Dizajn eksperimenta i ANOVA procedure

Radan, Matea

Sveučilište J. J. Strossmayera u Osijeku
Odjel za matematiku

Matea Radan

Dizajn eksperimenta i ANOVA procedure

Diplomski rad

Osijek, 2020.

Sveučilište J. J. Strossmayera u Osijeku
Odjel za matematiku

Matea Radan

Dizajn eksperimenta i ANOVA procedure

Diplomski rad

Mentor: prof. dr. sc. Mirta Benšić

Osijek, 2020.

Sadržaj

Uvod

U ovom radu upoznat ćemo se sa statističkim metodama dizajniranja eksperimenta, analizom varijance jednofaktorskog modela te slučajnim dizajnom punih blokova. Prije nego se posvetimo teorijskom i praktičnom pojašnjavanju ovih modela, proučit ćemo smjernice za dizajniranje eksperimenata. Metode dizajna eksperimenta primjenjive su u mnogim područjima te su važan alat za poboljšavanje proizvodnih procesa. Njihova upotreba u proizvodnim procesima može rezultirati proizvodima koji su jednostavniji za proizvodnju, a koji poboljšavaju izvedbu i pouzdanost samog proizvoda na terenu. Osim toga kraće vrijeme razvoja proizvoda te smanjenje troškova proizvodnje još su samo neke od prednosti do kojih statističko dizajniranje dovodi.
Za početak ćemo kratko objasniti kako se kroz povijest razvijao statistički dizajn eksperimenta.

Povijest statističkog dizajna eksperimenta

Moderan razvoj statističkog dizajna eksperimenta možemo podijeliti na četiri razdoblja. Prvo razdoblje nazivamo poljoprivrednom erom i započelo je radom Ronald A. Fishera u 1920-tim i početkom 1930-tih godina. U tom periodu Fisher je radio na poljoprivrednom institutu (nekad: Rothmasted Agricultural Experimental Station) u Rothamstedu blizu Londona. Kroz svoje poljoprivredne eksperimnte zapazio je da način na koji su podaci prikupljeni često onemogućuje analizu podataka. U suradnji s mnogim istraživačima i znanstvenicima iz raznih područja razvio je prve spoznaje koje su dovele do tri osnovne metode dizajna eksperimenta: randomizacija, replikacija i blokiranje. Fisher je kontinuirano uvodio statističke principe u dizajniranje eksperimentalnih istraživanja, uključujući faktorijalni dizajn i analizu varijance. Njegove dvije knjige (vidi [6] i [7]) su imale veliki utjecaj na upotrebu statistike u poljoprivredi i drugim znanostima. Druga ili industrijska era učvrtila je upotrebu statističkog dizajniranja eksperimenata u industriji i znanosti. Obilježena je razvojem metodologije odzivne površine (eng. response surface methodology, nadalje u tekstu: "RSM metodologija") koju su opisali George E.P. Box i K.B. Wilson. Uočili su da se mnogi industrijski eksperimenti razlikuju od poljoprivrednih u dvije najbitnije stavke:

1. ovisna varijabla se uglavnom može odmah primijetiti u industriji

2. u industriji se brzo u malom broju ponavljanja mogu uočiti bitne informacije za ponavljanje sljedećeg eksperimenta.

Box ove dvije karakteristike industijskih eksperimenata naziva neposrednost i posljedičnost. U idućih 30 godina, RSM metodologija i druge slične metode dizajniranja proširile su se na kemijska i industrijska istraživanja i razvojne procese, no ipak ne značajno. Razlog tome je manjak znanja o osnovnim statističkim koncepatima i metodama kod inžinjera i znanstvenika, kao i nedostatak statističkih softvera koji bi podržali primjenu dizajniranih eksperimenata. Krajem 1970-tih pojavljuje se sve veći interes za boljom kvalitetom proizvoda čime počinje treća era statističkog dizajna. Rad japanskog inžinjera i statističara Genichi Taguchija izazvao je širenje interesa za korištenjem dizajniranih eksperimenata. Taguchi se zalagao za korištenje dizajniranih eksperimenata kojima bi:

1. proces postao neosjetljiv na uvjete okoline ili bilo koje druge faktore koje je teško kontrolirati

2. finalni proizvodi bili neosjetljivi na promjene u komponentama

3. pronašao parametre je čija srednja vrijednost na razini željene vrijednosti, a istodobno smanjenih fluktuacija oko srednje vrijednosti.

Metode koje je Taguchi predložio (podijeljeni faktorizacijski dizajn i ostale) naišle su na mnogo rasprava i neslaganja. Krajem osamdesetih, stručnim pregledom zaključeno je da su Tuguchijevi inžinjerski ciljevi i koncepti bili osnovani, ali je problem uočen pri eksperimentalnom procesu i statističkoj obradi podataka. Iz osporavanja Taguchijeva rada proizašlo je nekoliko pozitivnih ishoda. Dizajnirani eksperimenti privukli su interes raznih industrija, pretežito tehničkih, poput automobilske i zrakoplovne proizvodnje te elektroničke industrije. Time je počela i četvrta era dizajniranih eksperimenata. Ovo doba obilježeno je novim zanimanjem industrijskog svijeta za statistički dizajn i razvojem mnogih novih korisnih pristupa eksperimentalnim problemima. Pronađene su alternative Taguchijevim tehničkim metodama koje omogućuju da se njegove ideje učinkovito provode u praksi. U posljednje vrijeme došlo je do značajnih primjena statistički dizajniranih eksperimenata u mnogim područjima uključujući uslužni sektor, financije, marketing, ekonomiju, itd.

Poglavlje 1

1 Smjernice za dizajniranje eksperimenata

Upotreba statističkog pristupa dizajniranju eksperimenta zahtijeva da sve osobe uključene u proces eksperimentiranja imaju jasnu ideju što se točno proučava, o tome na koji način se prikupljaju podaci te da razumiju kako bi se podaci trebali analizirati. U tablici 1.1 dan je pregled koraka koje je poželjno slijediti kako bi eksperiment dizajnirali koristeći statističke pristupe. Koraci 2 i 3 često se izvode istovremeno ili obrnutim redoslijedom. Svi koraci su detaljnije opisani u nastavku.

1. Prepoznavanje i definiranje problema
2. Odabir ovisne varijable
3. Izbor varijabli, raspona i granica
4. Izbor dizajna eksperimenta
5. Izvođenje eksperimenta
6. Statistička obrada podataka
7. Zaključak i preporuke

Tablica 1.1: Smjernice za dizajniranje eksperimenta

Korak 1. Prepoznavanje i definiranje problema

Uočavanje problema u okolini se možda čini previše jednostavno, ali u praksi često nije tako. Shvatiti da postoji problem nerijetko iziskuje posebnu posvećenost, a isto tako nekada nije lako uočiti da se rješenje tog problema može pronaći eksperimentirajući. Kada se problem zapazi potrebno je razviti jasne i općeprihvaćene činjenice o samom problemu, a zatim razraditi sve ideje o ciljevima eksperimenta. Vrlo je važno zatražiti informacije od svih strana uključenih u proces kao što su npr. inžinjeri, testeri kvalitete, proizvođači, marketing, kupci i prodajno osoblje. Upravo zato preporučuje se oformiti tim koji će uključivati osobe iz svakog područja. Tim zatim priprema popis pitanja i/ili problema koji se trebaju riješiti eksperimentom. Detaljno opisivanje problema doprinosi razumijevanju fenomena koji proučavamo i konačnom rješavanju problema. Postaviti prava pitanja imajući na umu ciljeve eksperimenta

jedan je od glavnih zadataka u predeksperimentalnom radu. U nastavku navodimo nekoliko primjera kako bi propitivanje procesa trebalo izgledati; npr. "Je li ovo novi proces?", "Nadograđujemo li ovim eksperimentom već postojeći proces?", "Radi li sustav na isti način kao nekada?", "Što ako u sustav ubacimo nove materijale?", itd.
Važan dio formuliranja problema je i prepoznavanje da jedan veliki sveobuhvatni eksperiment možda neće odgovoriti na sva ključna pitanja. Ako se eksperiment postavi u prevelikom obujmu, to bi značilo da tim zadužen za eksperiment mora moći odgovoriti na mnoštvo raznih pitanja, a ukoliko su u krivu rezultat može biti razočaravajući. To bi istovremeno značilo da su potrošeni resursi, materijali i vrijeme, a da nismo dobili zadovoljavajuće odgovore na istraživačka pitanja. U ovakvim slučajevima od iznimne je važnosti raspodijeliti proces na niz manjih eksperimenata, od kojih je svaki određen specifičnim ciljevima.

Korak 2. Odabir ovisne varijable

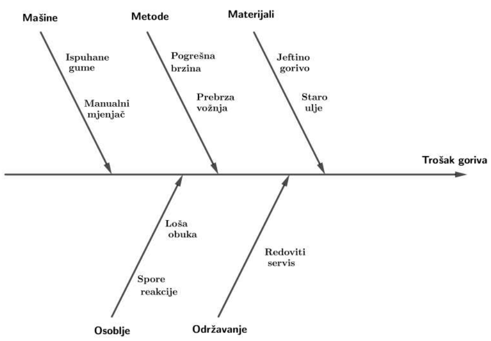
Pri odabiru ovisne varijable osoba zadužena za eksperiment bi doista trebala biti sigurna da ova varijabla daje korisne informacije o proučavanom procesu. Najčešće će najbolji odabir za ovisnu varijablu biti prosječno ili standardno odstupanje (ili oboje). Mjerna pogreška i adekvatnost mjernih instrumenata su isto tako vrlo bitni faktori koji imaju utjecaj na ovisnu varijablu. Ako su mjerni instrumenti ograničeni i/ili ako mjeritelj nije precizan, eksperimenetom će biti uočeni samo veliki učinci ili će biti potrebno provesti dodatna mjerenja. U nekim slučajevima može se izmjeriti svaku eksperimentalnu jedinicu više puta te za daljnje postupanje upotrijebiti prosjek svih mjerenja. Kako bi se smanjio utjecaj na ovisnu varijablu važno je identificirati probleme vezane uz područje interesa i proučiti na koje se načine provode mjerenja prije provođenja eksperimenta. Ponekad se dizajniranje eksperimenata koristi i za proučavanje i poboljšavanje mjernih sustava.

Korak 3. Izbor varijabli, raspona i granica

Pri određivanju varijabli koje imaju utjecaj na proces ili sustav, eksperimentatori su primjetili da se te varijable mogu podijeliti u dvije skupine *potencijalne varijable dizajna* i *faktore smetnji*. Potencijalne varijable dizajna su one koje želimo varirati tijekom eksperimenta. Kako se tijekom eksperimenta najčešće pojavi puno potencijalnih varijabli dizajna, dobro je napraviti i detaljniju podjelu. Jedna od klasifikacija koja je često u upotrebi je podjela na bazne varijable dizajna, fiksirane varijable te na varijable kojima su dopuštena odstupanja. Bazne varijable dizajna su one koje su odabrane za proučavnje u eksperimentu. Fiksirane varijable su one koje mogu imati utjecaj na rezultat, ali za potrebe ovog eksperimenta nisu zanimljive pa će se držati na nekoj konstantnoj razini. Pretpostavljamo da su učinci fiksiranih varijabli i varijabli kojima se dopušta odstupanje relativno mali.
S druge strane, iako nas faktori smetnje neće zanimati u okviru rezultata eksperimenta, oni mogu imati veliki utjecaj na rezultat pa ih je potrebno uzeti u obzir. Često ih klasificiramo kao kontrolirajuće, nekontrolirajuće i kao faktore šuma. Kontrolirajući faktori smetnje su oni čiju vrijednost ispitivač može postaviti i samim time se mogu uspoređivati. Za primjer kontolirajućih faktora smetnje možemo navesti dan u tjednu kada će eksperiment biti proveden. Princip blokova je često koristan za suočavanje sa kontrolirajućim faktorima smetnje, te je isti detaljnije pojašnjen u Poglavlju [3]. Za procjenu utjecaja nekontorlirajućih faktora smjetnje koji se mogu izmjeriti koristimo analizu varijance (detaljnije u Poglavlju [2]). Za primjer nekontrolirajućih faktora smetnje možemo navesti količinu oborina u nekom mjesecu, koja se ne može kontrolirati, ali se može izmjeriti. Faktorom šuma nazivaju se fakori koji su

po prirodi nekontrolirajući i variraju, ali se za potrebe eksperimenta isti mogu kontrolirati. U ovim slučajevima potrebno je pronaći postavke kontolirajućih faktora koji umanjuju varijabilnost faktora šuma. Takvi procesi, koji su neosjetljivi na šumove, nazivaju se robusni procesi (eng. robust desgin problem).

Nakon što su odabrane varijable dizajna potrebno je odabrati granice u kojima će se te varijabe kretati. Također, potrebno je razmisliti o načinu mjerenja odabranih varijabi. Za to je potebno poznavati proces i imati adekvatna znanja o procesu koja su najčešće kombinacija praktičnih znanja i teorijskog razumijevanja. Važno je proučiti sve bitne parametre na koje ne smijemo utjecati s iskustvom iz prethodnih studija, naručito u ranim fazama eksperimenta. Kada je cilj eksperimenta provjera ili karakterizacija procesa najbolje je zadržati mali broj varijabli. Područje interesa je važan čimbenik te bi u početku procesa trebalo biti relativno veliko, odnosno raspon iz kojeg se biraju varijable bi trebao biti širok. Kako proces napreduje te kako se prikupljaju informacije o važnim varijablama tako se simultano područje interesa sužava. Koristan alat za organiranje nekih prikupljenih podataka u predeksprimentalnom planiranju može biti dijagram uzoraka i posljedica, poznat još kao i dijagram riblje kosti. Ovisna varijabla se crta duž kralježnice dijagrama, a potencijalni uzroci ili varijable dizajna organizirani su u nizu rebara. Na Slici 1.1 prikazan je dijagram uzroka i posljedica u eksperimentu koji je za rezultat imao proučiti potrošnju goriva automobila.



Slika 1.1: Dijagram uzroka i posljedica

Dijagram uzroka i posljedica koristi tradiocinalne uzroke mjerena, materijale, ljude, okoliš, metode, strojeve i slično za organiziranje informacija i potencijalnih varijabli dizajna. Neki će pojedinačni uzroci biti uključeni u eksperiment (poput vrste mjenjača) kao varijable, neki kao nekontrolirani faktori i/ili šumovi, dok će drugi predstavljati područja za posebne studije. Uobičajno je da se sastavlja nekoliko različitih dijagrama uzroka i posljedica tijekom predeksperimentalnog planiranja. Većina uspjeha ovisi upravo o predeksperimentalnom planiranju koje obuhvaća korake od 1 do 3 te je zbog toga u njih iznimno bitno uložiti poseban trud.

Korak 4. Izbor dizajna eksperimenta

Ako je predeksperimentalno planiranje izvedeno kvalitetno, ovaj korak bi trebao biti relativno jednostavan. Prema [5], dizajn eksperimenta je statistička tehnika usvojena za izvođenje eksperimentalnih studija usredotočenih na poboljšanje proizvoda i procesa. Izbor dizajna eksperimenta ovisi o ciljevima eksperimenta i broju faktora koji se istražuju. Prema [12], tipovi dizajna su navedeni sukladno eksperimentalnom cilju, a koji može biti usporedba faktora, detekcija najbitnijih faktora te procjena interakcija i učinaka. Ukoliko imamo jedan ili više faktora za koje provodimo istraživanje, a pri tome je glavni cilj eksperimenta zaključivanje o značajnosti jednog faktora tj. postoji li značajna promjena u rezultatu ovisno o različitim razinama tog faktora, koristimo usporedne dizajne. U nastavku rada obradit ćemo upravo takva dva modela, jednofaktorski model analize varijance i model slučajnog dizajna punih blokova. Ukoliko je cilj eksperimenta odabrati nekoliko važnih glavnih učinaka od ostalih manje bitnih, koristimo dizajn glavnih efekata. Puni ili djelomični faktorski modeli su primjer modela dizajna glavnih efekata. Metoda površinske reakcije je dizajn koji nam omogućuje procjenu interakcija i kvadratnih učinaka. Za više o metodi površinske reakcije pogledati [9, str. 417-480]. Postoje i statistički paketi koji sadrže alate korisne za ovaj korak eksperimentalnog dizajna. Eksperimentator unese podatke o broju varijabli, rasponu i razini, a program ponudi izbor dizajna za upotrebu i razmatranje. Pri odabiru dizajna važno je imati na umu ciljeve eksperimenta. U nekim eksperimentima pokazalo se da su različite razine varijabli rezultirale značajnim razlikama u rezultatu. Sukladno tome, zanima nas identificiranje čimbenika koji uzrokuju ovu razliku i procjena veličine promjene u rezultatu.

Kao što smo naveli, u nastavku rada obradit ćemo dvije metode dizajna eksperimenta, analizu varijance za jednofaktorski model te analizu varijance za slučajni dizajn punih blokova. U primjeru 3.2 možemo uočiti do kakvih razlika dolazi prilikom upotrebe različitih dizajna.

Korak 5. Izvođenje eksperimenta

Prilikom izvođenja eksperimenta bitno je pažljivo nadgledati proces kako bi bili sigurni da sve teče prema planu. Pogreške u ovom koraku mogu osporiti eksperiment, a da do toga ne bi došlo bitno je pravovremeno planiranje. U složenim istraživanjima ili proizvodnjama lako se dogodi loša procjena logičkih i planskih aspekata izvođenja dizajniranog eksperimenta. Coleman i Montgomery [4] preporučuju izvođenje nekoliko probnih ponavljanja prije početka eksperimentiranja. Ova probna ponavljanja nam pružaju: informacije o kvaliteti eksperimentalnih materijala, okvirnu ideju o eksperimentalnoj pogrešci, mogućnost provjere mjernog sustava te šansu da preispitamo eksperimentalne tehnike i odluke donesene u prethodnim koracima 1 do 4.

Korak 6. Statistička obrada podataka

Upotreba statističkih metoda za obradu podataka je ključna kako bi zaključci i rezultati bili objektivni. Ukoliko je eksperiment pravilno osmišljen te izveden u skladu sa dizajnom, potrebne statističke metode nisu složene. Danas postoji mnoštvo softverskih paketa koje je moguće koristiti u analizi podataka, a mnogi programi koji se koriste za odabir dizajna pružaju i direktnu poveznicu sa statističkom analizom. Jednostavni grafovi i slične grafičke metode imaju važnu ulogu u statističkoj analizi i interpretaciji podataka. U najvećem broju slučajeva vrlo korisno je prikazati rezultate eksperimenta u empirijskom modelu, odnosno u obliku jednadžbe koja predstavlja vezu između ovisne varijable i varijabli dizajna. Preostali dio analize i provjera adekvatnosti odabranog modela su također bitni dijelovi statističke obrade podataka. Objektivnost u donošenju odluke je jedna od glavnih prednosti statističkih metoda. U kombinaciji sa kvalitetnim inžinjerskim znanjem te znanjem o procesu statističke metode dovode do ispravnih i zdravih zaključaka.

Korak 7. Zaključak eksperimenta i preporuke

Nakon statističke obrade podataka, eksperimentator mora donijeti zaključke o rezultatima i preporučiti djelovanje u skladu sa rezultatima. Najkorisniji alat za prezentiranje donesnih zaključaka ostalim sudionicima te zainteresiranima su grafičke metode. Jedna od preporuka koja se često donosi je provjera zaključaka eksperimenata u praksi. Time se zaključci i potvrđuju i sa sigurnijim pristupom se može ići u daljnje procese. Tijekom zaključivanja i donošenja nekih odluka važno je kontinuirano imati na umu da je eksperimentiranje važan dio procesa učenja, u kojemu formuliramo hipoteze o nekom sustavu ili procesu, izvodimo eksperimente kako bismo istražili te hipoteze, a zatim na temelju rezultata formuliramo nove hipoteze, koje je opet potrebno testirati. Zapravo je eksperimentiranje iterativan proces. Već smo spomenuli da je pogrešno dizajnirati samo jedan veliki proces na početku istraživačke studije. Uspješan eksperiment zahtijeva znanje o svim važnim čimbenicima, rasponima kretanja, odgovarajućem broju razina koje se mogu koristiti te odgovarajućim mjernim jedinicama. Na početku ne posjedujemo znanje o svemu navedenom, ali o tome, kao i o drugim utjecajima, učimo tijekom cijelog eksperimentalnog programa. Kako eksperiment napreduje tako odbacujemo neke ulazne varijable, neke ulazne varijable zamijenjujemo s drugima, mijenjamo razine i raspone kretanja. Upravo iz razloga kontinuranog učenja, u prvi eksperimentalni pristup nije preporučljivo uložiti više od dvadeset i pet posto raspoloživih novčanih sredstava. Takav pristup omogućit će dovoljnu količinu resursa za izvođenje potrebnih ponavljanja i novih manjih pokusa te u konačnici dovesti do željenog cilja.

Poglavlje 2

2 Analiza varijance

U ovom poglavlju bavit ćemo se analizom varijance ili skraćeno ANOVA (od eng. Analysis of Variance). Razvio ju je R. A. Fisher kao matematički model i tehniku za uspoređivanje i istraživanje bioloških modela. Danas se ANOVA metoda koristi i u mnogim drugim područjima te je izrazito popularna. Za razumijevanje daljnjeg sadržaja rada potrebno je određeno predznanje teorije vjerojatnosti i statistike, vidi [2] i [10]. Mi ćemo se u okviru ovog rada fokusirati na dizajn i objašnjenje jednofaktorskog eksperimenta sa a razina faktora. Također, pretpostavljamo da je eksperiment proveden nasumično.

2.1 Motivacija

Za lakše razumijevanje problematike promorit ćemo primjer potrošnje goriva.

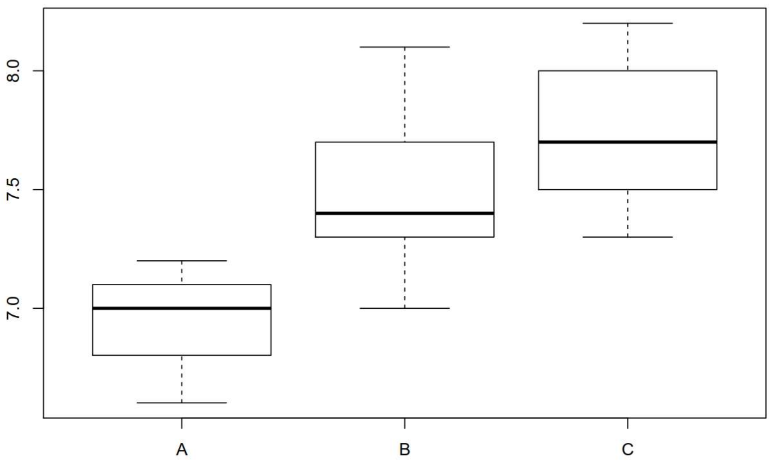
Primjer 2.1. Tri tvornice automobila A, B, C proizvode automobile približno jednake snage motora. Korisnik želi provjeri informaciju o potrošnji goriva, odnosno ovisi li potrošnja goriva automobila o tvornici koja ga proizvodi. Kako ćemo dizajnirati eksperiment?

Jedna od ideja za rješavanje ovog problema je promatranje nekoliko vrsti automobila istih specifikacija koje proizvode različite tvornice. Znamo da potrošnja goriva ovisi i o mnogo drugih čimbenika poput vrste ceste, vozača, vremenskih uvjeta i slično. Da bi testirali hipotezu o utjecaju tvornice na potrošnju goriva potrebno je sve ostale čimbenike eliminirati. Za tu svrhu isti vozač će voziti sve automobile, na istoj cesti, u istim vremenskim uvjetima i pod svim ostalim kontroliranim uvjetima.
Pretpostavimo da imamo $n = 5$ automobila tvornice A, B i C. Rezultati obavljenog eksperimenta prikazani su u tablici 2.1.

Tvornica	Potrošnja goriva					Ukupno	Aritmetička sredina
A	7,2	6,6	6,8	7,1	7,0	34,7	6,94
B	8,1	7,7	7,0	7,3	7,4	37,5	7,50
C	7,3	8,2	7,5	8,0	7,7	38,7	7,74

Tablica 2.1: Rezultati eksperimenta potrošnje goriva automobila

Postavlja se pitanje možemo li na temelju podataka u tablici 2.1 pretpostaviti da postoje značajne razlike u potrošnji goriva različitih tvornica. Radi lakšeg uočavanja razlika promotrimo podatke grafički prikazane na Slici 2.1.



Slika 2.1: Kutijasti dijagram potrošnje goriva automobila iz tvornica A, B, C

Promatrajući graf mogli bismo pretpostaviti da automobili tvornice C troše najviše goriva, no želimo objektivnost u analizi pa je potrebno izgraditi model. Želimo testirati razlike u očekivanju za sve tri tvornice. Pravilan pristup ovome problemu je analiza varijance.

2.2 Analiza varijance

Pretpostavimo da imamo a razina (tretmana) ili vrijednosti faktora koje želimo usporediti. Uz to pretpostavimo da za svaku razinu faktora imamo $n \geq 2$ mjerenja i da je za i-ti faktor, $i = 1, \ldots, a$ niz $y_{i1}, \ldots, y_{in}$ dobiven mjerenjem slučajne varijable y_i [1]. Podaci su prikazani u tablici 2.2. Primjetimo da je tablica 2.2 generalizirani zapis tablice 2.1 iz primjera 2.1. Model kojim možemo opisati podatke iz tablice glasi

$$y_{ij} = \mu_i + \epsilon_{ij}, \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases}, \tag{2.1}$$

gdje je y_{ij} ij-ta opservacija, μ_i je aritmetička sredina i-te razine faktora i ϵ_{ij} je slučajna greška. Slučajna greška uključuje sve druge izvore varijabilnosti u eksperimentu, poput varijabilnosti mjerenja, varijabilnosti koja proizlazi iz nekontrolirajućih faktora, razlika između

[1] U ovom radu slučajne varijable i njihove realizacije označavamo istom oznakom podrazumijevajući da se iz konteksta prepoznaje radi li se o slučajnoj varijabli ili realizaciji.

Tretman	Rezultati mjerenja				Ukupno	Aritmetička sredina
1	y_{11}	y_{12}	$\dots$	y_{1n}	y_1	$\overline{y}_1$
2	y_{21}	y_{22}	$\dots$	y_{2n}	y_2	$\overline{y}_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	$\dots$	y_{an}	y_a	$\overline{y}_a$

Tablica 2.2: Podaci za jednofaktorski eksperiment

eksperimentalnih jedinica na koje se primjenjuju razine ili tretmani faktora (npr. materijal). Slučajna greška uključuje i opće šumove u procesu, poput varijabilnosti vremena i utjecaja varijabli okoline. Pretpostavljat ćemo da je očekivanje slučajne greške jednako nuli, odnosno da vrijedi $E(\epsilon_{ij}) = 0$. Sada vrijedi $E(y_{ij}) = \mu_i$. Jednadžba (2.1) naziva se model očekivanja. Da bi zapisali model na drugačiji način definirat ćemo

$$\mu_i = \mu + \tau_i$$

pa sada jednadžba (2.1) postaje

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases}. \tag{2.2}$$

U ovom zapisu modela, μ je zajednički parametar svim razinama (tretmanima) faktora i naziva se ukupna aritmetička sredina, dok je parametar τ_i jedinstven za svaku razinu (tretman) faktora i i naziva se efekt tretmana. Jednadžba 2.2 naziva se modelom efekata.

Oba zapisa modela (2.1) i (2.2) su linerani statistički modeli, odnosno ovisna varijabla y_{ij} je linerna funkcija parametara. Iako su oba modela korisna, model efekata se češće koristi u literaturi vezanoj za problematiku dizajniranja eksperimenta. Razlog tome je intuitivnost u pristupu kroz konstantu μ, a dok τ_i posebno predstavlja devijaciju od konstantne vrijednosti μ za svaki tretman i.

Jednadžba (2.2) (ili (2.1)) proučava razlike u samo jednom faktoru pa je po tome i dobila ime **jednodimenzionalni** ili **jednofaktorski model analize varijance** (ANOVA). Nadalje ćemo pretpostaviti da se eksperiment izvodio slučajnim redoslijedom kako bi eksperimentalne jedinice tj. tretmani bili što ujednačeniji. Eksperimentalni dizajn je potpuno slučajan (eng.completely random design). Cilj ANOVE je testirati odgovarajuće hipoteze o očekivanjima varijable na tretman. Za potrebe testiranja hipoteza pretpostavljamo da su slučajne greške modela normalne nezavisne slučajne varijable sa očekivanjem nula i varijancom σ^2. Varijanca σ^2 je konstantna za sve tretmane faktora. Stoga možemo zapisati

$$\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2),$$

iz čega slijedi da su sve slučajne varijable y_{ij} nezavisne te da vrijedi:

$$y_{ij} \sim \mathcal{N}(\mu + \tau_i, \sigma^2).$$

Jednadžba (2.2) opisuje dvije različite situacije promatrajući tretmane faktora. U prvoj situaciji eksperimentator je odabrao a tretmana koje želi testirati. U ovoj situaciji želimo testirati hipoteze o očekivanjima tretmana te će se naš zaključak primijeniti samo na tretmane koje smo uzeli u obzir u analizi (kao što je prikazano u primjeru 2.1). Osim zaključka i testiranja hipoteza, možemo procijeniti parametre modela μ, τ_i i σ^2. Ovakav slučaj naziva se modelom fiksih efekata. S druge strane, a tretmana mogu biti izabrana slučajno iz populacije tretmana. U ovoj situaciji bismo trebali moći prošiti zaključke na sve jedinke populacije, bez obzira bile one uključene u analizu ili ne. Tada su τ_i slučajne varijable, a znanje o svakoj pojedinačnoj vrijednosti τ_i je relativno beskorisno. Umjesto toga, testiramo hipoteze o varijabilnosti τ_i i pokušavamo procijeniti tu varijabilnost. Ovakav model naziva se model slučajnih efekata ili model komponenti varijance. Za više o ovom modelu pogledati [9, str. 486 - 505], dok je u nastavku detaljnije pojašnjen model fiksnih efekata.

2.3 Model fiksih efekata

Podsjetimo da y_i predstavlja ukupnu sumu opservacija i-tog tretmana faktora, dok $\overline{y}_i$ predstavlja aritmetičku sredinu svih opservacija i-tog tretmana. Neka y predstavlja ukupnu sumu svih n opervacija svih a tretmana, a neka je $\overline{y}$ aritmetička sredina svih opservacija.
Zapišimo sada to matematički,

$$\begin{aligned} y_i &= \sum_{j=1}^{n} y_{ij} \qquad \overline{y}_i = \frac{y_i}{n} \qquad i = 1, 2, \ldots, a \\ y &= \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij} \qquad \overline{y} = \frac{y}{N} \end{aligned} \tag{2.3}$$

gdje je $N = an$ ukupan broj svih observacija.
Naš zadatak je testirati jednakost očekivanja a tretmana, a koje iznosi $E(y_i) = \mu + \tau i = \mu_i$, $i = 1, 2, \ldots, a$. Definirajmo hipoteze:

$$\begin{aligned} H_0 :& \mu_1 = \mu_2 = \cdots = \mu_a \\ H_1 :& \mu_i \neq \mu_k, \qquad \text{za neki par (i,k)} \end{aligned} \tag{2.4}$$

gdje su $i, k = 1, 2, \ldots, a$. U modelu efekata, očekivanje i-tog tretmana μ_i rastavimo na dvije komponente μ i τ_i tj. $\mu_i = \mu + \tau i$. O μ obično razmišljamo kao o ukupnoj aritmetičkoj sredini svih očekivanja pa slijedi:

$$\frac{\sum_{i=1}^{a} \mu_i}{a} = \mu$$

iz čega dobivamo sljedeće:

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

Prema tome, možemo reći da se tretmani faktora mogu shvatiti kao devijacije oko ukupnog očekivanja. Ekvivalentan način zapisa hipoteza 2.4 odnosi se na efekte tretmana τ_i:

$$\begin{aligned} H_0 :& \tau_1 = \tau_2 = \cdots = \tau_a \\ H_1 :& \tau_i \neq 0, \qquad \text{za barem jedan i.} \end{aligned} \tag{2.5}$$

Govorimo o testiranju jednakosti očekivanja tretmana ili o testiranju nulte vrijednosti efekata tretmana τ_i. Analiza varijance je pravi odabir za testiranje ovih hipoteza.

Ime analiza varijance je nastalo zbog rastava ukupne varijance na njezine komponente. Ukupna suma kvadrata

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2$$

se koristi za mjerenje ukupne varijabilnsoti podataka. Ako ukupnu sumu kvadrata SS_T podijelimo s brojem stupnjeva slobode, što je u ovom slučaju jednako $an-1 = N-1$, imamo uzoračku varijancu svih podataka. Primijetimo da ukupnu sumu kvadrata SS_T možemo zapisati kao

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n}[(\overline{y}_i - \overline{y}) + (y_{ij} - \overline{y}_i)]^2 \tag{2.6}$$

ili

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2 =& n\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2 \\ &+2\sum_{i=1}^{a}\sum_{j=1}^{n}(\overline{y}_i - \overline{y})(y_{ij} - \overline{y}_i)\end{aligned}$$

Primijetimo da je unakrasni član u gornjoj jednakosti jednak nuli. Razlog tome je

$$\sum_{j=1}^{n}(y_{ij} - \overline{y}_i) = y_i - n\overline{y}_i = y_i - n\left(\frac{y_i}{n}\right) = y_i - y_i = 0.$$

Sada imamo,

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2 = n\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2, \tag{2.7}$$

što nazivamo fundamentalni zapis analize varijnce. Govori nam da se ukupna varijabilnost podataka, mjerena ukupnom sumom kvadrata, može rastaviti na sumu kvadrata razlika prosjeka tretmana i ukupnog prosjeka te na ukupnu sumu kvadrata razlika opservacija tretmana i prosjeka tretmana. Jednadžbu (2.7) možemo zapisati i kao

$$SS_T = SS_{Tretman} + SS_E,$$

gdje je $SS_{Tretman}$ suma kvadrata tretmana, a SS_E suma kvadrata grešaka. Ukupno imamo $an = N$ opservacija pa SS_T ima $N-1$ stupnjeva slobode i imamo a tretmana faktora pa $SS_{Tretman}$ ima $a-1$ stupnjeva slobode. Imamo n ponavljanja svakog tretmana, što znači da $n-1$ stupanj slobode procjenjuje eksperimentalnu pogrešku. S obzirom da imamo a tretmana, imamo $a(n-1) = an - a = N - a$ stupnjeva slobode greške.
Korisno je detaljnije pogledati dva izraza na desnoj strani jednakosti fundamentalnog zapisa ANOVE (2.7). Promotrimo sumu kvadrata pogrešaka,

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2 = \sum_{i=1}^{a}\left[\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2\right].$$

Možemo primijetiti da ako izraz unutar uglatih zagrada podijelimo sa $n-1$, dobijemo uzoračku varijancu S_i^2 $i-tog$ tretmana

$$S_i^2 = \frac{\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2}{n-1}, \qquad i = 1, 2, \ldots, a.$$

Uzoračku varijancu možemo kombinirati kako bi dobili pojedinačnu procjenu varijance ukupne populacije kako slijedi:

$$\frac{(n-1)S_1^2+(n-1)S_2^2+\cdots+(n-1)S_a^2}{(n-1)+(n-1)+\cdots+(n-1)}=\frac{\sum_{i=1}^{a}\left[\sum_{j=1}^{n}(y_{ij}-\overline{y}_i)^2\right]}{\sum_{i=1}^{a}(n-1)}=\frac{SS_E}{(N-a)}.$$

Pa tako možemo reći da je $SS_E/(N-a)$ skupna procjena zajedničke varijance unutar svakog od a tretmana.
Slično, ukoliko ne postoje značajne razlike između očekivanja tretmana, možemo procijeniti σ^2 pomoću varijacije prosječnih tretmana od ukupnog prosjeka. Posebno,

$$\frac{SS_{Tretman}}{a-1}=\frac{n\sum_{i=1}^{a}(\overline{y}_i-\overline{y})^2}{a-1}$$

je procjena za σ^2 ukoliko su očekivanja tretmana jednaka.
Izraz

$$\frac{\sum_{i=1}^{a}(\overline{y}_i-\overline{y})^2}{(a-1)} \qquad \text{procjenjuje} \qquad \frac{\sigma^2}{n}$$

pa

$$\frac{n\sum_{i=1}^{a}(\overline{y}_i-\overline{y})^2}{(a-1)} \qquad \text{procjenjuje} \qquad \sigma^2,$$

ukoliko nema razlika u očekivanjima tretmana. Iz jednadžbe (2.7) smo dobili dvije procjene za σ^2, jednu na temelju varijabilnosti unutar tretmana, a drugu na temelju varijabilnosti između tretmana. Ukoliko ne postoje razlike u očekivanjima tretmana, ove dvije procjene bi trebale bit približno slične. Pokažimo to sada formalno. **Srednje kvadratno odstupanje** je dano formulama

$$MSE_{Tretman}=\frac{SS_{Tretman}}{a-1}$$

i

$$MS_E=\frac{SS_E}{N-a}.$$

Pogledajmo sada vrijednosti očekivanja tih srednje kvadratnih odstupanja.

$$\begin{aligned}
E(MS_E)=&E\left(\frac{SS_E}{N-a}\right)=\frac{1}{N-a}E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\overline{y}_i)^2\right]\\
=&\frac{1}{N-a}E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2-2y_{ij}\overline{y}_i+\overline{y}_i^2)\right]\\
=&\frac{1}{N-a}E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}y_{ij}^2-2n\sum_{i=1}^{a}\overline{y}_i^2+n\sum_{i=1}^{a}\overline{y}_i^2\right]\\
=&\frac{1}{N-a}E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}y_{ij}^2-\frac{1}{n}\sum_{i=1}^{a}y_i^2\right]
\end{aligned}$$

Uvrštavanjem jednadžbe (2.2) i $y_i = \sum_{j=1}^{n} y_{ij}$ dobivamo

$$E(MS_E) = \frac{1}{N-a} E\left[\sum_{i=1}^{a} \sum_{j=1}^{n} (\mu + \tau_i + \epsilon_{ij})^2 - \frac{1}{n} \sum_{i=1}^{a} \left(\sum_{j=1}^{n} (\mu + \tau_i + \epsilon_{ij}) \right)^2 \right].$$

Kvadriranjem i primjenjivanjem linearnosti očekivanja te posljedično zbog $E(\epsilon_{ij}) = 0$ i činjenice da svi dvočlani faktori nastali kvadriranjem sadrže ϵ_{ij} slijedi

$$E(MS_E) = \frac{1}{N-a}\left[N\mu^2 + n\sum_{i=1}^{a} \tau_i^2 + N\sigma^2 - N\mu^2 - n\sum_{i=}^{a} \tau_i^2 - n\sigma^2 \right]$$

i završno

$$E(MS_E) = \sigma^2.$$

Analogno ovom postupku, dobivamo

$$E(MS_{Tretman}) = \sigma^2 + \frac{n\sum_{i=1}^{a} \tau_i^2}{a-1}.$$

Ovime smo dokazali da $MS_E = SS_E/(N-a)$ procjenjuje σ^2 te da, ukoliko nema razlika u očekivanjima tretmana (što implicira da je $\tau_i = 0$), $MS_{Tretman} = SS_{Tretman}/(a-1)$ također procjenjuje σ^2. Primjetimo još, ukoliko se očekivanja tretmana razlikuju, očekivana vrijednost srednje kvadratnog odstupanja tretmana je veća od σ^2.
Jasno je sada da hipoteze o jednakosti očekivanja tretmana možemo testirati kroz usporedbu $MS_{Tretman}$ i MS_E.

Pogledajmo sada kako možemo formalno testirati hipoteze o različitosti očekivanja tretmana ($H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$ ili ekvivalentno $H_0 = \tau_1 = \tau_2 = \cdots = \tau_a$). S obzirom da smo pretpostavili da su greške ϵ_{ij} nezavisne normalno distribuirane s očekivanjem nula i varijancom σ^2, observacije y_{ij} su nezavisne normalno distribuirane sa očekivanjem $\mu + \tau_i$ i varijancom σ^2. Dakle, SS_T je suma kvadrata nezavisnih normalno distribuiranih varijabli, pa možemo dokazati da je

$$\frac{SS_T}{\sigma^2} \sim \chi^2(N-1).$$

Osim toga, može se pokazati da ukoliko ne odbacujemo hipotezu $H_0 : \tau_i = 0$ vrijedi

$$\frac{SS_E}{\sigma^2} \sim \chi^2(N-a) \qquad \text{i} \qquad \frac{SS_{Tretman}}{\sigma^2} \sim \chi^2(a-1).$$

Teorem 2.1 (Cochranov teorem). *Neka su Z_i međusobno nezavisne, $Z_i \sim \mathcal{N}(0,1)$ za $i = 1, 2, \ldots, \nu$ i neka je*

$$\sum_{i=1}^{\nu} Z_i^2 = Q_1 + Q_2 + \cdots + Q_s,$$

gdje je $s \leq \nu_i$, a Q_i ima ν_i stupnjeva slobode. Tada vrijedi da su $Q_1, Q_2, \ldots, Q_s$ su nezavisne i $Q_i \sim \chi^2(\nu_i)$ ako i samo ako je $\nu = \nu_1 + \nu_2 + \cdots + \nu_s$.

Za dokaz i argumentaciju o nezavisnosti suma vidi [11] i [9]. Zbrajajući stupnjeve slobode od $SS_{Tretman}/\sigma^2$ i SS_E/σ^2 slijedi $(N-a)+(a-1)=N-1$ što je broj stupnjeva slobode od SS_T/σ^2. Prema Teoremu 2.1 slijedi da su $SS_{Tretman}/\sigma^2$ i SS_E/σ^2 nezavisne *chi*-kvadrat distribucije. Ukoliko ne odbacujemo hipotezu o jednakosti očekivanja tretmana, omjer

$$F_0 = \frac{SS_{Tretman}/(a-1)}{SS_E/(N-a)} = \frac{MS_{Tretman}}{MS_E} \tag{2.8}$$

ima F distribuciju sa $a-1$ stupnjeva slobode brojnika i $N-a$ stupnjeva slobode nazivnika. Jednadžba (2.8) je test statistika za hipotezu o jednakosti očekivanja tretmana.

Izvor varijabilnosti	Suma kvadrata	Broj stupnjeva slobode	Srednje kvadratno odstupanje	F_0
Između tretmana	$SS_{Tretman} = n\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2$	$a-1$	$MS_{Tretman}$	$F_0 = \frac{MS_{Tretman}}{MS_E}$
Greška	$SS_E = SS_T - SS_{Tretman}$	$N-a$	MS_E	
Ukupno	$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2$	$N-1$		

Tablica 2.3: Analiza varijance jednofaktorskog modela

Postupak testiranja hipoteze dan je u tablici 2.3 koja se naziva **osnovna ANOVA tablica**. Sada kad smo objasnili jednofaktorski model te kad znamo kako bi trebala izgledati osnovna ANOVA tablica vratimo se na uvodni primjer 2.1 o potrošnji goriva automobila različitih tvornica. Izračunajmo sume kvadrata kako bi kreirali ANOVA tablicu za ovaj primjer.

$$\begin{aligned} SS_{Tretman} =& n\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2 \\ =& 5\Big[\Big(\frac{34,7}{5} - \frac{110,9}{15}\Big)^2 + \Big(\frac{37,5}{5} - \frac{110,9}{15}\Big)^2 + \Big(\frac{38,7}{5} - \frac{110,9}{15}\Big)^2\Big] = 1,685, \end{aligned}$$

Analogno, primjenjujući izraze iz osnovne ANOVA tablice 2.7 i vrijednosti prikazane u primjeru 2.1 slijedi:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2 = 3,149$$

$$SS_E = SS_T - SS_{Tretman} = 1,464.$$

U tablici 2.4 dan je prikaz analize varijance jednofaktorskog modela potrošnje goriva. Ukoliko uzmemo nivo značajnosti $\alpha = 0,05$, kritična vrijednost F testa je $F_{\alpha,2,12} = 3,89$. S obzirom da je $6,907 < 3,89$ zaključujemo da odbacujemo nul-hipotezu H_0 o jednakosti očekivanja potrošnje goriva. Isto nam potvrđuje i p-vrijednost koju smo dobili koristeći statistički alat R i funkciju *aov*. Shodno tome možemo odbaciti hipotezu o jednakosti očekivanja potrošnje goriva između tvornica i zaključiti kako automobili različitih tvornica imaju različitu potrošnju goriva.

2.4 Provjera adekvatnosti modela

Pretpostavili smo da su opservacije dobro opisane modelom

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

Izvor varijabilnosti	Suma kvadrata	Broj stupnjeva slobode	Srednje kvadratno odstupanje	F_0	P-vrijednost
Između tretmana	1,685	2	0,843	6,907	0,01009
Greška	1,464	12	0,122		
Ukupno	3,149	14			

Tablica 2.4: Analiza varijance jednofaktorskog modela

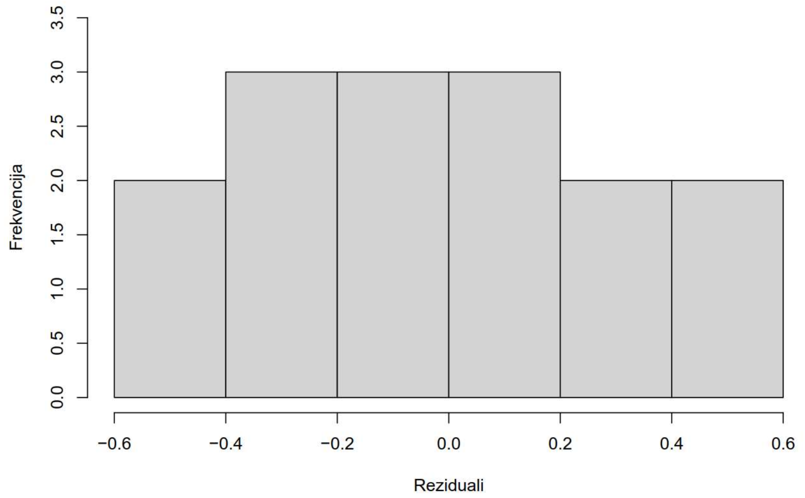
te da su greške nezavisne jednako distribuirane normalne slučajne varijable sa očekivanjem 0 i varijancom σ^2. Ukoliko su ove pretpostavke zadovoljene, analiza varijance je egzaktni test za testiranje hipoteza o jednakosti očekivanja tretmana.
Ipak, ove pretpostavke nije uvijek lako zadovoljiti u praksi pa je sukladno tome uvijek dobro prvo provjeriti valjanost pretpostavki, a tek zatim se osloniti na analizu varijance. Adekvatnost modela, odnosno jesu li su osnovne pretpostavke ispunjene možemo provjeriti pomoću reziduala. Rezidual j-te opservacije i-tog tretmana definira se kao:

$$e_{ij} = y_{ij} - \hat{y}_{ij}, \tag{2.9}$$

gdje je $\hat{y}_{ij}$ procjena opservacije y_{ij} dobivena kao:

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i = \overline{y} + (\overline{y}_i - \overline{y}) = \overline{y}_i. \tag{2.10}$$

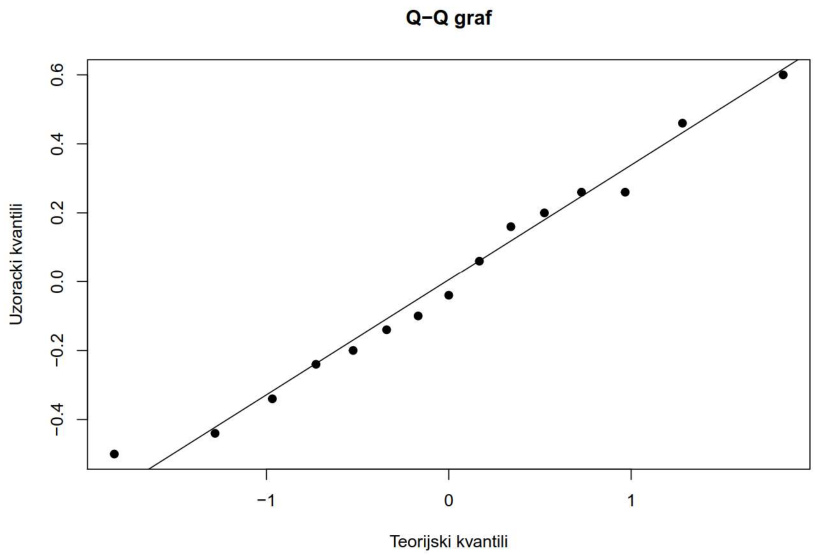
Jednadžba (2.10) govori nam da je procjena opservacije y_{ij} jednaka prosjeku i-tog tretmana. Ispitivanje svojstava reziduala bi trebalo biti dio svake analize varijance. Reziduali bi trebali biti nestrukturirani, odnosno ne bi trebali imati obrazac pronašanja. Analizom reziduala mogu se otkriti neadekvatnosti modela i može se uočiti neispunjavanje pretpostavki modela. Sada ćemo pokazati kako se adekvatnost modela može relativno jednostavno provjeriti grafičkom analizom reziduala.



Slika 2.2: Histogram reziduala

Provjera pretpostavke normalnosti može se izvršiti crtanjem histograma reziduala. Ako je zadovoljena pretpostavka o normalnoj distribuiranosti $N(0, \sigma^2)$ i jednakoj distribuiranosti grešaka modela ϵ_{ij}, histogram bi trebao imati oblik kao uzorak iz normalne distribucije centrirane oko nule. Ukoliko uzorak nije velik, često se javljaju velike fluktuacije u obliku histograma pa pojava umjerenog odstupanja histograma od grafa normalne distribucije ne mora nužno značiti kršenje pretpostavki. Svakako svako odstupanje zahtijeva daljnju analizu.

Na slici 2.2 prikazan je histogram iz primjera 2.1, a s obzirom da naš primjer ima ukupno 15 opservacija, fluktuacije u obliku histograma su, kako smo naveli, česta pojava. Koristan alat za provjeru normalnosti je Q-Q graf (eng. Q-Q plot). Ukoliko je distribucija grešaka normalna, ovaj graf će biti skoro pa pravac. Odstupanja reziduala od pravca sa lijeve ili desne strane grafa ukazuju na moguće postojanje teških repova. Umjerena odstupanja je potrebno promotriti s posebnom pažnjom u analizi modela fiksnih efekata.



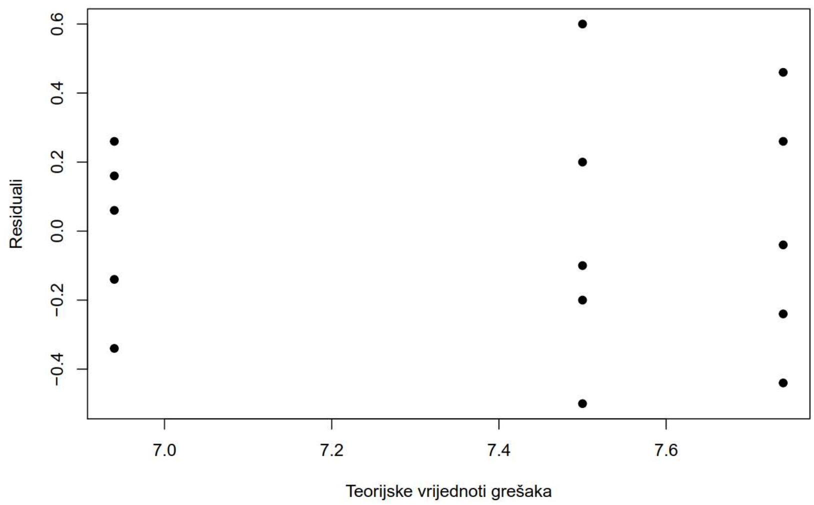
Slika 2.3: Q-Q graf

Na slici 2.3 je prikazan Q-Q graf pa možemo zaključiti kako naš primjer ispunjava pretpostavku o normalnoj distribuiranosti. Nedostatak koji često ima utjecaj na Q-Q graf je rezidual koji ima vrijednost koja znatno odstupa od vrijednosti svih ostalih rezidula. Takav rezidual nazivamo **stršeća vrijednost** (eng. outlier). Prisutnost jedne ili više stršećih vrijednosti može ozbiljno narušiti analizu varijance. Potrebno je detektirati stršeće vrijednosti te ih pažljivo analizirati. Vrlo često je pogreška u mjerenju ili greška pri unosu podataka razlog nastanka stršećih vrijednosti. U tom slučaju ti podaci se mogu korigirati, ukoliko je to moguće, ili izbaciti iz analize. Ako to nije uzrok nastanka stršeće vrijednosti, potrebno je proučiti eksperimentalne okolnosti ovog podatka. Nekada nam informacije proizašle iz analize stršećih vrijednosti mogu biti veoma značajne za buduće zaključke, a u nekim slučajevima će možda biti potrebno provesti dvije analize, sa i bez stršećih vrijednosti. Jedan od jednostavnijih "grubih" načina provjere stršećih vrijednosti je ispitivanje standardiziranih reziduala d_{ij}:

$$d_{ij} = \frac{e_{ij}}{\sqrt{MS_E}}.$$

Ukoliko su greške e_{ij} normalno distribuirane sa očekivanjem 0 i varijancom σ^2, onda su standardizirani reziduali također normalni sa očekivanjem nula i homogeniziranom varijancom.

Ukoliko standardizirani rezidual odstupa više od četiri standardne devijacije od nule, pretpostavljamo da je riječ o stršećoj vrijednosti.



Slika 2.4: Graf reziduala i teorijskih vrijednosti grešaka

Potrebno je još provjeriti homoskedastičnost modela, odnosno pretpostavku da su varijance grešaka konstantne. Može se dogoditi da varijanca grešaka nije konstantna nego da ovisi o vrijednostima ulaznih varijabli. Takvo svojstvo varijance grešaka naziva se heteroskedastičnost. Ukoliko je prisutna heteroskedatičnost potrebno je potražiti adekvatniji model ili primjeniti transformacije kako bi se model sveo na homoskedastičan. Za provjeru je korisno nacrtati graf reziduala usporedno sa teorijskim vrijednostima grešaka modela. Ovdje treba imati na umu da je za jednofaktroski model teorijska vrijednost greške i-tog tretmana jednaka aritmetičkoj sredini i-tog tretmana. Na grafu reziduala usporedno sa teorijskim vrijednostima grešaka modela ne bi trebao postojati očiti obrazac ponašanja reziudala, a ukoliko uočimo nekakav obrzac ponašanja znači da sumnjamo na narušenost pretpostavke. Na slici 2.4 dan je graf reziduala usporedno sa teorijskim vrijednostima modela za naš primjer 2.1 te ne možemo uočiti nikakav obrazac koji bi ukazivao na kršenje pretpostavke homoskedastičnosti. Koristan alat za provjeru je i Shapiro Wilkov test ugrađen u statistički alat R kojim se testira normalnost distribucije. U slučaju p-vrijednosti veće od nivoa značajnosti α (najčešće 0,05 ili 0,01) ne odacujemo hipotezu o normalnosti distribucije. Ukoliko Shapiro Wilkovim testom ne odbacujemo hipotezu o normalnosti distribucije rezudiala, zaključujemo kako reziduli dolaze iz normalne distribucije.

Poglavlje 3

3 Slučajni dizajn punih blokova

3.1 Motivacija

Varijabilnost koja proizlazi iz faktora smetnji može utjecati na rezultat. Općenito, faktor smetnje definiramo kao faktor dizajna koji ima utjecaj na rezultat, ali nas taj utjecaj ne zanima, odnosno želimo ga eliminirati. Ponekad je faktor smetnje skriven i ne možemo kontrolirati što posljedično može dovesti do promjena u rezultatima eksperimenta. Slučajnost je tehnika dizajniranja koja se koristi za izbjegavanje utjecaja takvog skrivenog slučaja smetnje. U drugim slučajevima je faktor smetnje poznat, ali ga također ne možemo kontrolirati. Ukoliko možemo barem promatrati vrijednosti koje faktor smetnje postiže, njegov utjecaj možemo eliminirati iz rezultata eksperimenta korištenjem analize kovarijance. Za više o analizi kovarijance pogledati u [9, str.542 - 577]. Kada je izvor smetnje poznat i kada ga možemo kontrolirati, njegov utjecaj možemo sustavno ukloniti koristeći dizajn slučajnih blokova. Kroz sljedeći primjer ćemo ilustrirati opću ideju dizajna slučajnih blokova.

Primjer 3.1. Promotrimo eksperiment ispitivanja tvrdoće metala. Imamo stroj za ispitivanje tvdoće koji šipku sa šiljastim vrhom pritisne u metalnu leguru. Mjerenjem dubine udubljenja uzrokovanog šiljastim vrhom određuje se tvrdoća metala. Pretpostavimo sada da želimo utvrditi uzrokuju li četiri različita vrha različita mjerenja dobivena strojem za ispitivanje tvrdoće. Eksperimentator je odlučio svakoj metalnoj leguri dodijeliti jednu od četiri opservacije promatranja tvdoće, koristeći Rockwell C skale. Obzirom da postoje četiri tipa šiljastog vrha, da bi se postigao potpuni slučajni jednofaktorski dizajn svaki tip vrha je nasumično dodjeljen jednoj od $4x4 = 16$ eksperimentalnih jedinica, tj. metalnih legura za koje ćemo mjeriti rezultat ispitivanje tvrdoće.

Dizajn prikazan u tablici 3.1 naziva se **slučajni dizajn punih blokova** (eng, randomized complete block design, dalje u tekstu kao: "RCBD"). Ovaj dizajn modeliran je na način da su svi blokovi popunjeni sa jednakim brojem tretmana. Ovim dizajnom blokovi nam formiraju homogenu eksperimentalnu jedinicu. U primjeru (3.1) je to uzorak metala. Ovaj dizajn omogućava bolju usporedbu vrhova šipke eliminirajući varijabilnost između uzoraka metala. Unutar svakog pojedinog bloka se određuje redoslijed testiranja četiri različita vrha šipke. Ako je ovaj primjer testiran na samo dva različita vrha šipke, onda ga možemo modelirati t-testom dvije nezavisne varijable (eng. paired t-test). Više o tome u [9, str. 48 - 50]. Slučajni dizajn punih blokova je generalizacija $pairedt - testa$.

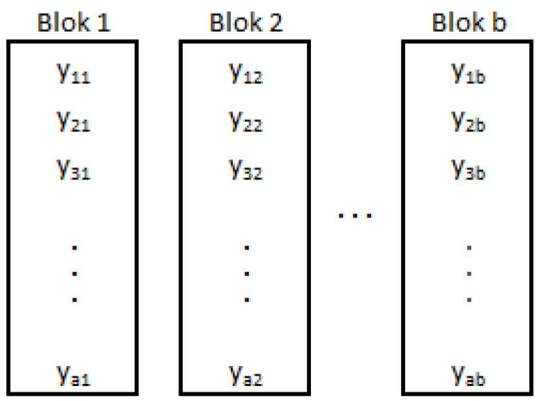
Metal			
1	2	3	4
Vrh 3	Vrh 3	Vrh 2	Vrh 1
Vrh 1	Vrh 4	Vrh 1	Vrh 4
Vrh 4	Vrh 2	Vrh 3	Vrh 2
Vrh 2	Vrh 1	Vrh 4	Vrh 3

Tablica 3.1: Slučajni dizajn punih blokova

RCBD je jedan od najčešće korištenih eksperimenatalnih dizajna. Situacije za koje je RCBD odgovarajući odabir su mnogobrojne. Mjerni strojevi i oprema se često razlikuju u svojim karakteristikama pa se pojavljuju i razni faktori smetnje. Sirovine, ljudi i vrijeme su također varijable koje uzrokuju varijabilnost u eksperimentu, a iste se mogu kontrolirati upotrebom RCBD-a.

3.2 Slučajni dizajn punih blokova

Kao u poglavlju 2 pretpostavimo da imamo a tretmana te dodatno pretpostavimo da imamo b blokova. Slučajni dizajn punih blokova je prikazan na slici (3.1). U svakom bloku postoji jedna opservacija po svakom tretmanu, a redoslijed kojim se tretmani unutar bloka raspoređuju je slučajan, tj. nasumičan. Budući da je jedina slučajnost postignuta unutar blokova, kažemo da blokovi predstavljaju restrikciju slučajnosti.



Slika 3.1: Slučajni dizajn punih blokova

Statistički model RCBD-a može se zapisati na više načina, ali najčešći način zapisa naziva se **model efekata** i dan je jednadžbom kako slijedi:

$$y_{ij} = \mu + \tau_i + \beta j + \epsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}, \tag{3.1}$$

gdje je μ ukupna aritmetička sredina, τ_i efekt i-tog tretmana, a β_j efekt j-tog bloka, dok je ϵ_{ij} normalna slučajna greška s očekivanjem 0 i varijancom σ^2. Pretpostavit ćemo i da su tretmani i blokovi fiksni faktori te da su efekti tretmana τ_i i efekti blokova β_j odstupanja od ukupne aritmetičke sredine tako da vrijedi:

$$\sum_{i=}^{a} \tau_i = 0 \qquad \text{i} \qquad \sum_{j=}^{b} \beta_j = 0.$$

Osim modela efekata možemo koristiti i model očekivanja za RCBD koji je dan jednadžbom:

$$y_{ij} = \mu_{ij} + \epsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases},$$

gdje je $\mu_{ij} = \mu + \tau_i + \beta_j$. Ipak u okviru ovog rada zadržat ćemo se na modelu efekata. Želimo testirati hipoteze o jednakosti očekivanja tretmana:

$$\begin{aligned} H_0 :& \mu_1 = \mu_2 = \cdots = \mu_a \\ H_1 :& \mu_i \neq \mu_j, \qquad \text{za barem jedan par (i,j)} \end{aligned}$$

Kako je

$$\mu_i = \frac{1}{b} \sum_{j=1}^{b} (\mu + \tau_i + \beta_j) = \mu + \tau_i,$$

ekvivalentan zapis ovih hipoteza je:

$$\begin{aligned} H_0 :& \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\ H_1 :& \tau_i \neq 0, \qquad \text{za barem jedan i.} \end{aligned}$$

Analizu varijance možemo jednostavno proširiti na RCBD.
Neka je y_i suma svih opservacija za tretman i

$$y_i = \sum_{j=1}^{b} y_{ij} \qquad i = 1, 2, \ldots, a \tag{3.2}$$

te neka je y_j suma svih opservacija u bloku j

$$y_j = \sum_{i=1}^{a} y_{ij} \qquad j = 1, 2, \ldots, b, \tag{3.3}$$

a y suma svih ukupnih opservacija

$$y = \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij} = \sum_{i=1}^{a} y_i = \sum_{j=1}^{b} y_j. \tag{3.4}$$

Tada je $N = ab$ ukupan broj svih opservacija.
Slično, y_i je prosjek svih opservacija tretmana i, y_j je prosjek svih opservacija unutar bloka j, dok je y ukupan prosjek svih N opservacija,

$$\overline{y}_i = \frac{y_i}{b} \qquad \overline{y}_j = \frac{y_j}{a} \qquad \overline{y} = \frac{y}{N}. \tag{3.5}$$

Sada možemo ukupnu sumu kvadrata zapisati kao:

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \overline{y})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}[(\overline{y}_i - \overline{y}) + (\overline{y}_j - \overline{y}) + (y_{ij} - \overline{y}_i - \overline{y}_j + \overline{y})]^2. \tag{3.6}$$

Raspišimo sada desnu stranu jednakosti:

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \overline{y})^2 =& b\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2 + a\sum_{j=1}^{b}(\overline{y}_j - \overline{y})^2\\ &+\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \overline{y}_i - \overline{y}_j + \overline{y})^2\\ &+2\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{y}_i - \overline{y})(\overline{y}_j - \overline{y})\\ &+2\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{y}_i - \overline{y})(y_{ij} - \overline{y}_i - \overline{y}_j + \overline{y})\\ &+2\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{y}_j - \overline{y})(y_{ij} - \overline{y}_i - \overline{y}_j + \overline{y}).\end{aligned}$$

Koristeći formule (3.2), (3.3) i (3.4) te jednostavni algebarski račun može se pokazati da su zadnja tri člana jednakosti jednaka nuli. Pa sukladno tome,

$$\begin{aligned}\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \overline{y})^2 =& b\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2 + a\sum_{j=1}^{b}(\overline{y}_j - \overline{y})^2\\ &+\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \overline{y}_i - \overline{y}_j + \overline{y})^2\end{aligned} \tag{3.7}$$

gornja formula (3.7) predstavlja rastav ukupne sume kvadrata, te je možemo zapisati i simbolično kao:

$$SS_T = SS_{Tretman} + SS_{Blok} + SS_E. \tag{3.8}$$

Jednadžbe (3.7) i (3.8) predstavljaju fundamentalne ANOVA jednadžbe za Slučajni dizajn punih blokova, tj. RCBD. S obzirom da u modelu imamo N opservacija, SS_T ima $N - 1$ stupnjeva slobode. Imamo a tretmana i b blokova pa $SS_{Tretman}$ i SS_{Blok} imaju $a-1$, odnosno $b - 1$ stupnjeva slobode. Suma kvadrata grešaka je suma kvadrata između svih opservacija minus suma kvadrata tretmana i blokova pa SS_E ima $ab-1-(a-1)-(b-1) = (a-1)(b-1)$ stupnjeva slobode. Ukoliko zbrojimo stupnjeve slobode desne strane jednadžbe (3.8), slijedi $(a - 1) + (b - 1) + (a - 1)(b - 1) = ab - 1 = N - 1$ te dodatno, uz pretpostavku normalano distribuiranih nezavisnih grešaka sa očekivanjem 0 i varijancom σ^2, zadovoljene su sve pretpostavke Teorema 2.1 iz čega slijedi da su $SS_{Tretman}/\sigma^2$, SS_{Blok}/σ^2 i SS_E/σ^2 međusobodno nezavisne *chi*-kvadrat distribucije. Svaka od navedenih suma podijeljena sa pripadajućim brojem stupnjeva slobode čini srednje kvadratno odstupanje. Očekivano srednje kvadratno odstupanje je prikazano kako slijedi:

$$E(MS_{Tretman}) = \sigma^2 + \frac{b\sum_{i=1}^{a}\tau_i^2}{a-1}, \qquad E(MS_{Blok}) = \sigma^2 + \frac{a\sum_{j=1}^{b}\beta_j^2}{b-1}, \qquad E(MS_E) = \sigma^2.$$

Prema [9], za testiranje postavljenih hipoteza o jednakosti očekivanja koristimo test statistiku:

$$F_0 = \frac{MS_{Tretman}}{MS_E} \tag{3.9}$$

koja ima $F_{a-1,(a-1)(a-1)}$ distribuciju, ukoliko ne odbacujemo nul hipotezu o jednakosti očekivanja. Kritičko područje značajnosti α je desni rep F distribucije, odnosno obacit ćemo hipotezu H_0 ukoliko vrijedi $F_0 > F_{\alpha,a-1,(a-1)(b-1)}$. Zanima nas i usporedba srednjih vrijednosti blokova. Ukoliko se srednje vrijednosti blokova ne razlikuju značajno, kreiranje blokova u budućim eksperimentima možda nije nužno potrebno. Promatrajući očekivanja srednje kvadranih vrijednosti čini se da bi hipotezu $H_0 : \beta_j = 0$ mogli testirati uspoređivanjem statistike $F_0 = MS_{Blok}/MS_E$ sa $F_{\alpha,b-1,(a-1)(b*1)}$. Međutim već smo prije komentirali kako se slučajnost primjenjuje samo unutar blokova, tj. da blokovi predstavljaju ograničenje slučajnosti. Pojasnimo sada kako ograničenje slučajnosti može utjecati na statistiku $F_0 = MS_{Blok}/MS_E$. Prema [3] F test može biti opravdan samo na temelju slučajnosti, bez upotrebe pretpostavke o normalnosti. Kako je u RCBD modelu prisutna slučajnost samo unutar blokova, ali ne i između blokova, Box, Hunter i Hunter ([3]) zaključuju da se usporedba jednakosti srednjih vrijednosti blokova ne može vršiti F testom, osim u slučaju kada su greške nezavisne i normalno distribuirane sa očekivanjem 0 i varijancom σ^2. S druge strane, Anderson i McLean [1] navode kako ograničenje slučajnosti spriječava statistiku F_0 da bude ispravan test za uspoređivanje srednjih vrijednosti blokova te da je F statistika zapravo test za jednakost srednjih vrijednosti i ograničenje slučajnosti. Pa možemo reći, da ukoliko je pretpostavka normalnosti zadovoljena, možemo koristiti F test. Bitno je zanti da u praksi to ipak često nije tako. Zbog toga, isključujemo F test iz tablice analize varijance. Međutim, za približan postupak ocjenjivanja efekta blokova ipak koristimo omjer $MS_{Tretman}/MS_E$. Ukoliko je ovaj omjer velik, podrazumijeva se da faktor blokova ima veliki utjecaj i da su blokovi pomogli pri smanjivanju utjecaja faktora šuma. Ovaj postupak je prikazan u Tablici (3.2).

Pojasnimo još sume kvadrata iz tablice (3.2):

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y^2}{N} \tag{3.10}$$

$$SS_{Tretman} = \frac{1}{b}\sum_{i=1}^{a} y_i^2 - \frac{y^2}{N} \tag{3.11}$$

$$SS_{Blok} = \frac{1}{a}\sum_{j=1}^{b} y_j^2 - \frac{y^2}{N} \tag{3.12}$$

dok sumu kvadrata grešaka možemo izračunati iz izraza:

$$SS_E = SS_T - SS_{Tretman} - SS_{Blok}. \tag{3.13}$$

Izvor varijabilnosti	Suma kvadrata	Broj stupnjeva slobode	Srednje kvadratno odstupanje	F_0
Između tretmana	$SS_{Tretman}$	$a-1$	$\frac{SS_{Tretman}}{a-1}$	$F_0 = \frac{MS_{Tretman}}{MS_E}$
Unutar blokova	SS_{Blok}	$b-1$	$\frac{SS_{Blok}}{b-1}$	
Greška	SS_E	$(a-1)(b-1)$	$\frac{SS_E}{(a-1)(b-1)}$	
Ukupno	SS_T	$N-1$		

Tablica 3.2: Analiza varijance za Slučajni dizajn punih blokova

Primjer 3.2. Proizvađač medicinskih proizvoda izrađuje vaskularne graftove - umjetne vene. Ovi graftovi se proizvode u cijevi od politetrafluoretilen (PTFE) smole zajedno sa dodatnim lubrikantom. Tijekom proizvodnog procesa te cijevi često sadrže male, tvrde izbočine na vanjskoj površini. Ove nepravilnosti su često razlog odbacivanja grafta. Proizvođač sumnja da tlak pri ekstruziji utječe na pojavu izbočina te namjerava provesti eksperiment kako bi istražio ovu hipotezu. Smolu proizvodi vanjski dobavljač te je isporučuje u serijama. Proizvođač sumnja da mogu postojati varijacije između serija. Smola bi trebala biti u skaldu s parametrima kao što su molekularna težina, srednja veličina čestica i slično, ali zbog odstupanja pri izradi i prirodnog odstupanja materijala vjerojatno nisu uvijek u potpunosti u skladu s parametrima. Sukladno sumnjama proizvođač odlučuje istražiti učinak četiri različite razine tlaka pri ekstruziji na izbočine u cijevi koristeći slučajni dizajn punih blokova (RCBD), a kako je prikazano u tablici 3.3. U dizajnu punih blokova imamo četiri razine tlaka (tretmani) i šest serija isporuke smole (blokovi).

Razina tlaka	Serija isporuke smole 1	2	3	4	5	6	Tretman ukupno
8500	90,3	89,2	98,2	93,9	87,4	97,9	556,9
8700	92,5	89,5	90,6	94,7	87,0	95,8	550,1
8900	82,5	90,8	89,6	86,2	88,0	93,4	533,5
9100	82,5	89,5	85,6	87,4	78,9	90,7	514,6
Blok ukupno	350,8	359,0	364,0	362,2	341,3	377,8	$y = 2155,1$

Tablica 3.3: Slučajni dizajn punih blokova za eksperiment vaskularnih graftova

Sjetimo se da je redoslijed odabira tlaka pri ekstruziji untar bloka slučajan. Opervacije predstavljaju postotak cijevi koje u proizvodnom procesu nisu sadržavale nikakve izbočine.

Izračunajmo sada sume kvadrata koje su nam potrebne za analizu varijance za slučajni dizajn punih blokova.

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y^2}{N} = 193.991,31 - \frac{(2.155,1)^2}{24} = 480,31$$

$$SS_{Tretman} = \frac{1}{6}\Big[(556,9)^2 + (550,1)^2 + (533,5)^2 + (514,6)^2\Big] - \frac{(2.155,1)^2}{24} = 178,17$$

$$SS_{Blok} = \frac{1}{4}\Big[(350,8)^2 + (359,0)^2 + \cdots + (377,8)^2\Big] - \frac{(2.155,1)^2}{24} = 192,25$$

Tablica 3.4 je tablica analize varijance za RCBD. Ako uzmemo $\alpha = 0,05$, kritična vrijednost F-testa je $F_{\alpha,3,15} = 3,29$, pa zaključujemo da razina tlaka pri ekstruziji utječe na postotak broja izbočina u cijevi. Također primjetimo da je srednje kvadratno odstupanje blokova ima relativno visoku vrijednost u odnosu na srednje kvadratno odstupanje grešaka.

Izvor varijabilnosti	Suma kvadrata	Broj stupnjeva slobode	Srednje kvadratno odstupanje	F_0	P-vrijednost
Između tretmana	178,17	3	59,39	8,11	0,0019
Unutar blokova	192,25	5	38,45		
Greška	109,89	15	7,33		
Ukupno	480,31	23			

Tablica 3.4: Analiza varijance za RCBD za eksperiment vaskularnih graftova

Promotrimo sada kako bi izgledali rezultati da nismo koristili slučajni dizajn punih blokova, već jednofaktorski model analize varijance. Za jednofaktroski model opservacije ne bi bile podijeljene u serije. U tom slučaju promatrali bi $n = 6$ opservacija tretmana a, pri čemu $a = 1,2,3,4$. Koristeći izraz iz talice 2.3 izračunajmo sumu kvadrata tretmana:

$$\begin{aligned} SS_{Tretman} =& n\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2 \\ =& 6\Big[\Big(\frac{556,9}{6} - \frac{2155,1}{24}\Big)^2 + \cdots + \Big(\frac{514,6}{6} - \frac{2155,1}{24}\Big)^2\Big] = 178,17 \end{aligned}$$

Analogno, uvršavajući brojeve iz tablice 3.3 u izraze za ukupnu sumu kvadrata i sumu kvadrata grešaka dobivamo:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \overline{y})^2 = 480,31$$

$$SS_E = SS_T - SS_{Tretman} = 480,31 - 178,17 = 302,14$$

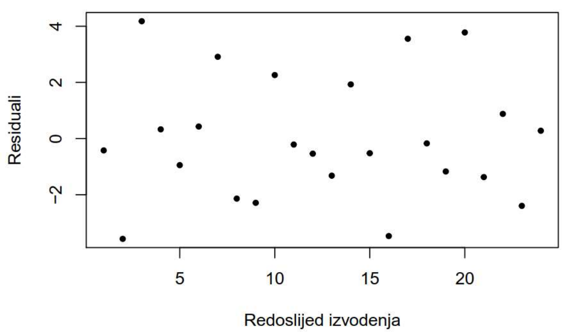
Izvor varijabilnosti	Suma kvadrata	Broj stupnjeva slobode	Srednje kvadratno odstupanje	F_0	P-vrijednost
Između tretmana	178,17	3	59,39	3,95	0,0235
Greška	302,14	20	15,11		
Ukupno	480,31	23			

Tablica 3.5: Analiza varijance jednofaktorskog modela za eksperiment vaskularnih graftova

U tablici 3.5 dan je prikaz analize varijance jednofaktorskog modela za eksperiment vaskularnog grafta. Ukoliko uzmemo da je nivo značajnosti $\alpha = 0,05$ također ćemo odbaciti nul hipotezu o jednakosti očekivanja tretmana, odnosno možemo prihvatiti hipotezu da razina tlaka pri ekstruziji utječe na postotak broja izbočina u cijevi. Međutim, zanimljivo za promotranje jesu razlike između srednje kvadratnih odstupanja grešaka dobivenih primjenom RCBD-a i jednofaktorskog modela. Uočavamo da je srednje kvadratno odstupanje grešaka u jednofaktorkom modelu iznosilo $15,11$ što je dvostruko više od srednje kvadratnog odstupanja grešaka u RCBD modelu koje je iznosilo $7,33$. Sva varijabilnost, koja je postojala između blokova u RCBD modelu, je u jednofaktorskom modelu sadržana unutar varijabilnosti grešaka. Iz ove usporedbe očito je zašto RCBD nazivamo i tehnikom dizajniranja za smanjivanje faktora šuma. RCBD povećava preciznost kojom uspoređujemo tretmane. Ova usporedba također dovodi do zaključka da ako ne primjenimo dizajn blokova u eksperimentima gdje smo trebali, učinak faktora šuma može biti toliko značajan da posljedično poveća srednje kvadratno odstupanje eksperimentalnih grešaka do te mjere da se razlike u očekivanju tretmana ne mogu utvrditi.

3.3 Provjera adekvatnosti modela

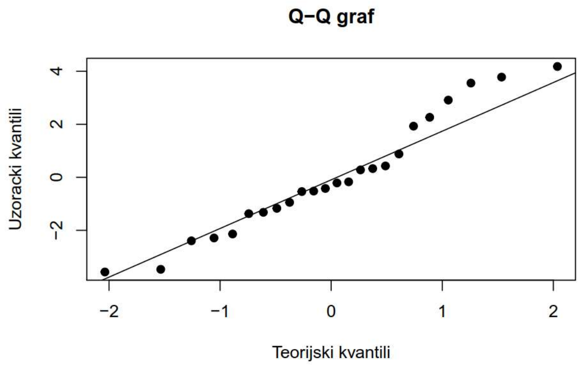
Kao u poglavlju 2, potrebno je provesti provjeru adekvatnosti modela kako bi provjerili postavljene pretpostavke. Potencijalni problemi pojavljuju se kod pretpostavki o normalnosti, kod različitosti varijance greške tretmana ili blokova te kod interakcije između blokova i tretmana. Jednako kao i kod jednofaktorskog modela analize varijance, reziduali su glavni alat za provjeru adekvatnosti.



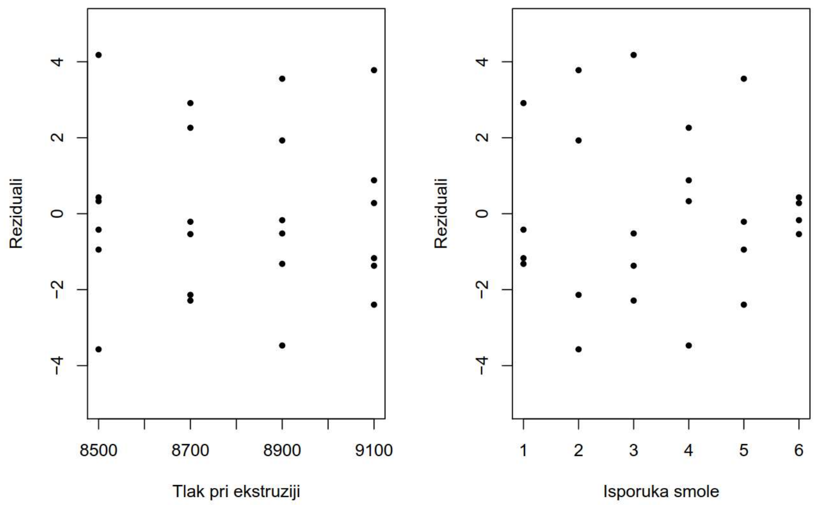
Slika 3.2: Graf rezidala eksperimenta vaskularnih graftova

Graf reziduala za eksperiment vaskularnih graftova dan je na slici 3.2 te promatrajući ga nemamo razloga sumnjati u kršenje pretpostavke o nezavisnoti reziduala. Q-Q graf reziduala dan je na slici 3.3. Shapiro Wilkov testom, koristeći statistički alat R, je testirana normalnost reziduala, te dobivena p-vrijednost iznosi 0,3689 pa ne odbacujemo hipotezu o normalnosti distribucije reziduala. Ne postoje ozbiljni razlozi za sumnju u normalnost distribucije reziduala niti postoje dokazi da postoje stršeće vrijednosti.

U analizi reziduala slučajnog dizajna punih blokova dobro je još promotriti i odnose reziduala i tretmana, te odnos reziduala i blokova. Ako je za određeni broj tretmana više varijabilnosti između reziduala, to bi moglo značiti da ovaj tretman stvara više pogrešnih opservacija od ostalih tretmana. Više varijabilnosti u rezidualima u odnosu na blokove, može značiti da blok nije homogen. U našem primjeru sa slike 3.4 ne primjećujemo nejadnokost varijance reziduala između tretmana, no primjećujemo nešto manju varijancu reziduala za isporuku broj šest u odnosu na sve ostale isporuke. Međutim kako su svi ostali grafovi zadovoljavajući, smatrat ćemo da je model adekvatan.



Slika 3.3: Q-Q graf rezidala eksperimenta vaskularnih graftova



Slika 3.4: Graf rezidala u odnosu na tretmane i blokove

Literatura

[1] V.L.ANDERSON i R.A.MCLEAN, *Design of Experiment: A Realistic Approach Dekker*, Dekker, New York, 1974.

[2] M. BENŠIĆ i N. ŠUVAK, *Uvod u vjerojatnost i statistiku*, Sveučilište J.J. Strossmayera, Odjel za matematiku, Osijek, 2014.

[3] G.E.P. BOX, J.S.HUNTER i W.G.HUNTER, *Statistics for Experimenters*, Wiley, 2nd Edition, 2005.

[4] D.E.COLEMAN i D.C.MONTGOMERY, *A Systematic Approach to Planning for Designed Industrial Experiment*, Technometrics, 1993., Vol. 35

[5] N.R.P.COSTA, A.R.PIRES I C.O.RIBEIRO, *Guidelines to help practitioners of design of experiments*, The TQM Magazine, Srpanj 2006.

[6] R. A. FISHER, *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 1958.

[7] R. A. FISHER, *The design of Experiments*, Hafner Publishing Company, New York, 1966.

[8] M.G.KENDALL i A.STUART, *The Advanced Theory Of Statistics, Volume 3*, Charles Griffin & Company Limited, London, 1966.

[9] DOUGLAS C. MONTGOMERY, *Design and analysis of experiments*, Wiley, 7th edition, New York, 2008.

[10] ŽELJKO PAUŠE, *Uvod u matematičku statistiku*, Školska knjiga, Zagreb, 1993.

[11] G. A. F. SEBER i A. J. LEE, *Linear Regression Analysis*, Wiley, 2nd edition, New Jersey, 2003.

[12] *Handbook of Statistical Methods*, www.itl.nist.gov/div898/handbook/

Sažetak

U ovom radu predstavljene su ANOVA procudure i dizajn eksperimenta. Na početku rada dan je kratak osvrt na povijest statističkog dizajna eksperimenta. Nadalje, u prvom poglavlju dane su smjernice za dizajniranje eksperimenta kroz sedam koraka. U nastavku rada bavili smo se ANOVA procedurama. Opisan je jednofaktroski model analize varijance te je pojašnjen način provjere adekvatnosti modela. U slučaju nešto složenijih dizajna, kada želimo otkloniti dodatne utjecaje faktora smetnji, koristimo slučajni dizajn punih blokova koji je opisan u poglavlju tri. Kroz primjer su pojašnjene razlike između jednofaktorskog modela analize varijance i modela slučajnog dizajna punih blokova. Za kraj je provedena provjera adekvatnosti modela slučajnog dizajna punih blokova.

Ključne riječi: Dizajn eksperimenta, ANOVA, jednofaktorski model, slučajni dizajn punih blokova, provjera adekvatnosti modela

The design of experiment and ANOVA procedures

Summary

This graduate thesis introduce the design of experiment and ANOVA procedures. A brief history of statistical design is given at the beginning. Furthermore, the first chapter describes guidelines for designing experiments trow seven steps. In the next two chapters are written ANOVA procedures. Chapter two includes analysis of fixed effects model with model adequacy checking. Random block design, which we use to eliminate the nuisance factor is explained in chapter three. Differences between ANOVA procedures given in chapter two and three are explained throw the example. Finally, model adequcy checking for random block design is given at the end.

Keywords: The design of experiment, ANOVA procedures, Fixed effect model, Random Block design, Model adequcy checking

Životopis

Matea Radan je rođena 2. srpnja 1993. godine u Dubrovniku. U prvim godina formalnog obrazovanja pohađala je Osnovnu školu Ivana Gundulića u Dubrovniku, a zatim 2008. godine upisuje prirodoslovno matematički smjer Gimnazije Dubrovnik. Po završetku srednje škole 2012. godine odlučuje se za Preddiplomski studij matematike na Odjelu za matematiku Sveučilišta J.J. Strossmayera u Osijeku koji završava 2016. godine izradom završnog rada Cobb-Douglasova funkcija uz mentorstvo doc. dr. sc. Snježane Majstorović čime stječe akademski naziv prvostupnice matematike. Iste godine nastavlja školovanje na Odjelu za matematiku u Osijeku odabirom Diplomskog studija financijske matematike i statistike. U srpnju 2018. godine preko studentskog servisa počinje raditi u Zagrebačkoj banci u poslovnici u Osijeku kao promotorica digitalnih platformi. U ožujku 2019. godine je primljena na stručnu praksu " ZABA Banking Academy" Zagrebačke banke u odjelu Korporativnog bankarstva. Praksu je obavljala u Regionalnom centru Istočne Hrvatske u Osijeku na poziciji Pomoćnice voditelja poslovnog odnosa korporativnog bankrstva. Trenutačno je zaposlena u Zagrebačkoj banci, na istoj poziciji s fokusom rada u području kreditnih analiza poslovnih subjekata korporativnog bankarstva.