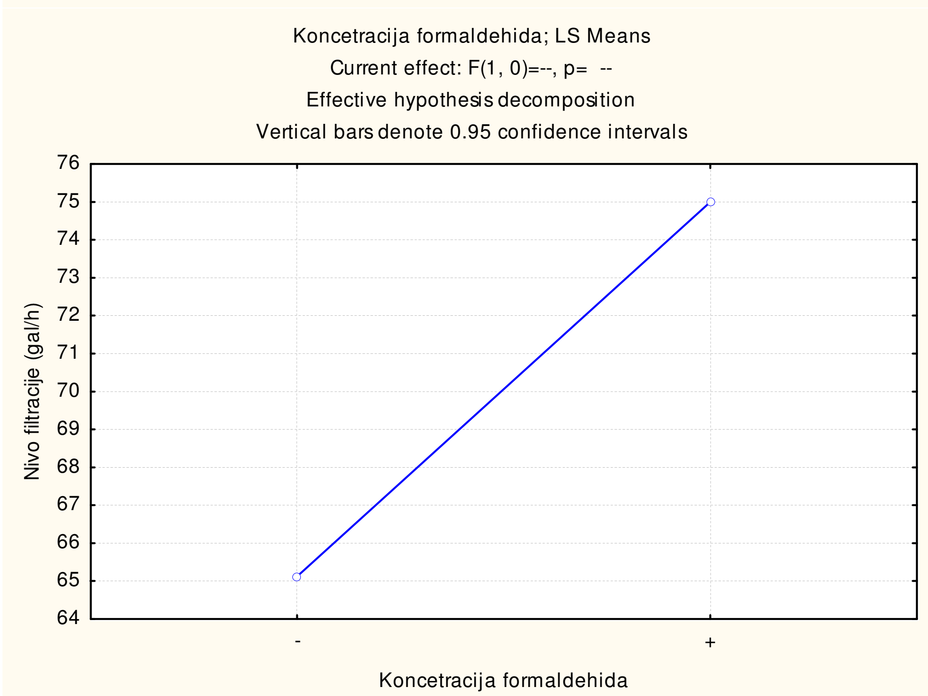
Grafik 3.6. prikazuje normalnu raspodelu i dobijene efekte faktora i njihovih intrakcija

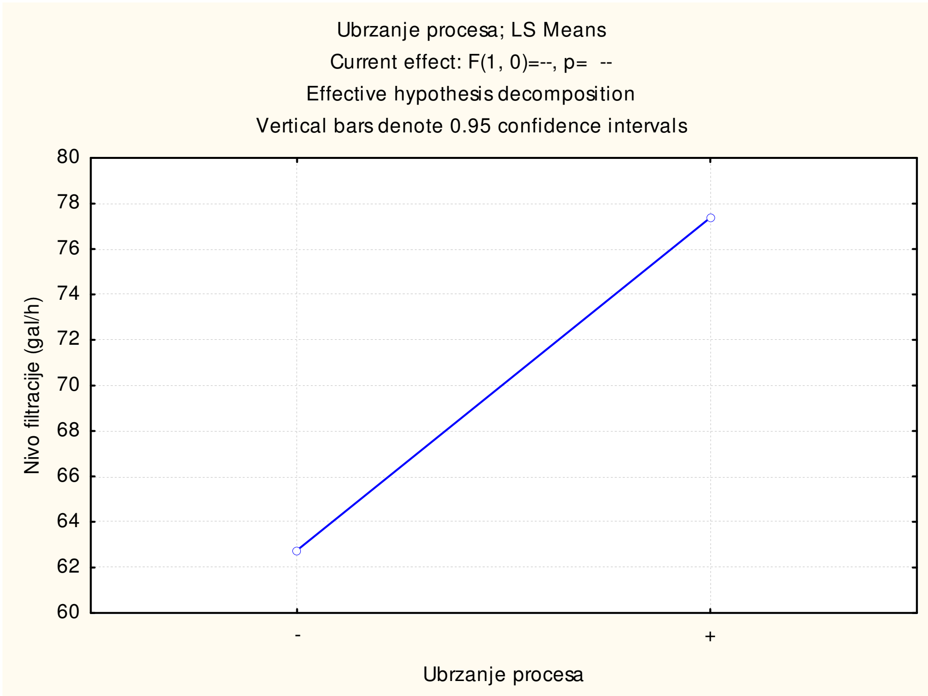
Grafik 3.6.

Svi efekti koji su blizu prave su zanemarljivog uticaja, dok oni koji imaju značajan utica su daleko od prave. Kao što se vidi veliko odstupanje od prave imaju efekti faktora A, C i D, i interakcija AC i AD. Na Grafiku 3.7. su srednje vrednosti efekata za faktor A, C i D.

Grafik 3.7.

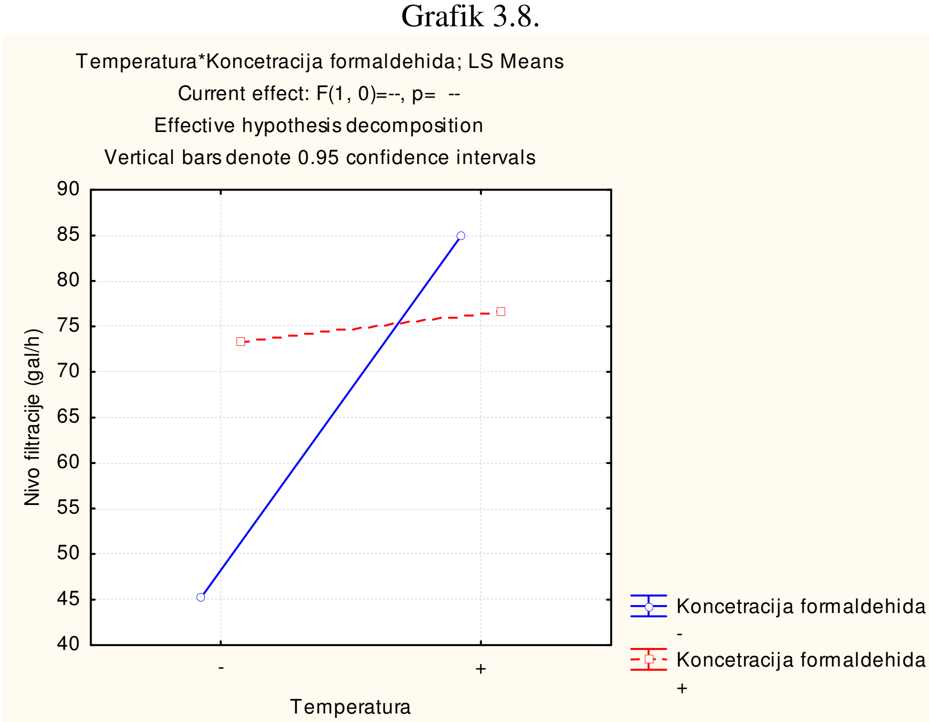

Temperatura; LS Means
Current effect: F(1, 0)=--, p= --
Effective hypothesis decomposition
Vertical bars denote 0.95 confidence intervals
Nivo filtracije (gal/h)
85
80
75
70
65
60
55
-
+
Temperatura


Koncetracija formaldehida; LS Means
Current effect: F(1, 0)=--, p= --
Effective hypothesis decomposition
Vertical bars denote 0.95 confidence intervals
Nivo filtracije (gal/h)
76
75
74
73
72
71
70
69
68
67
66
65
64
-
+
Koncetracija formaldehida

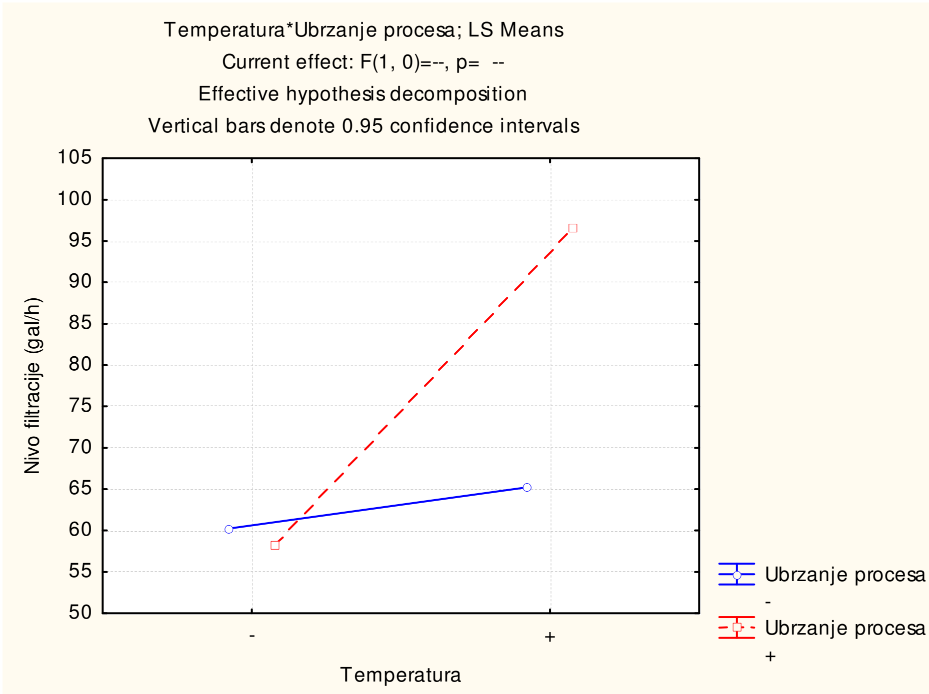


Svi ovi efekti su pozitivni i možemo da zaključimo da je u svrhu povećanja nivo filtracije dovoljno da se primenjuju ovi faktori na višem nivou. Međutim, uvek je neophodno uzeti u obzir i interakciju između faktora. Podsetimo da efekti faktora nemaju većeg značaja kad je interakcija značajna. Na Grafiku 3.8. i Grafiku 3.9. prikazane su interakcije *AC* i *AD*. Ove interakcije su od ključne važnosti.

Grafik 3.8.



Grafik 3.9.



Iz interakcije *AC* vidi se da je efekat temperature dosta mali kada je koncetracija formaldehida na višem nivou, a veoma veliki kada je koncetracija na nižem nivou i da je najbolji rezultat dala kombinacija višeg nivoa temperature i nižeg nivoa koncetracije formaldehida. Iz interakcije *AD* vidi se da je efekat ubrzanja procesa mali kad je teperatura na nižem nivou i veliki kada je temperatura na višem nivou.

S toga, najviši nivo filtracije će se ostvariti kad su faktori *A* i *D* na višem nivou i faktor *C* na nižem nivou. Ovo bi omogućilo smanjenje koncetracije formaldehida na niži nivo, što je i bio početni cilj inžinjera.

Kako je uticaj faktora *B* dosta mali, kao i svi efekti interakcija u kojima učestvuje pomenuti fakor, možemo odbaciti faktor *B* i za preostale faktore uraditi ponovo faktorijalni ogled, ali sada 2^3 faktorijalni ogled sa dva ponavljanja. Ovu anlizu ponovićemo softverskim programom *Statisica 7.0,* a rezultati su dati u Tabeli 3.11. i potvrđuju zaključke prethodne analize.

Tabela 3.11.

	Univariate Tests of Significance for Nivo filtracije (gal/l Sigma-restricted parameterization Effective hypothesis decomposition				
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	78540.06	1	78540.06	3500.393	0.000000
Temperatura	1870.56	1	1870.56	83.368	0.000017
Koncetracija formaldehida	390.06	1	390.06	17.384	0.003124
Ubrzanje procesa	855.56	1	855.56	38.131	0.000267
Temperatura*Koncetracija formaldehida	1314.06	1	1314.06	58.565	0.000060
Temperatura*Ubrzanje procesa	1105.56	1	1105.56	49.273	0.000110
Koncetracija formaldehida*Ubrzanje procesa	5.06	1	5.06	0.226	0.647483
Temperatura*Koncetracija formaldehida*Ubrzanje procesa	10.56	1	10.56	0.471	0.512032
Error	179.50	8	22.44		

Glava 4

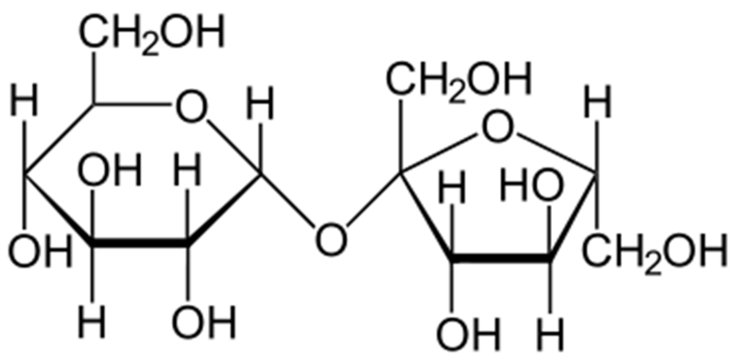
4. Primena faktorijalnih ogleda u poljoprivredi

4.1. Proizvodnja šećera iz šećerne repe

4.1.1. Uvod

Šećer je opšti pojam za klasu kristalnih i praškastih hemijskih supstanci karakterističnog slatkog ukusa koji se najčešće koriste u ishrani. Sama reč šećer potiče od persijske reči ***shakar.*** Hemijski šećeri su jednostavni ugljen hidrati: monosaharidi, disahariti i trisaharidi. U ishrani se najčešće koristi takozvani granularni šećer, ***saharoza,*** koji spada u disaharide i njegova hemijska formula je $C_{12}H_{22}O_{11}$.

Slika 4.1. Hemijska formula saharoze



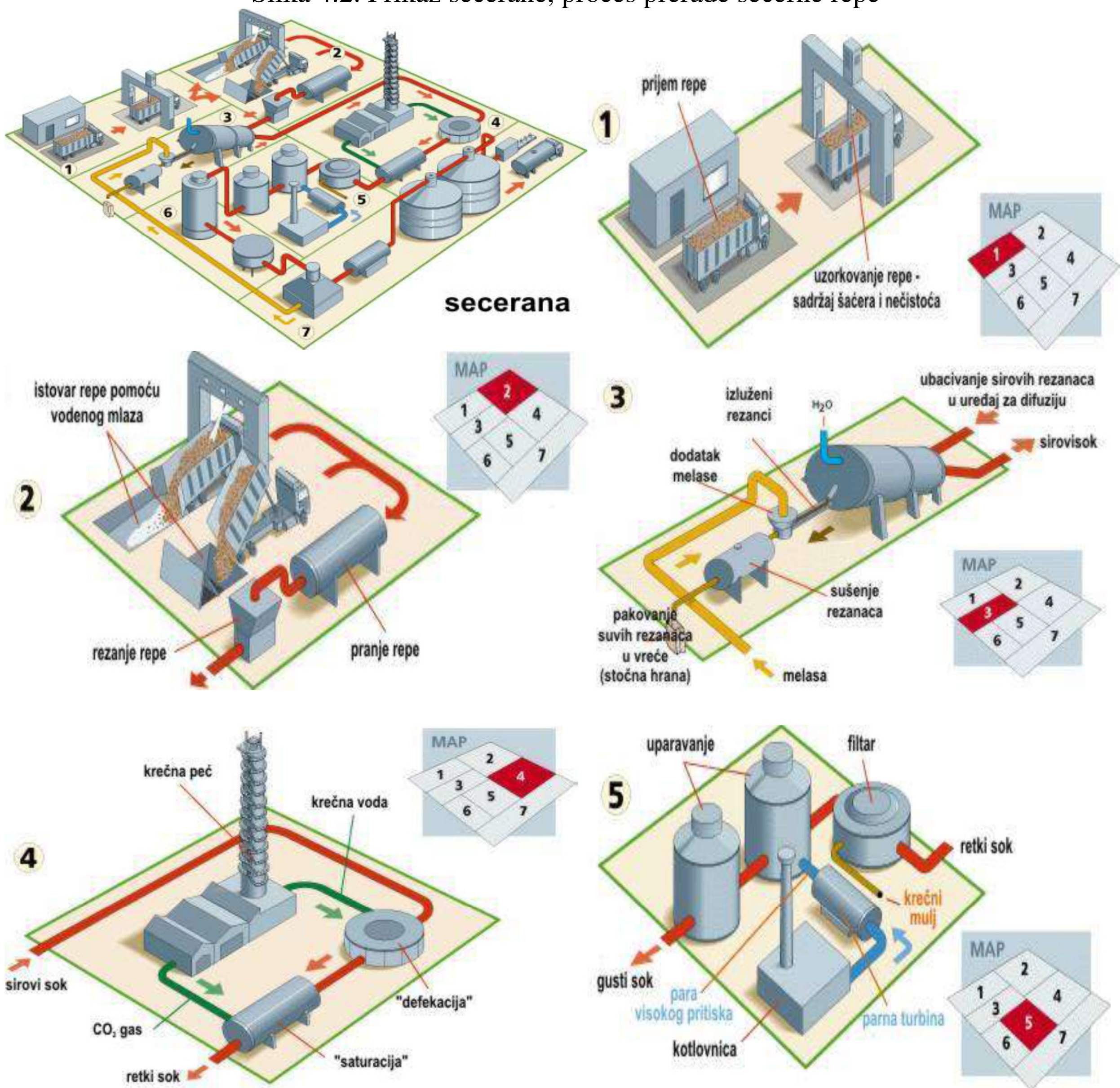
Prvi pisani podaci o šećeru datiraju iz 520. godine pre Hrista. Naime, nakon jednog pohoda na Indiju Persijski car Darije svedoči o postojanju neke vrste meda koji se dobija iz trske bez sudjelovanja pčela. U početku šećer se konzumirao žvakanjem sirove trske, da bi oko 350. godine bila otkrivena kristalizacija šećera i tada počinje razvoj proizvodnje čistog šećera na bazi trske. Proizvodnja šećera se vremenom širila na zapad i u Evrpu dolazi u srednjem veku. Tu se vršila prerada sirovog trščanog šećera koji se dopremao iz prekomorskih zemalja. Od 1747. godine počinje da se koristi šećerna repa kao sirovina za dobijanje šečera, da bi oko 1880. godine i definitivno potisnula trstiku kao sirovinu. [7]

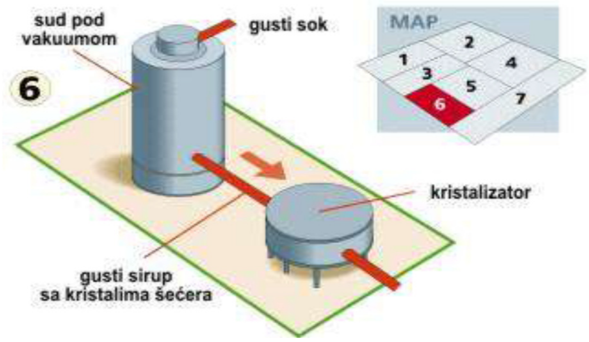
4.1.2. Tehnologija šećera

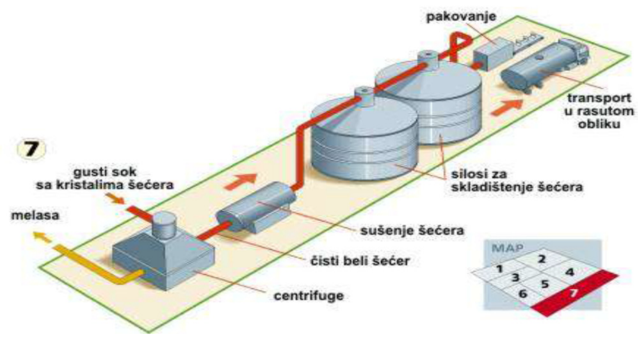
Šećer se danas u svetu prozvodi u više od 100 zemalja. Kao krajnji proivod dobija se preradom iz dve osnovne sirovine, a to su šećerna trska koja ima udeo od 78% i šećerna repa koja ima preostali udeo od 22%. Ukupna svetska proizvodnja šećera prelazi 160 miliona tona godišnje (2008), a najveći pojedinačni proizvođači su Brazil, Indija, Kina, Tajland, SAD, Meksiko te zemlje Evropske unije. U Srbiji je proizvodnja šećera organizovana u devet šećerana i sve one šećer dobijaju preradom isključivo šećerne repe. Godišnja proizvodnja je oko 400000 tona od čega je 220000 tona namenjeno izvozu.

Šećer se dobija preradom šećerne repe u šećeranama. To su fabrike koje imaju kapacitet da prerade i više hiljada tona sirove repe. Proizvodnja je tipična kampanjska, počinje u jesen, vađenjem repe na obližnjim popljoprivrednim dobrima i traje neprekidno, dan i noć, bez i jednog sekunda zastoja sve dok se ne preradi čitava količina raspoložive šećerne repe. Kampanja traje obično od 90 do 120 dana što ostavi značajnog traga na fabriku, i njoj pedstoji ozbiljan remont do sledeće kampanje. Tokom remonta sve mašine i aparati se rastavljaju, te zamenjuju, modernizuju ili popravljaju, tako da je fabrika posle svakog remonta praktično potpuno nov objekat, odnosno, nova šećerana. Tok proizvodnje šećera iz šećerne repe u šećeranama [7]:

Slika 4.2. Prikaz šećerane, proces prerade šećerne repe







4.1.3. Proizvodnja i priprema šećerne repe za preradu

Šećerna repa ili na latinskom ***Beta vulgaris saccharifera,*** je osnovna sirovina koja se u Srbiji koristi za dobijanje šećera. Šećerom najbogatiji deo šećerne repe jeste njen koren. Kod uzgoja šećerne repe vodi se računa da ona bude dobrog tehnološkog kvaliteta, sa što većim sadržajem saharoze i što većim prinosom mase korena po jedinici površine.

Slika 4.3. Šećerna repa

Kada je u pitanju tehnološki kvalitet, misli se na to koja je sorta repe u pitanju, kvalitet semena, osobine zemljišta, klimatski uslovi, primena agrotehničkih mera, te obrada zemljišta, đubrenje i zaštita repe.

Međunarodno priznata kvalifikacija sorti šećerne repe poznaje tri tipa sorti:

- prinosne sorte (E) koje daju najveći prinos mase korena i lišća po hektaru, kasnije postižu tehnološku zrelost, a sadržaj šećera, digestija, niža im je u odnosu na normalne sorte,
- normalne sorte (N) koje se po masi korena i lišća, te po sadržaju šećera nalaze između prinosnih i šećernih sorti, i
- šećerne sorte (Z) koje imaju znatno veći sadržaj šećera, manji prinos korena i lišća po hektaru i pogodne su za gajenje u područjima sa kraćim vegetacionim periodom, jer ranije dostižu tehnološku zrelost.

Seme šećerne repe nalazi se u plodu koje donosi biljka u drugoj godini vegetacije. Plod je složen, naziva se ***klube*** i u njemu se nalazi od 3 do 7 semenki. Najvažnija osibina semena jeste njgova klijavost koja mora biti veća od 80%. Seme u toku čuvanja ne sme da sadrži više od 15% vlage, a čistoća semena ne sme biti manja od 97%.

Što se tiče osobina zemljišta neophodnih za uspešnu proizvodnju šećerne repe neophodna su neutralna i slabo kisela plodna zemljišta:

- ph ~ 7.2,
- dobrih fizičkih i hemijskih svojstava, mora sadržati mineralne materije (P, N, K, Ca, Mg, S, Fe i druge),
- u minimalnim količinama mora sadržati mikroelemente za stimulativni rast biljke (Mn, Al, Si, Cu, B, Cl), te
- mora se voditi računa o plodoredu kultura, tj. šećerna repa se može sejati na istom zemljištu svake 4. ili 5. godine.

Klimatski uslovi su važni faktori za nicanje biljke i njen razvoj i u tom smislu najveću ulogu imaju tempratura i vlažnost zemljišta. Vegetacioni period šećerne repe je od 150 do 180 dana i u tom periodu temperatura ne bi smela da padne ispod $4\text{-}5^0$C i da pređe maksimalnih $28\text{-}30^0$C.

Setva šećerne repe kod nas s obavlja u periodu od 15. marta do 15. aprila kada su temperature vazduha i zemlje dovoljno visoke da seme može da klija. Seje se u redovima, sa razmakom od 40 do 50 cm i razmakom između biljaka od 20 do 25 cm. Najpovljnija dubina setve je od 2 do 4 cm, i zavisi da li je zemljište zbijeno ili je rastresito. Kada šećerna repa nikne razvije prve listove, pristupa se prvom okopavanju između redova. Nakon toga sledi proređivanje tako da biljke ostanu na odgovarajućim udaljenostima. Od tog perioda pa do sklapanja lišća, šećerna repa se okopava još najmanje 2 do 3 puta.
Visoki prinosi šećerne repe preko 40 tona po hektaru postižu se upotrebom mineralnog i stajskog đubriva. Mineralna đubriva koja se upotrebljavaju su azotna, fosforna kalijumova đubriva. Količina, vrsta i vreme unošenja đubriva zavisi od prethodne hemijske analize zemljišta. Šećerna repa od nicanja pa sve do prerade izložena je napadima različitih štetočina, te bolesti koje uzrokuju virusi, gljivice i bakterije. Za suzbijanje štetočina i bolesti primenjuju se prskanja ili zaprašivaja preparatima kao što su: DDT na bazi dihloridfeniltrihlor etana i HCH preparati.

Šećerna repa je industrijska biljka koja botaničku zrelost dostiže u drugoj godini života nakon sazrevanja ploda i semena. Međutim, šećerna repa se uzgaja zbog šećera, a njegov sadržaj u repi je najveći u prvoj godini vegetacije, pa tada biljka dostiže tehnološku zrelost. Data je masa korena i sadržaj šećera u njoj najveći, tj. najmanji je odnos lisne mase u odnosu na masu korena i tada je sadržaj nešećernih materija najpogodniji za tehnološku preradu.
Vađenje repe iz zemlje se obavlja mehanički upotrebom specijalnih mašina, doprema se do šećerana gde preostaje proces prerade i dobijanje finalnog proizvoda, šećera, opisanog u prethodnoj šemi. Repa se u šećeranama prvo istovara i čisti, odnosno odstranjuju se nečistoće kao što su zemlja, kamenje, lišće i slično. Nakon toga slede procesi u kojima se koren repe reže i dobijaju se takozvana rezanca. To se obavlja pomoću specijalnih noževa kako bi se dobili rezanci odgovarajućeg oblika i veličina. Nakon toga sledi procesi kojima se iz rezanaca izdvaja saharoza, tj. ekstakcija soka iz repe, čišćenje difuzionog soka, uparavanje soka i na kraju kristalizacija šećera. Kao krajnji proizvod dobija se sitni šećer u kristalnom obliku spreman za upotrebu u ishrani. [7]

4.2. Eksperimentalni deo

4.2.1. Prinos i kvalitet šećerne repe

Prinos šećerne repe ostvaren po hektaru kao i digestija, tj. kvalitet proizvedene sirovine zavisi od više faktora. Tu se pre svega misli na pravilnu agrotehniku, povoljne pedološke i klimatske uslove, te dobar izbor kvalitetnih hibrida. Za pravilan izbor hibrida potrebno je dobro poznavanje njegovih bioloških i proizvodnih svojstava, otpornost prema uzročnicima bolesti i reakciju prema osobinama tla. [8]

Pravilnim izborom hibrida šećerne repe može se bez dodatnih ulaganja povećati prinos i kvalitet korena repe, te pojeftiniti i olakšati preradu u fabrikama šećera, odnosno u šećeranama. Zbog toga, ulaganja u istraživanje proizvodnih osobina hibrida u pojedinim klimatskim i zemljišnim uslovima svakako ima opravdanje. Kada je u pitanju izbor hibrida, proizvođači najčešće u prvi plan stavljaju one sorte koje daju najveći prinos korena repe, ali sve češće uvažavaju i potrebe fabrika šećera za većim sadržajem šećera u korenu (digestijom) na što ih one podstiču cenom repe koja je stimulativna za veći sadržaj šećera. Takođe, neophodna je što veća tolerantnost hibrida na uzročnike bolesti koji dovode do značajnih smanjenja prinosa korena i šećera, te kvalitet repe. Još uvek nije dostignuta zadovoljavajuća tolerantnost prema uzročniku pegavosti lišća (C. Beticola Sacc.) koja i uz primenu fungocida redovno svake godine uzrokuje veće ili manje gubitke. Na kraju treba spomenuti i tip tla kao i primenu agrotehničkih mera. [9]

4.2.2 Primena faktorske analize u ispitivanju zavisnosti prinosa šećerne repe u odnosu na vrstu hibrida i količine padavina

Ovde ćemo razmotriti jedno istraživanje čiji je cilj da se otkrije koji hibrid šećerne repe ima najveću opravdanost za sejanjem na nekom području u različitim klimatskim uslovima, tj. posmatrano u različitim godinama. Iskoristićemo istraživanje koje je sprovedeno na lokaciji Dalj u periodu od 2009. do 2012. godine. [8]
Naime, izabrano je deset hibrida koji su najčešće zastupljeni. Vremenske prilike u godinama istraživanja su se znatno razlikovale. Jedna je bila s prosečnom, jedna s povećanom, a dve sa malim količinama padavina u odnosu na višegodišnji prosek. U sve četiri godine temperature vazduha tokom vegetacije bile su povišene i približno jednake.

Proizvodne vrednosti deset hibrida šećerne repe istraživane su putem sortnih mikro eksperimenata, postavljenih u razdoblju od 4 godine na plodnom tlu, černozemu na lokaciji Dalj. Ocena rodnosti izvršena je preko pokazatelja prinosa korena i šećera, ali i kvaliteta korena. Eksperiment je postavljen po šemi slučajnog blok rasporeda u 5 ponavljanja, a osnova svake parcele iznosila je 36 kvadratnih metara. Setva je u sve četiri godine obavljena u trećoj dekadi marta. Takođe, tretiranje fungocidima izvršeno je u približno istim terminima i to tri puta. Gnojidba je izvršena prema zalihaa hraniva u zemlji i potrebama biljaka šećerne repe. Tlo na kome je izvršen eksperiment je plodno tlo, tipa černozem, slabo je kiselo (pH~6,21), te sadrži 2,43% humusa. Godine u kojima su vršeni eksperimenti su se međusobno znatno klimatski razlikovali po količini padavina. Dve su godine bile sušne (2009. i 2011. godina), jedna je bila kišna (2010. godina), a jedna je bila prosečna (2012. godina). Što se tiče temperatura, one su bile više u odnosu na dugogodišnji prosek, i približno su bile iste za sve godine.

S obzirom da je eksperiment izvršen u realnim uslovima, treba imati na umu činjenicu da rezultati koji su dobijeni nisu sa sagirnošću pouzdani, jer veliki broj klimatskih činilaca je zanemaren, odnosno smatrani su fiksiranim za sve godine. Međutim, ovakav ekseriment može biti dobar prvi korak ka tome da se iz jednog velikog broja različitih sorti hibrida šećerne repe izdvoji manji broj (2 ili 3), koja će kasnije biti predmet dodatnih istraživanja. Uslovi u kojima su gajeni hibridi su tipični za poslednju deceniju koju karakterišu visoke, nadprosečne temperature sa razlčitim količinama padavina.

Na osnovu sprovedenog istraživanja dobijeni su rezultati prikazani u Tabeli 4.1. Iz rezultata dobijenih na terenu jasno se uočava velika razlika u prinosima svih sorti šećerne repe u zavisnosti od količine padavina. Najmanje prinose hibridi su davali kad je količina padavina bila ispod proseka, a drastično najveće onda kada je količina padavina bila znatno iznad dugogodišnjeg proseka. Međutim, razlike u prinosima nisu bile linearne, pa su tako hibridi različito reagovali. Narednom analizom prvo ćemo ustanoviti da li ima bitnih razlika među sortama hibrida, da li ima značajne interakcije između sorte hibrida i količine padavina, te koje sorte su najzahvalnije u smislu da pri različitim vremeskim uslovima daju najbolje rezultate.

Tabela 4.1. Prinos korena šećerne repe (u tonama po hektaru) u odnosu na sortu hibrida i količine padavina

		Količina padavina						
		Ispod proseka		**Prosečna**		**Iznad proseka**		$X_{i..}$
Sorte hibrida	**Libera**	56.5		81.64		113.06		
		55		80.12		111.51		
		54.5	**277.5**	79.64	**403.15**	111.05	**560.2**	**1240.85**
		55.1		80.24		111.68		
		56.4		81.51		112.9		
	Santino	43.6		84.54		104.3		
		42.1		83.02		102.75		
		41.6	**213**	82.54	**417.65**	102.29	**516.4**	**1147.05**
		42.2		83.14		102.92		
		43.5		84.41		104.14		
	Colonia KWS	52		80.74		106.55		
		50.5		79.22		105		
		50	**255**	78.74	**398.65**	104.54	**527.65**	**1181.3**
		50.6		79.34		105.17		
		51.9		80.61		106.39		
	Severina	63.6		93.88		121.89		
		62.1		92.36		120.34		
		61.6	**313**	91.88	**464.35**	119.88	**604.35**	**1381.7**
		62.2		92.48		120.51		
		63.5		93.75		121.73		
	Boomerang	65.3		98.56		111.22		
		63.8		97.04		109.67		
		63.3	**321.5**	96.56	**487.75**	109.21	**551**	**1360.25**
		63.9		97.16		109.84		
		65.2		98.43		111.06		
	Coyote	54.2		94.1		103.22		
		53.1		92.58		101.67		
		54.6	**268.5**	92.1	**465.45**	101.21	**511**	**1244.95**
		52.7		92.7		101.84		
		53.9		93.97		103.06		
	Predator	62.9		80.06		107.89		
		61.8		78.54		106.34		
		63.3	**312**	78.06	**395.25**	105.88	**534.35**	**1241.6**
		61.4		78.66		106.51		
		62.6		79.93		107.73		
	Asketa	53.3		84.86		106.09		
		52.2		83.34		104.54		
		53.7	**264**	82.86	**419.25**	104.08	**525.35**	**1208.6**
		51.8		83.46		104.71		
		53		84.73		105.93		
	Gazeta	54.5		91.65		99.69		
		53.4		90.13		98.14		
		54.9	**270**	89.65	**453.2**	97.68	**493.35**	**1216.55**
		53		90.25		98.31		
		54.2		91.52		99.53		
	Protekta	55.1		86.17		93.75		
		54		84.65		92.2		
		55.5	**273**	84.17	**425.8**	91.74	**463.65**	**1162.45**
		53.6		84.77		92.37		
		54.8		86.04		93.59		
	$X_{.j.}$	**2767.5**		**4330.5**		**5287.3**		$X_{...}$**=12385.3**

Prvo ćemo izračunati sume kvadrata totala, faktora, interakcije i greške na način opisan u poglavlju 2.1 :

$$SS_T = \sum_{i=1}^{10}\sum_{j=1}^{3}\sum_{k=1}^{5} X_{ijk}^2 - \frac{X_{...}^2}{10\cdot 3\cdot 5} = (56,5)^2 + (55)^2 + (55,1)^2 + \cdots + (93,59)^2 - \frac{(12385,3)^2}{150} = 71211,52$$

$$SS_{sorta\ hibrida} = \sum_{i=1}^{10} \frac{X_{i..}^2}{3\cdot 5} - \frac{X_{...}^2}{10\cdot 3\cdot 5} = \frac{(1240,85)^2 + (1147,05)^2 + \cdots + (1162,45)^2}{15} - \frac{(12385,3)^2}{150} = 3612,03$$

$$SS_{količina\ padavina} = \sum_{j=1}^{3} \frac{X_{.j.}^2}{10\cdot 5} - \frac{X_{...}^2}{10\cdot 3\cdot 5} = \frac{(2722,5)^2 + (4270,5)^2 + (5242,3)^2}{10\cdot 5} - \frac{(12385,3)^2}{150} = 64718,85$$

$$SS_{interakcija} = \sum_{i=1}^{10}\sum_{j=1}^{3} \frac{X_{ij.}^2}{5} - \frac{X_{...}^2}{10\cdot 3\cdot 5} - SS_{vrsta\ hibrida} - SS_{količina\ padavina}$$
$$= \frac{(277,5)^2 + (213)^2 + \cdots + (463,65)^2}{5} - \frac{(12385,3)^2}{150} - 3612,03 - 64718,85$$
$$= 2788,67$$

i na kraju $SS_E = SS_T - SS_{vrsta\ hibrida} - SS_{količina\ padavina} - SS_{interakcija} =$
$71211,52 - 3612,03 - 64718,85 - 2788,67 = 91,97$

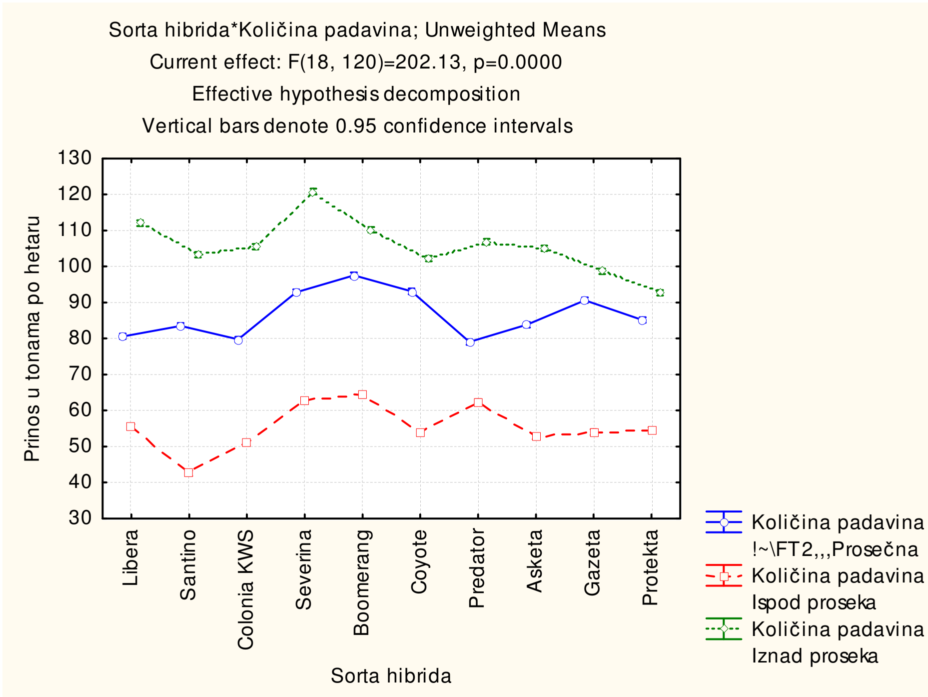
Rezultate analize varijanse prikazaćemo u Tabeli 4.2.

Tabela 4.2. Rezultati analize varijanse

Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	F_0
Vrsta hibrida	3612,03	9	401,34	523.63
Količina padavina	64718,85	2	32359,42	42219.88
Interakcija	2788,67	18	154,93	202.13
Greška	91,97	120	0,77	
Ukupno	71211,52	149		

Iz tablice za Fišerovu raspodelu dobije se da je $F_{0,05;18;120} = 1,69$ i iz toga sledi da postoji značajna interakcija između faktora, odnosno sorte hibrida i količine padavina. Šta više, $F_{0,05;9;120} = 1,96$ i $F_{0,05;2;120} = 3,07$ pa su i uticaji samih faktora značajni. Najveća razlika između terijskih i realizovanih vrednosti Fišerove raspodele je kod drugog faktora, odnosno količine padavina. To je i sasvim razumljivo. Međutim, kako se u realnoj proizvodnji šećerne repe na ovaj faktor ne može uvek delovati, nas će interesovati prvi faktor, sorta hibrida. Drugim rečima, daljom analizom pokušaćemo da izdvojimo najbolje sorte koje daju i najviši prinos za sve nivoe faktora, količina padavina. Da bi se bolje interpretirali rezultati analize, od pomoći je nacrtati grafik sa srednjim vrednostima za svaku kombinaciju nivoa.

Grafik 4.1.



Iz grafika se jasno uočava značajan uticaj koji ima količina padavina na prinose, tj. sve sorte imaju povećanje prinosa kako se povećava količina padavina. To povećanje je manje ili veće, zavisi od sorte koja se posmatra, ali je bitno istaći da je ono prisutno uvek i kod svih sorti šećerne repe.

Kada analiza varijanse ukazuje da postoje razlike između srednjih vrednosti faktora, obično je poželjno napraviti poređenje između pojedinačnih nivoa da bi se otkrile značajne razlike. Da bismo ovo izveli, primenićemo Dankanov test. Kako je u eksperimentu interakcija bila značajna, poređenja srednjih vrednosti nivoa za jedan faktor, u ovom slučaju sorti hibrida, mogu biti sakrivena u interakciji dva faktora. Dankanov test ćemo sprovesti pomoću softverskog programa *Statistica 7.0,* a rezultati su prikazani u Tabeli 4.3.

Tabela 4.3. Dankanov test

	Duncan test; variable Prinos u tonama po hetaru (Prvi primer 2) Approximate Probabilities for Post Hoc Tests Error: Between MS = .76645, df = 120.00										
Cell No.	Vrsta hibrida	{1} 82.723	{2} 76.470	{3} 78.753	{4} 92.113	{5} 90.683	{6} 82.997	{7} 82.773	{8} 80.573	{9} 81.103	{10} 77.497
1	Libera		0.000025	0.000045	0.000028	0.000045	0.425243	0.876063	0.000059	0.000104	0.000028
2	Santino	0.000025		0.000059	0.000017	0.000016	0.000017	0.000020	0.000045	0.000028	0.001835
3	Colonia KWS	0.000045	0.000059		0.000017	0.000020	0.000025	0.000028	0.000103	0.000059	0.000239
4	Severina	0.000028	0.000017	0.000017		0.000119	0.000059	0.000045	0.000020	0.000025	0.000016
5	Boomerang	0.000045	0.000016	0.000020	0.000119		0.000103	0.000059	0.000025	0.000028	0.000017
6	Coyote	0.425243	0.000017	0.000025	0.000059	0.000103		0.486260	0.000028	0.000045	0.000020
7	Predator	0.876063	0.000020	0.000028	0.000045	0.000059	0.486260		0.000045	0.000060	0.000025
8	Asketa	0.000059	0.000045	0.000103	0.000020	0.000025	0.000028	0.000045		0.100058	0.000059
9	Gazeta	0.000104	0.000028	0.000059	0.000025	0.000028	0.000045	0.000060	0.100058		0.000045
10	Protekta	0.000028	0.001835	0.000239	0.000016	0.000017	0.000020	0.000025	0.000059	0.000045	

Između sorti Libera, Coyote i Predator ne postoje značajne razlike, kao ni između sorti Asketa i Gazeta. U ostalim kombinacijama postoje značajne razlike. Na osnovu svega možemo zaključiti da se po prinosu šećerne repe značajno izdvajaju dve sorte hibrida: Severina i Boomerang. Za sve nivoe količine padavina ove dve sorte su uvek imale najviše prinose. Izuzetak je samo slučaj kad je količina padavina bila iznad proseka, i tada je sorta Boomerang imala nezntno manji prinos u odnosu na sortu Libera. Međutim, kada se uporede i drugi nivoi količine padavina, uočavju se značajne razlike među sortama u korist pomenutih sorti Severina i Boomerang.

Na kraju treba reći da rezultate ovog eksperiment treba pažljivo uzimati u obzir, jer je sproveden na posebnom zemljištu, tj. pod specifičnim pedološkim uslovima i o tim faktorima treba voditi računa pri izboru sorte hibrida šećerne repe. Kako ekonomska opravdanost gajenja ove kulture zahteva ne samo visok prinos, već i njen dobar kvalitet, odnosno procenat šećera koji sadrži koren repe, u nastavku ćemo se baviti analizom izdvojenih sorti hibrida iz ovog eksperimenta kako bismo utvrdili dodatno da li prinose prati digestija, odnosno količina šećera koja se može ekstraktovati.

4.2.3. Digestija šećerne repe

Mnogi faktori utiču na prinos i kvalitet šećerne repe. Istraživači su ispitivali uticaj mnogih faktora, kako njihov pojedinačni tako i njihov međusobni, interakcijski uticaj. Danas je sasvim jasno da se visoki prinosi u proizvodnji šećerne repe ne mogu ostvariti bez izolovanja bitnih faktora, kako bi se njihovim upravljanjem postigli maksimalni rezultati . [10]

U ovom odeljku biće reči o digestiji, tj. procentu šećera zastupljenog u korenu šećerne repe. Ovo je vrlo bitan parametar o kome treba voditi računa, jer utiče na otkupnu cenu repe, pa tako i na samu ekonomsku isplativost gajenja šećerne repe.
Do danas je razvijeno više metoda kvantitatine analize saharoze u šećernoj repi. Postoje direktne i indiretne analize. Najzastupljenije metode su: polarimetrija (saharimetri), rastvaranje izotopa, magnetna rezonanca (spoktroskopi), hromatografi, kolorimetri, spektrofotometri, analiza enzima i drugi. [11]

Tehnološki postupak proizvodnje šećerne repe je kontinuirani proces i odvija se po fazama koje su međusobno povezane u tehnološku celinu. Pored osnovnog tehnološkog procesa proizvodnje šećera, tehnologija uključuje i pomoćne procese i operacije u pogonima za proizvodnju energije, pomoćnih materijala i nusproizvoda. Tokom same prerade šećerne repe na više mesta u tehnološkom postupku vrše se određene laboratorijske analize.

Proizvodnja šećera započinje prijemom šećerne repe, njenim vaganjem i uzimanjem prvih uzoraka za analizu digestije (postotka šećera u repi) i relevantnih elemenata (azot, kalijum, natrijum), te za određivanje nečistoća koje dolaze sa repom (zemlja, trava, kamen). Uzorkovanje se vrši automatskom sondom, a dalji proces je poluautomatski, a kao rezultat se dobija postotak nečistoće repe i repna kaša koja se dalje analizira. Analizom repne kaše određuje se digestija, odnosno postotak saharoze, ali i količina kalijuma, natrijuma i drugih elemenata. Kontrolom šećerne repe na njenom samom prijemu u šećeranama omogućava se da tehnička služba izvrši neophodnu optimizaciju tehnološkog procesa za preadu sirovine kako bi se ostvarilo ekonomično poslovanje fabrike.

Daljim procesom nastaje više poluproizvoda i nusproizvoda. Ponovo se vrše analize na repnim rezancima, te se određuje digestija, te prisustvo kalcijumove soli, kiseonika, ugljendioksida, ugljenmonoksida i drugog.

Najbitniji proces u šećeranama je ekstrakcija šećera iz repe. Osnovni cilj ekstrakcije je izvlačenje najveće moguće količine saharoze iz sirovine. Ovim postupkom se iz rezanaca šećerne repe, odnosno iz repnog tkiva ekstahuje se šećer. Nakon toga slede niz procesa kao što su: procesi čišćenja, karbonizacije, te na kraju kristalizacije. Kao krajnji proizvod dobija se šećer u kristalnom obliku, ali i melasa koja predstavlja matični sirup iz centrifuga iz kojeg se više ne može na ekonomičan način iscrpeti saharoza. Sastoji se od vode, saharoze, glukoze, fruktoze, te raznih nešećera. Koristi se u proizvodnji stočne hrane, kvasca i alkohola.

4.2.4. Primena 2^k faktorske analize u ispitivanju prinosa i digestije šećerne repe

S obzirom da je šećerna repa industrijska biljka i da je kao takva namenjena fabričkoj preradi, odnosno proizvodnji šećera, proizvođači su opredeljeni za uzgajanje sorti šećerne repe koje daju ne samo najviše prinose već imaju i najveću digestiju.

U prethodnom primeru, faktorskom analizom smo utvrdili da postoje značajne razlike između sorti šećerne repe, količine padavina, kao i da postoji značajna interakcija ova dva navedena faktora kada su u pitanju prinosi šećerne repe po hektaru zasađene povšine. Ponovo ćemo se vratiti na ovaj eksperiment, ali će nas sada interesovati digestija, odnosno procenat šećera koji se može dobiti iz ovih sorti.

Primenićemo 2^2 faktorsku analizu na dobijene rezultate količine čistog šećera u tonama po hektaru, i to tako što ćemo za prvi faktor uzeti sortu hibrida sa dva nivoa: Severina i Boomerang. To su sorte za koje je prethodno utvrđeno da daju najveće prinose. Za drugi faktor ćemo opet uzeti količinu padavina, ali ovaj put samo dva nivoa: ispod proseka i prosečna. Razlog tome je taj što su prinosi kod svih sorti za nivo, iznad proseka, bili ekstremno visoki, te nije za očekivati da se u narednih nekoliko godina takvi vremenski uslovi ponove.

Na osnovu sprovedenog istraživanja dobijeni su rezultati koji su prikazani u Tabeli 4.4. U pitanju su dva faktora sa po dva nivoa i sa 5 ponavljanja.

Tabela 4.4. Količina čistog šećera u tonama po hektaru

		Količina padavina	
		Ispod proseka	**Prosečna**
Sorta hibrida	**Severina**	7.46 6.89 7.05 6.7 6.82	8.6 8.67 9.42 9.83 9.18
		(1)=34.92	**b**=45.7
	Boomerang	8.59 9.23 8.67 9.72 8.99	12.62 11.96 12.3 12.4 11.5
		a=45.2	**ab**=60.78

Prvo ćemo odrediti efekte faktora na način opisan u odeljku 3.1.

$$Sorta\ hibrida = \frac{1}{2n}[(ab-b)+(a-(1))] = \frac{1}{2n}[ab+a-b-(1)] = 25{,}36$$

$$Količina\ padavina = \frac{1}{2n}[(ab-a)+(b-(1))] = \frac{1}{2n}[ab+b-a-(1)] = 26{,}36$$

$$Interakcija\ faktora = \frac{1}{2n}[(ab-b)-(a-(1))] = \frac{1}{2n}[ab+(1)-a-b] = 4{,}8$$

Efekat prvog faktora (sorta hibrida) je pozitivan što ukazuje da prelazak sa nižeg nivoa faktora (sorta Severina) na viši nivo (sorta Boomerang), utiče na povećanje vrednosti obeležja, odnosno na povećanje količine čistog šećera koji se može dobiti po jednom hektaru.

Efekat drugog faktora (količina padavina) je takođe pozitivan i približno je iste veličine, što opet upućuje da povećanje količine padavina, odnosno prelazak sa nižeg nivoa faktora (ispod proseka) na viši nivo faktora (prosečna) utiče na povećanje vrednosti obeležja. Vrednost efekta interakcije faktora je 4,8 i upućuje na relativno mali uticaj u odnosu na efekte faktora sorta hibrida i količina padavina.

Dalje računamo sume kvadrata odstupanja za faktore interakciju faktora. Tako dobijamo da je:

Suma kvadrata odstupanja za faktor Sorta hibrida

$$SS_{Sorta\ hibrida} = \frac{[ab+a-b-(1)]^2}{4\cdot n} = 32{,}15648$$

suma kvadrata odstupanja za faktor Količina padavina

$$SS_{Količina\ padavina} = \frac{[ab+b-a-(1)]^2}{4\cdot n} = 34{,}74248$$

suma kvadrata odstupanja za interakciju faktora

$$SS_I = \frac{[ab+(1)-a-b]^2}{4\cdot n} = 1{,}152 \quad \text{, te}$$

sumu kvadrata totala ćemo izračunati na standardan način:

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} X_{ijk}^2 - \frac{X_{\cdots}^2}{4n} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{5} X_{ijk}^2 - \frac{X_{\cdots}^2}{20} = 71{,}0716$$

Generalno, SS_T ima $4n-1$ stepen slobode dok suma kvadrata greške SS_E ima $4(n-1)$ stepen slobode i uobičajeno se računa preko ostalih suma kvadrata:

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 3{,}02064$$

Kompletni rezultati analize varijanse predstavljeni su u sledećoj Tabeli 4.5.

Tabela 4.5. Rezultati analize varijanse

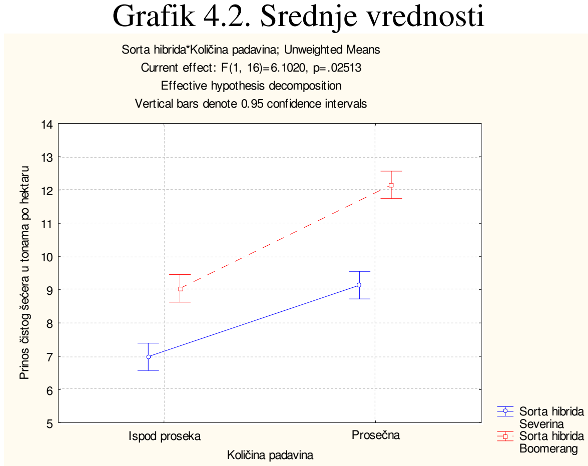
Izvori varijacije	Zbir kvadrata	Stepeni slobode	Sredina zbira kvadrata	F_0
Vrsta hibrida	32,15648	1	32,15648	170,3294
Količina padavina	34,74248	1	34,74248	184,0271
Interakcija	1,152	1	1,152	6,102018
Greška	3,02064	16	0,18879	
Ukupno	71,0716	19		

Iz tablice za Fišerovu raspodelu dobije se da je $F_{0,05;1;16} = 161{,}4476$ i iz toga sledi da ne postoji značajna interakcija između faktora, odnosno sorte hibrida i količine padavina. S druge strane uticaji samih faktora jesu značajni. Kako nas interesuje prvi faktor, odnosno sorta hibrida, uradićemo Dankanov test da potvrdimo tvrdnju da postoje značajne razlike među sortama i nacrtaćemo grafik srednjih vrednosti.

Tabela 4.6. Dankanov test (*Statistica7.0*)

	Duncan test; variable Prinos čistog šećera u tonama po hek Approximate Probabilities for Post Hoc Tests Error: Between MS = .18879, df = 16.000		
Cell No.	Sorta hibrida	{1} 8.0620	{2} 10.598
1	Severina		0.000159
2	Boomerang	0.000159	

Grafik 4.2. Srednje vrednosti



Dankanov test je potvrdio da postoje značajne razlike između sorti hibrida Severina i Boomerang, a iz grafika vidimo da je ta razlika u korist sorte Boomerang.

Na kraju da zaključimo. Ovim eksperimentom testirano je deset sorti hibrida u kombinaciji sa tri količine padavina. Prvim testiranjem izdvojene su dve sorte koje su davale značajno veće prinose šećerne repe u odnosu na ostale, a treba istaći da je i među njima bilo značajnih razlika u korist sorte Severina. Drugim rečima ako se gleda prinos šećerne repe po hektaru zasada, prednost treba dati upravo ovoj sorti.
U drugom testiranju pažnja je bila usmerena na količinu čistog šećera koji se mogao dobiti iz dve izdvojene sorte prethodnog testiranja. Posmatrali su se rezultti količine šećera u uslovima manje količine padavina i prosečne. Rezultati testiranja su ponovo pokazali da postoji značajan uticaj klimatskih razlika, odnosno količine padavina, ali takođe i da postoje značajne razlike među sortama. Ovaj put ta prednost bila je na strani druge sorte, Boomerang.

Ispostavilo se da bez obzira što sorta Severina daje najveću količinu šećerne repe, prednost ipak treba dati sorti Boomerang, jer sadrži znatno manje nečistoća i značajno veću količinu čistog šećera. Time je ta sorta ekonomski isplativija za sadnju.

Zaključak

Statistika je u današnjem svetu veoma značajna naučna disciplina, a s tim i svaka grana statistike dobija na značaju. Danas se većina stvari posmatra kroz statistiku; rade se procene kretanja cena na berzi, procenjuje se potencijaalno tržište, takođe i ponašanje ljudi se sagledava kroz statistiku, ona je opšte prisutna. Računski postupak je ponekad komplikovan i obiman, ali se sve većim razvojem razvojem računara i programa za statističku obradu podataka, više ne predstavlja problem.

Faktorijalni ogledi predstavljaju jedan od najznačajnijih statističkih metoda proučavanja uticaja više faktora na vrednost neke pojave koja se posmatra. U radu su prikazani neki specijalni slučajevi faktorijalnih ogleda kao i često korištena Analiza varijanse kao jedna od statističkih metoda koja je takođe našla svoje mesto u današnjim istraživanjima.

Primena faktorijalnih ogleda je veoma široka i raznovrsna. Uspešno se može primeniti u ekonomiji, medicini, tehnici, psihologiji, sociologiji, poljoprivredi, biologiji, vojnoj nauci i tako dalje. Faktorijalni ogledi su uneli revolucionarnost u eksperimentalnim istraživanjima masovnih pojava, posebno ekonomskih, gde je veoma teško ispitivanu pojavu razdvojiti od uticaja pretpostavljenih faktora.

Literatura

[1] Douglas C. Montgomer Design and Analysis of Experiments, 5th Edition, John Wiley & Sons, INC., 2001

[2] Box, G.E., Hunter,W.G., Hunter, J.S., *Statistics for Experimenters: Design, Innovation, and Discovery*, 2nd Edition, Wiley, 2005,

[3] Vladan Todić*, Analiza varijanse kroz primere, 2009.*

[4] Zagorka Lozanov-Crvenković, Danijela Rajter-Ćirić „Zbirka rešenih zadataka iz verovatnoće i statistike“, Prirodno matematički fakultet, Novi Sad, 1999.

[5] Stevan Hadživuković „Statističke metode“, Poljoprivredni fakultet, Novi Sad, 1991.

[6] Zoran A. Ivković ,,Matematička statistika“, Naučna knjiga, Beograd, 1980.

[7] Trochim, William M.K. 2006. "Factorial Designs." Research Methods Knowledge Base.

[8] Dražen Jurišić, Andrija Kristek, *Prinos i kvaliteta novih KWS hibrida šećerne repe,* Poljoprivreda, 2010.

[9] A. Kristek, Suzana Kristek, Renata Glavaš-Tokić, Manda Antunović, Sanda Rašić, I. Rešić, Ivana Varga, *Prinos i kvaliteta koijena istraženih hibrida šećerne repe*, Poljoprivreda 2010.

[10] George K. Ryser, *A regression study on tare samples of sugar beets in relation to factor influencing productivity and quality,* Journal of the A.S.S.B.T., 1965

[11] Douglas W. Lowman, *Bibliography: methods of sucrose analysis,* Journal of the A.S.S.B.T., 1987

Biografija

Vladan Todić rođen je 15.4.1986. godine u Tuzli. Osnovnu školu „Vuk Karadžić“, je završio u Bijeljini, 2001. godine kao nosilac Vukove diplome i đak generacije. Gimnaziju „Filip Višnjić“, takođe završava u Bijeljini četiri godine kasnije, odnosno 2005. godine. Tokom školovanja više puta je učestvovao na Republičkim i Saveznim takmičenjima iz fizike, matematike i istorije. U Novom Sadu 2005. godine upisuje Prirodno-matematčki fakultet, konkretno na Departmanu za matematiku i informatiku, upisuje matematiku, smer Matematika finansija. Osnovne studije završava 2009. godine , a iste godine upisuje master studije na smeru Primenjena matematika. Od 2010. godine radi kao asistent na Pedagoškom fakultetu u Bijeljini, pri Univerzitetu u Istočnom Sarajevu.

Univerzitet u Novom Sadu
Prirodno-matemtaički fakultet
Ključna dokumentacijska inforamacija

Redni broj:
RBR
Identifikacioni broj:
IBR
Tip dokumentacije: Monografska dokumentacija
TD
Tip zapisa: Tekstualni štampani materijal
TZ
Vrsta rada: Diplomski master rad
VR
Autor: Vladan Todić
AU
Mentor: Prof. dr Zagorka Lozanov-Crvenković
MN
Naslov rada: Primena faktorijalnih ogleda
NR
Jezik publikacije: Srpski (latinica)
JP
Jezik izvoda: s/e
JI
Zemlja publikovanja: Republika Srbija
ZP
Uže geografsko područje: Vojvodina
UGP
Godina: 2014.
GO
Izdavač: Autorski reprint
IZ
Mesto i adresa: Novi Sad, Departman za matematiku i informatiku, PMF, Trg Dositeja Obradovića 4.
MA
Fizički opis rada: (4, 89, 11, 54, 3, 37, 0)
Broj poglavlja, broj strana, br. literarnih citata, br. tabela, br. slika, br. grafika, br priloga.
FO
Naučna oblast: Matematika
NO
Naučna disciplina: Statistika
ND
Ključne reči: Faktorksa analiza, 2^k faktorijalni ogledi, analiza varijanse
PO
UDK:

Čuva se: U biblioteci Departmana za matematiku i informatiku
ČU
Važna napomena: Nema
VN
Izvod:
U radu se govori o faktorijalnim ogledima i rad sadrži četiri glavne celine. U prvom delu se govori o motivaciji za upotrebu faktorijalnih ogleda, osnovnim principima statističkog planiranja eksperimenta – ponavljanje, randomiziranje i pravljenje blokova. Takođe dati su osnovni pojmovi jednofaktorske i dvofaktorske analize varijanse.
U drugom delu rada date su matematičke osnove faktorijalnih ogleda, kroz oglede sa dva faktora. Takođe prikazano je višestruko poređenje, testiranje adekvatnosti modela i ocena parametara modela. U trećem delu obrađeni su 2^k faktorijalni ogledi - ogledi sa k faktora, od kojih svaki ima samo dva nivoa. U poslednjem delu, četvrtom, prikazana je primena faktorijalnih ogleda u poljoporivredi, na proizvodnji šećerne repe.
IZ
Datum prihvatanja teme od strane NN veća:
DP
Datum odbrane:
DO
Članovi komisije:
KO
Predsednik: dr Ljiljana Gajić, redovni profesor Prirodno-matematičkog fakulteta u Novom Sadu
Član: dr Zagorka Lozanov-Crvenković, redovni profesor Prirodno-matematičkog fakulteta u Novom Sadu, mentor
Član: dr Ivana Štajner-Papuga, vanredni profesor Prirodno-matematičkog fakulteta u Novom Sadu

University of Novi Sad
Fakulty of natural sciences and mathematics
Key words documentation

Accession number:
ANO
Identification number:
INO
Document type: Monograph type
DT
Type of record: Printed text
TR
Contents Code: Master's thesis
CC
Author: Vladan Todić
AU
Mentor: Prof. dr Zagorka Lozanov-Crvenković
MN
Title: Application of Factorial design
TI
Language of text: Serbian
LT
Language of abstract: English
LA
Country of publication: Serbia
CP
Locality of publication: Vojvodina
LP
Publication year: 2014
PY
Publisher: Author's reprint
PU
Publ. place: Novi Sad, Faculty of Natural Sciences and Mathematics, Trg Dositeja Obradovića 4
PP
Physical description: (4, 89, 11, 54, 3, 37, 0)
PD
Scientific field: Mathematics
SF
Scientific discipline: Statistics
SD
Key words: Factorial analysis, 2^k factorial design, analysis of variance
SKW
UC:
Holding data: In library of Department of Mathematics
HD

Note:
N
Abstract:
AB
This paper deals with factorial experiments with four main parts. The first part deals with the motivation for the use of factorial experiments, the basic principles of statistical experiment design - repetition, randomiziranje and making blocks. Also, there are the basic concepts of single-factor and two-factor analysis of variance.
The second part of the paper presents a mathematical basis factorial experiment, through experiments with two factors. Also shown is a multiple comparison, testing adequacy models and model parameters. In the third part analyzed the 2^k factorial experiments - experiments with k factors, each of which has only two levels. In the final section, the fourth shows the application of factorial experiments in agriculture, the sugar beet production.
Accepted by the Scientific Board on:
ASB
Defended:
De
Thesis defend board:
DB
President: Ph.D. Ljiljana Gajić, Full professor, Faculty of Natural Sciences and Mathematics, University of Novi Sad
Member: Ph.D. Zagorka Lozanov-Crvenković, Full professor, Faculty of Natural Sciences and Mathematics, University of Novi Sad, menthor
Member: Ph.D. Ivana Štajner-Papuga, Associate professor, Faculty of Natural Sciences and Mathematics, University of Novi Sad