

SVEUČILIŠTE U ZAGREBU AGRONOMSKI FAKULTET

# Dizajn eksperimenata s primjenom u poljoprivredi

---

Priručnik s primjerima u R-u

**Jerko Gunjača**

**1/23/2018**



# Sadržaj

Uvod .....	3
Literatura: .....	4
Linearni modeli – provjera pretpostavki .....	5
Primjer 1.1 .....	6
Primjer 1.2 .....	11
Primjer 1.3 .....	17
Literatura: .....	22
Višestruke usporedbe .....	23
Primjer 2. ....	24
Literatura: .....	28
Višefaktorijsni pokusi .....	30
Primjer 3.1 .....	32
Primjer 3.2 .....	34
Literatura: .....	38
Dizajni s potpunim blokovima .....	40
Primjer 4.1 .....	41
Primjer 4.2 .....	46
Literatura: .....	48
Dizajni s nepotpunim blokovima .....	49
Primjer 5.1 .....	50
Primjer 5.2 .....	54
Literatura: .....	59
Dizajni s razdijeljenim parcelama .....	60
Primjer 6. ....	60
Literatura: .....	64
Dodatak A: Modeli .....	66
Dodatak B: Primjeri .....	67

# Uvod

Osnovna je namjera ovog priručnika prikazati mogućnosti primjene alata za statističku analizu iz programskog okruženja R na nekoliko primjera iz područja dizajna eksperimenata. Priručnik je primarno namijenjen studentima diplomskih studija Sveučilišta u Zagrebu Agronomskog fakulteta, i to kao nastavni materijal za predmet „Eksperimentalni dizajn“. To je izborni predmet za studente diplomskih studija, a u ovom je obliku sastavni dio studijskih programa Sveučilišta u Zagrebu Agronomskog fakulteta od 2006. godine. U studijskim programima iz prethodnih desetljeća, većina programskih cjelina od kojih se sastoji ovaj predmet bila je sadržana u predmetima koji su osim dizajna eksperimenata obuhvaćali i osnove biometrike (statistike), na kojima su nastavu izvodili nastavnici Zavoda za oplemenjivanje bilja, genetiku i biometriku. Osamdesetih i devedesetih godina dvadesetog stoljeća taj se predmet zvao „Planiranje i statistička analiza pokusa“, a nakon reforme studija 1997. godine, uz neznatne izmjene sadržaja, nosio je naziv „Biometrika i planiranje pokusa u bilinogojstvu“. Nakon reforme sustava visokog obrazovanja („Bolonjski proces“), osnove biometrike i dizajn eksperimenata se razdvajaju u dva zasebna predmeta.

Sadržaj predmeta „Eksperimentalni dizajn“ osim cjelina prethodno obuhvaćenih predmetom „Biometrika i planiranje pokusa u bilinogojstvu“ uključuje i nekoliko novih tema: „Linearni modeli“, „Provjera pretpostavki“, „Višestruke usporedbe“ i „Dizajni s nepotpunim blokovima“. Zbog toga se umjesto udžbenika „Biometrika i eksperimentiranje u bilinogojstvu“ prof. Đurđice Vasilj, napisanog za potrebe ranijeg predmeta, u nastavi iz „Eksperimentalnog dizajna“ kao udžbenik koristi „Design of Experiments“ Roberta O. Kuehla. Upravo je to i razlog zašto se ova knjiga navodi kao primarni literaturni izvor na popisu literature za svako od poglavlja u ovom priručniku.

U svakom se poglavlju navode i modeli koji se koriste u statističkoj analizi opisanih primjera. Umjesto simboličkog načina, koristi se deskriptivni sustav zapisa modela koji su predložili Piepho i sur. (2003), uz jednu modifikaciju opisanu u poglavlju „Višefaktorijelni pokusi“. Poglavlja su koncipirana tako da se sastoje od uvodnog dijela i jednog ili više primjera. Uvodni dio sadrži vrlo kratak sažetak

teoretskih osnova čije je poznavanje nužan preduvjet za razumijevanje postupaka koji se primijenjuju u statističkoj analizi, a detaljna se objašnjenja mogu pronaći u dva ranije navedena udžbenika. Za svaki je primjer na početku ukratko opisana struktura podataka, zatim se redak po redak objašnjava R kod koji se koristi za njihovu analizu, te na kraju interpretiraju rezultati. Detaljnije upute se za funkcije iz osnovnog paketa R-a mogu pronaći u dvije reference s popisa literature na kraju ovog poglavlja, a za specijalizirane su pakete reference navedene na kraju poglavlja u kojima se koriste.

Većina primjera potječe iz znanstvenih istraživanja, pa je za svaki takav primjer naveden i izvor, odnosno publikacija u kojoj su objavljeni rezultati istraživanja (osim za Primjer 5.2 iz istraživanja čiji rezultati još nisu objavljeni). Primjer 1.1 temelji se podacima koje su prikupili studenti u sklopu nastave na predmetu „Osnove biometrike“, a dva primjera (4.1 i 4.2) na podacima iz službenih sortnih pokusa. Podatci za preostali primjer (1.3) ustupljeni su ljubaznošću izv. prof. dr. sc. Vesne Židovec sa Zavoda za ukrasno bilje i krajobraznu arhitekturu Sveučilišta u Zagrebu Agronomskog fakulteta.

## Literatura:

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.
2. Vasilj Đ. (2000) Biometrika i eksperimentiranje u bilinogojstvu. Hrvatsko agronomsko društvo, Zagreb.
3. Piepho H.P., Büchse A., Emrich K. (2003) A Hitchhiker's Guide to Mixed Models for Randomized Experiments. Journal of Agronomy & Crop Science, 189: 310-322.
4. Venables W.N., Smith D.M., R Core Team (2004) Uvod u korištenje R-a. <https://cran.r-project.org/doc/contrib/Kasum+Legovic-UvodUr.pdf>
5. The R Core Team (2017) R: A Language and Environment for Statistical Computing – Reference Index. <https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf>

## Linearni modeli – provjera pretpostavki

Opći linearni model naziv je koji obuhvaća različite statističke modele koji jednu (zavisnu) varijablu definiraju kao linearnu funkciju jedne ili više (nezavisnih) varijabli. Između ostalih specifični oblici općeg linearnog modela su: regresijski model, u kojem su i zavisna i nezavisne varijabla kontinuirane; te model analize varijance, u kojem je zavisna varijabla kontinuirana, a nezavisna kategorijska. Model analize varijance osnovni je model, koji će se u različitim oblicima pojavljivati u ovom priručniku, a u najjednostavnijem se obliku može definirati ovako:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (1)$$

U ovom izrazu  $y_{ij}$  je oznaka za  $j$ -to opažanje zavisne varijable unutar  $i$ -te kategorije,  $\mu$  je oznaka za ukupni prosjek,  $\tau_i$  je oznaka za učinak  $i$ -tog tretiranja (odnosno razlika između prosjeka  $i$ -te kategorije i ukupnog prosjeka), a  $\varepsilon_{ij}$  je slučajna pogreška (rezidual). Kod definiranja modela u većini se statističkih programa podrazumijeva da u načelu svaki model sadrži ukupni prosjek i pogrešku, pa se ti efekti u programskom kodu ne moraju posebno zadavati. To načelo vrijedi i za R, pa će se u ovom priručniku koristiti ova pojednostavnjena definicija osnovnog modela analize varijance:

$$Y = TRET \quad (2)$$

Valjanost statističke analize provedene prema nekom linearnom modelu ovisi o ispunjenju tri pretpostavke koje se odnose na pogreške, odnosno razlike između procjena i opažanja. Te tri pretpostavke su.

1. pogreške su međusobno neovisne,
2. njihova distribucija je normalna, s prosječnom vrijednošću nula,
3. varijance pogrešaka unutar svih kategorija su jednake (homogene).

Ukoliko je bilo koja od ovih pretpostavki narušena, zaključci doneseni na temelju rezultata provedene analize se ne mogu smatrati valjanima.

Osnovni alat za provjeru ovih pretpostavki su dijagnostički grafikoni pogrešaka (reziduala). Najjednostavniji dijagnostički grafikon prikazuje odnos između procjena i

pogrešaka (reziduala), na temelju kojeg se manje ili više mogu provjeriti sve navedene pretpostavke. Za provjeru pretpostavke o normalnoj distribuciji pogrešaka specijaliziran je grafikon distribucije reziduala, a za provjeru pretpostavke o homogenosti varijanci grafikon disperzije reziduala. Osim pomoću ovih grafikona pretpostavke se mogu provjeriti i uz pomoć specijaliziranih testova kao što je npr. Shapiro-Wilkov test (Shapiro i Wilk, 1965).

Ukoliko se utvrdi da pri analizi izvornih podataka pretpostavke nisu zadovoljene, podaci se mogu modificirati, odnosno transformirati. Jedan od mogućih razloga za narušavanje pretpostavki je postojanje atipičnih podataka (outlier-a), koji na rezultate analize utječu znatno više nego ostali podatci, te ih čine pristranima. Najbolje rješenje u takvim situacijama je izbacivanje atipičnih podataka, te provođenje analize na reduciranom setu podataka. Kada pogreške ne slijede normalnu distribuciju, kao i kada varijance nisu homogene, transformacijom se izvornih varijabli ovi problemi mogu otkloniti. Pomoć pri traženju optimalnih transformacija nude Tukeyjeve ljestve (Tukey, 1977) ili Box-Coxove transformacije (Box i Cox, 1964).

## Primjer 1.1

U uzorku od 561 biljke soje sorte Hrvatica prikupljeni su podatci o broju mahuna i broju zrna za svaku biljku. Mahuna soje može sadržavati do četiri sjemenke, najčešće ih ima dvije do tri, a odnos broja zrna i broja mahuna u ovom primjeru ispitati će se regresijskom analizom.

### *1. Učitavanje podataka i analiza ovisnosti broja zrna o broju mahuna*

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
soja <- read.csv("Primjer1_1.csv")
```

Zatim se provodi regresijska analiza ovisnosti broja zrna o broju mahuna, primjenom funkcije `lm` (funkcija koja provodi analizu prema linearnom modelu koji je naveden kao prvi argument funkcije):

```
soja.lm1 <- lm(broj.zrna~broj.mahuna,data=soja)
```

Procijenjene vrijednosti konstantnog člana (intercept) i regresijskog koeficijenta mogu se ispisati primjenom generičke funkcije `coefficients`:

```
coefficients(soja.lm1)
(Intercept) broj.mahuna
-1.852709    2.374125
```

Na temelju procijenjene vrijednosti regresijskog koeficijenta može se zaključiti da je prosječan broj zrna po mahuni u ovom primjeru 2.37.

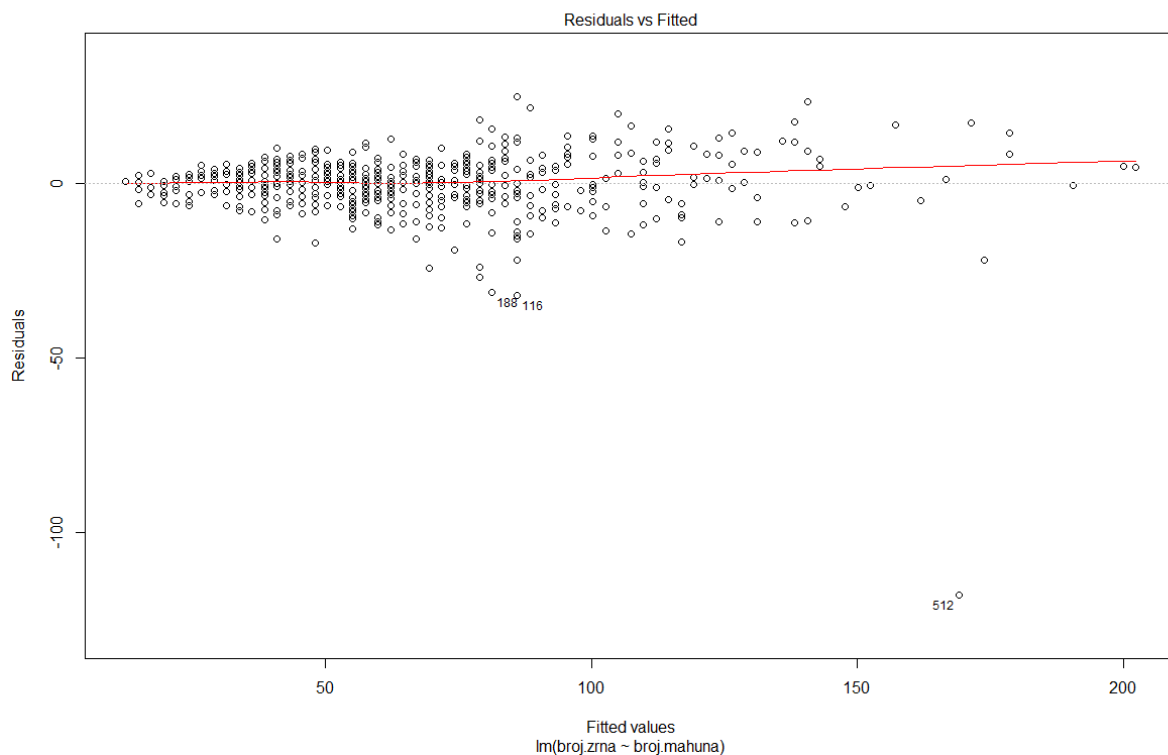
## 2. *Provjera pretpostavki*

Dijagnostički grafikoni generiraju se primjenom generičke funkcije `plot` ako joj je se kao argument navede ime objekta u kojem su spremljeni rezultati analize:

```
plot(soja.lm1)
```

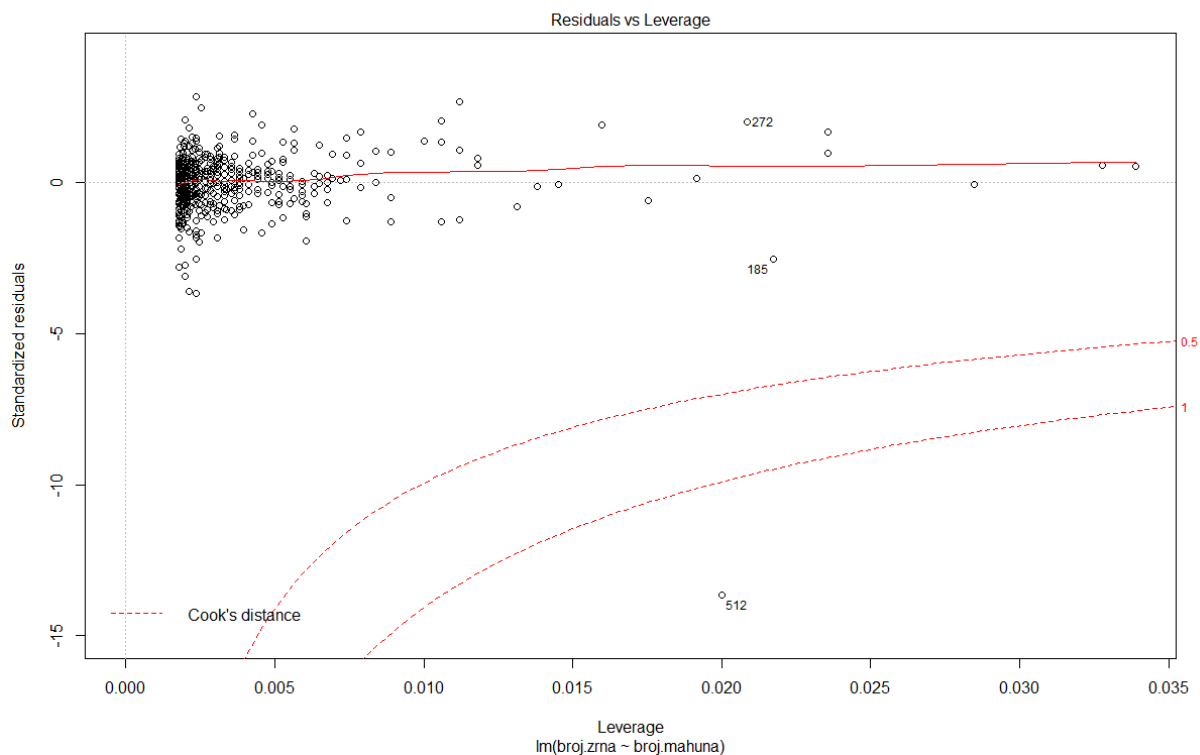
Ovdje će biti prikazana samo dva od četiri dijagnostička grafikona koji se generiraju primjenom funkcije `plot`. Prvi grafikon prikazuje odnos procjena (`Fitted values`) i pogrešaka (`Residuals`):





Slika 1. Odnos procjena i pogrešaka u regresijskoj analizi odnosa broja zrna i broja mahuna kod soje

Na ovom je grafikonu jasno uočljiv jedan atipičan podatak. To je podatak za biljku soje pod rednim brojem 512. koja ima oko 100 zrna manje od očekivanog broja procijenjenog na temelju broja mahuna. Koliki je utjecaj (**Leverage**) te atipične biljke na rezultate analize može se provjeriti uz pomoć četvrtog dijagnostičkog grafikona:



Slika 2. Odnos utjecaja i pogrešaka u regresijskoj analizi odnosa broja zrna i broja mahuna kod soje

Cookova distanca za određeno opažanje mjerilo je ukupne promjene svih vrijednosti u regresijskom modelu ukoliko se taj podatak izostavi iz analize. Za atipičnu biljku br. 512 vrijednost ove distance nalazi se izvan dogovornih graničnih vrijednosti, pa je potrebno razmotriti opciju izostavljanja ovog podatka iz analize. Podatci o broju zrna i broju mahuna za 512. biljku mogu se ispisati na ovaj način:

```
soja[512,]
  broj.mahuna broj.zrna
512          72       51
```

Iako se ovaj podatak doima nelogičnim, jer je broj mahuna veći od broja zrna, nije sasvim nemoguće da je na ovoj biljci velik broj mahuna bio potpuno prazan. No, razlog za izostavljanje nekog opažanja nije točnost ili ispravnost podatka, nego upravo njegov utjecaj na rezultate analize, pa je i u ovom primjeru najbolje rješenje izostavljanje podataka za biljku br. 512, te ponovno analiziranje ovako reduciranog seta podataka.

### 3. Analiza reduciranog seta podataka

Reducirani se set podataka može spremi u novi objekt:

```
soja.corr<-soja[-512,]
```

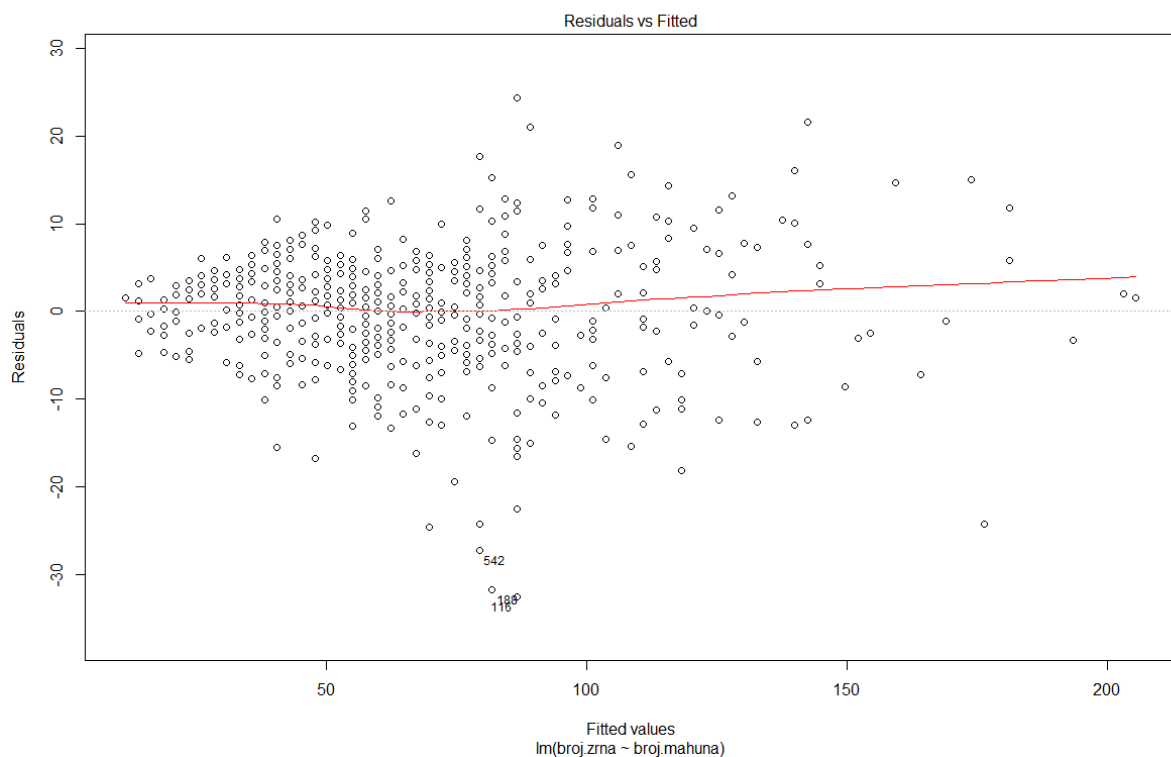
Zatim se ponovno provodi analiza prema istom modelu:

```
soja.lm2 <- lm(broj.zrna~broj.mahuna,data=soja.corr)
```

Prije pregleda rezultata, provjere se pretpostavke:

```
plot(soja.lm2)
```

Prvi dijagnostički grafikon sada izgleda ovako:



Slika 3. Odnos procjena i pogrešaka u regresijskoj analizi odnosa broja zrna i broja mahuna kod soje za reducirani set podataka

Na ovom, kao i na ostalim dijagnostičkim grafikonima koji nisu prikazani sada se više ne uočavaju atipični podatci koji bi mogli utjecati na rezultate analize. Procjene vrijednosti parametara modela sada iznose:

```
coefficients(soja.lm2)
(Intercept) broj.mahuna
-3.144028      2.425525
```

Prosječan broj zrna po mahuni sada je veći za 0.05 što i nije velika procjena u odnosu na prethodnu. Vrijednost konstantnog člana u oba slučaja je negativna, zbog čega se nema konkretnu biološku interpretaciju, nego jednostavno predstavlja korekciju procjene ukupnog broja zrna po biljci kao produkta broja mahuna i prosječnog broja zrna po mahuni. Promjena vrijednost konstantnog člana znatno je veća: u analizi izvornog seta podataka je ta procjena od produkta broja mahuna i prosječnog broja zrna po biljci bila manja za 1.85 zrna, a u analizi reduciranog seta je skoro dvostruko veća (3.14).

## Primjer 1.2

U istraživanju brzine rasta tjelesne mase purana bilježene su tjelesne mase životinja u početnim tjednima života (Janječić i sur. 2009). Priloženi podatci odnose se na mase (g) mužjaka crne rase u prvih trinaest tjedana života.

### 1. Učitavanje podataka i analiza rasta tjelesne mase

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
purani <- read.csv("Primjer1_2.csv")
```

Zatim se provodi regresijska analiza ovisnosti tjelesne mase o tjednu uzgoja, primjenom funkcije `lm`:

```
purani.lm1 <- lm(masa~tjedan,data=purani)
```

Procijenjene vrijednosti konstantnog člana (intercept) i regresijskog koeficijenta mogu se ispisati primjenom generičke funkcije `coefficients`:

```
coefficients(purani.lm1)
(Intercept)      tjedan
-1420.0482      455.4607
```

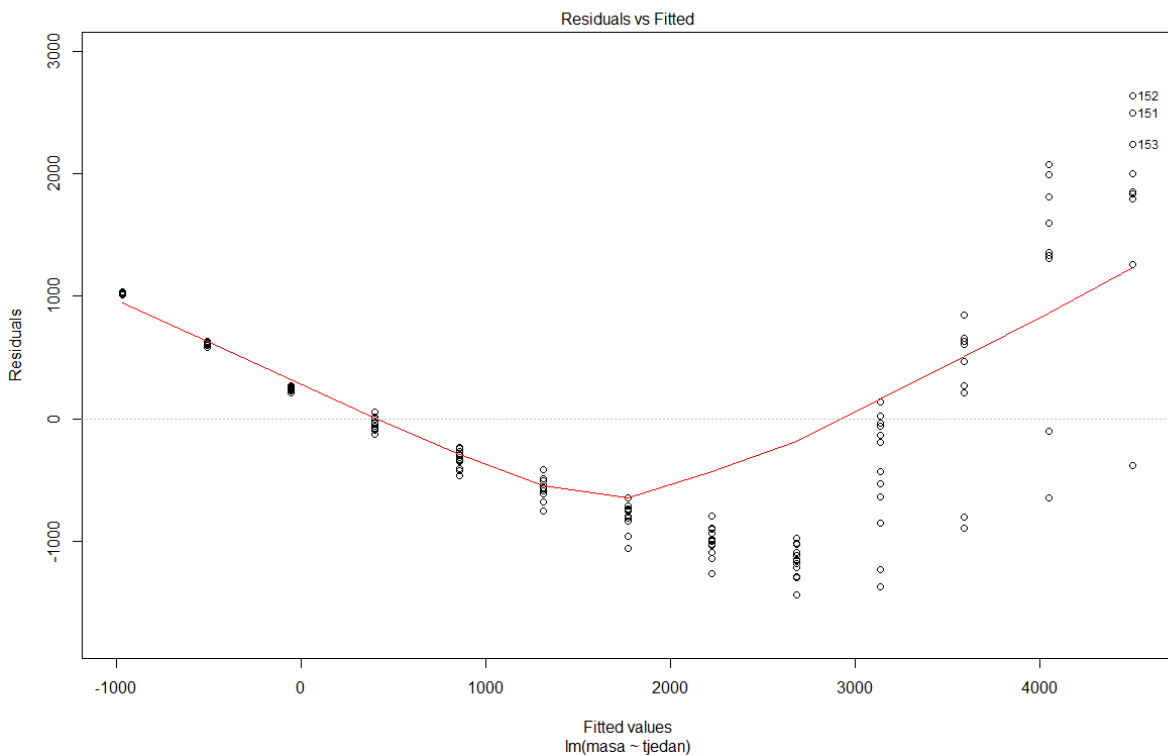
Na temelju procijenjene vrijednosti regresijskog koeficijenta može se zaključiti da prosječni tjedni prirast tjelesne mase purana iznosi 455.46 grama.

## 2. Provjera pretpostavki

Dijagnostički grafikoni generiraju se primjenom generičke funkcije `plot`:

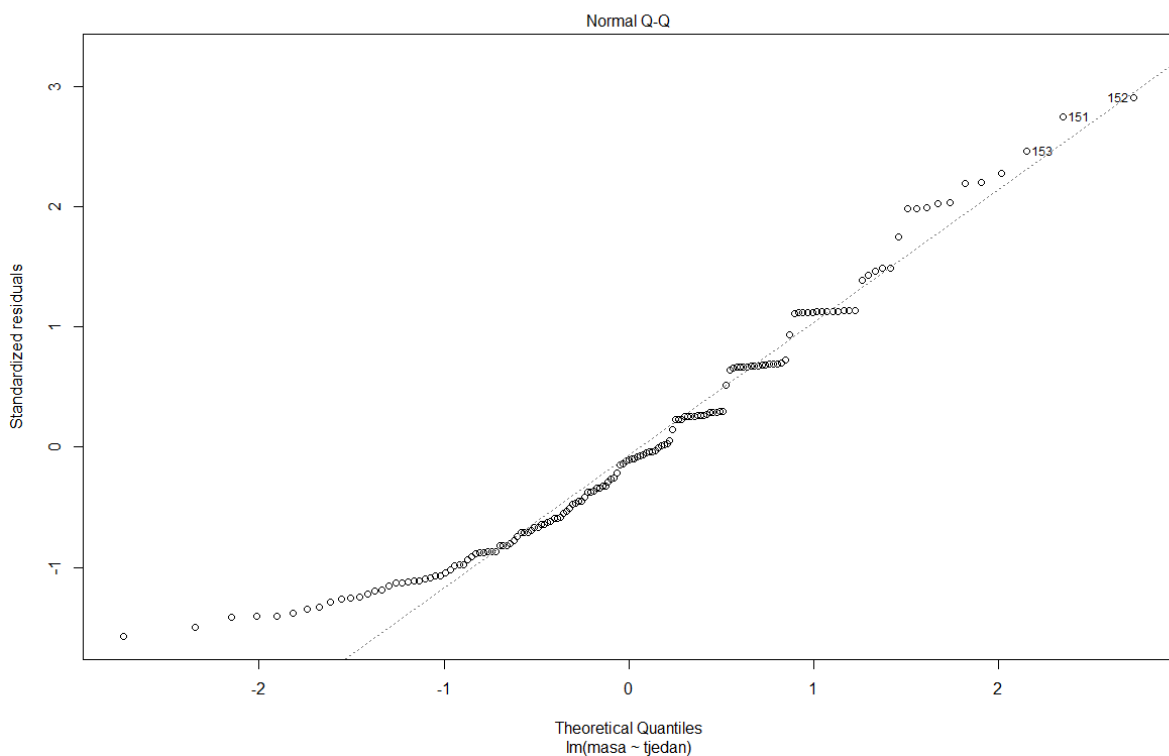
```
plot(purani.lm1)
```

U ovom će se primjeru narušavanje pretpostavki prikazati na prvom i drugom od četiri dijagnostička grafikona. Prvi grafikon, koji prikazuje odnos procjena i pogrešaka izgleda ovako:



Slika 4. Odnos procjena i pogrešaka u regresijskoj analizi rasta tjelesne mase purana

Na ovom se grafikonu može uočiti narušavanje pretpostavki koje se manifestira na dva načina. Varijance reziduala rastu iz tjedna u tjedan, pa je očito narušena treća pretpostavka o homogenosti varijanci. Ta se pretpostavka može provjeriti pomoću trećeg grafikona koji prikazuje disperziju reziduala, koji se ovdje neće prikazati, jer je vrlo sličan prvom. Drugi potencijalni problem je nelinearnost odnosa duljine uzgoja i tjelesne mase, zbog čega treba provjeriti i pretpostavku o normalnosti distribucije reziduala. Ta se pretpostavka provjerava pomoću drugog grafikona koji prikazuje distribuciju reziduala (normal probability plot ili Q-Q plot):



Slika 5. Distribucija pogrešaka u regresijskoj analizi rasta tjelesne mase purana

Na grafikonu se uočava odstupanje od normalne distribucije u većem broju točaka. Značajnost uočenog odstupanja može se provjeriti Shapiro-Wilkovim testom normalnosti:

```
shapiro.test(residuals(purani.lm1))
```

```
Shapiro-wilk normality test
```

```
data: residuals(purani.lm1)
W = 0.95029, p-value = 2.025e-05
```

Test je signifikantan, što potvrđuje značajno odstupanje od normalne distribucije, pa je potrebno pronaći odgovarajuću transformaciju izvornih podataka, kojom bi se postiglo ispunjenje pretpostavki za analizu.

### *3. Pronalaženje odgovarajuće transformacije*

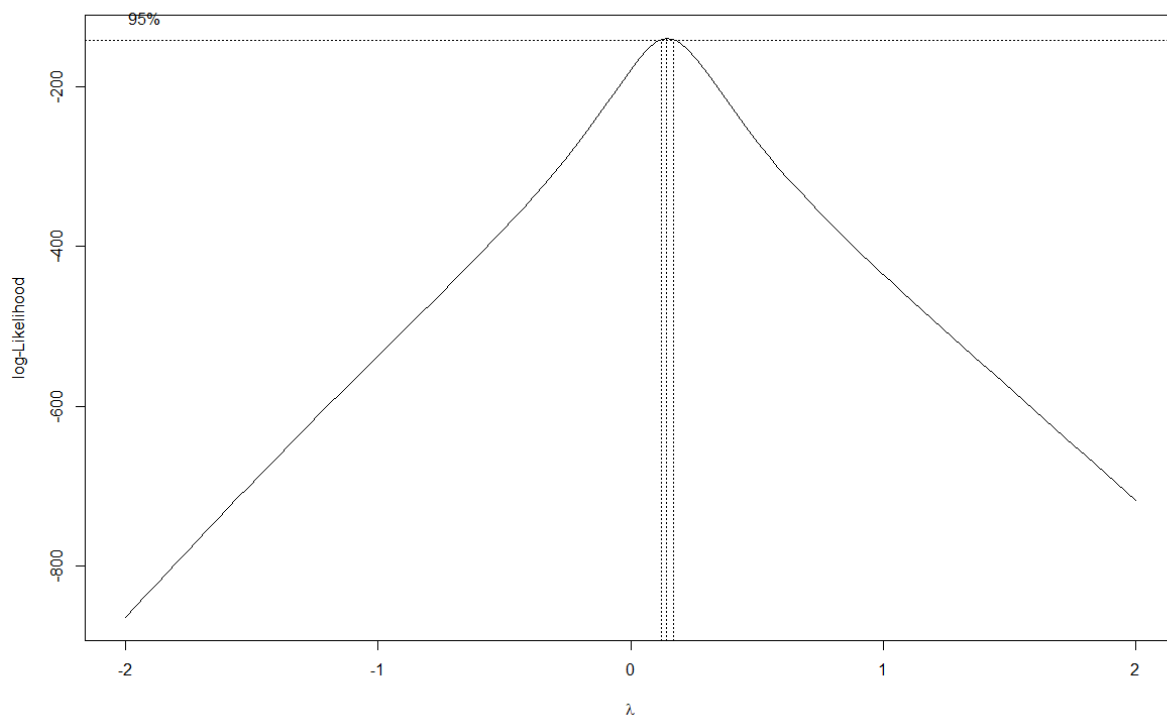
Odgovarajuća se transformacija može pronaći primjenom metode čiji su autori George Box i David Cox, primjenom funkcije `boxcox` iz paketa `MASS`. Najprije je potrebno aktivirati paket `MASS`:

```
library(MASS)
```

Zatim se na objekt u kojem su spremljeni rezultati analize primjenjuje funkcija `boxcox`:

```
boxcox(purani.lm1)
```

Rezultat primjene ove funkcije je grafikon s logaritmima vjerodostojnosti (log Likelihood) za analize prema transformiranim vrijednostima, koje nastaju potenciranjem vrijednosti izvorne varijable na parametar  $\lambda$ :



Slika 6. Vjerodostojnosti transformacija uz različite vrijednosti parametra  $\lambda$  za analizu prirasta mase purana

U načelu se za transformacije uzimaju vrijednosti  $\lambda$  koje su cijeli brojevi ili cijeli brojevi i polovine najbliže području najveće vjerodostojnosti. U ovom primjeru najveću vjerodostojnost imaju transformacije za vrijednost parametra  $\lambda$  koji je vrlo blizu nuli, pa se transformacija izvornih vrijednosti tjelesne mase može provesti za  $\lambda = 0$ . Prema Tukeyjevim ljestvama, odnosno Boxovoj i Coxovoj metodi, ova je vrijednost jedina iznimka u smislu da ne predstavlja vrijednost potencije (što bi sve vrijednosti izvorne varijable pretvorilo u jedinice), nego podrazumijeva njihovo logaritmiranje.

#### 4. Analiza transformiranih vrijednosti zavisne varijable

Ponovna se analiza provodi tako da se u model umjesto izvornih uvrste transformirane vrijednosti zavisne varijable:

```
purani.lm2 <- lm(log10(masa)~tjedan,data=purani)
```



Slijede li reziduali u ovoj analizi normalnu distribuciju može se provjeriti Shapiro-Wilkovim testom:

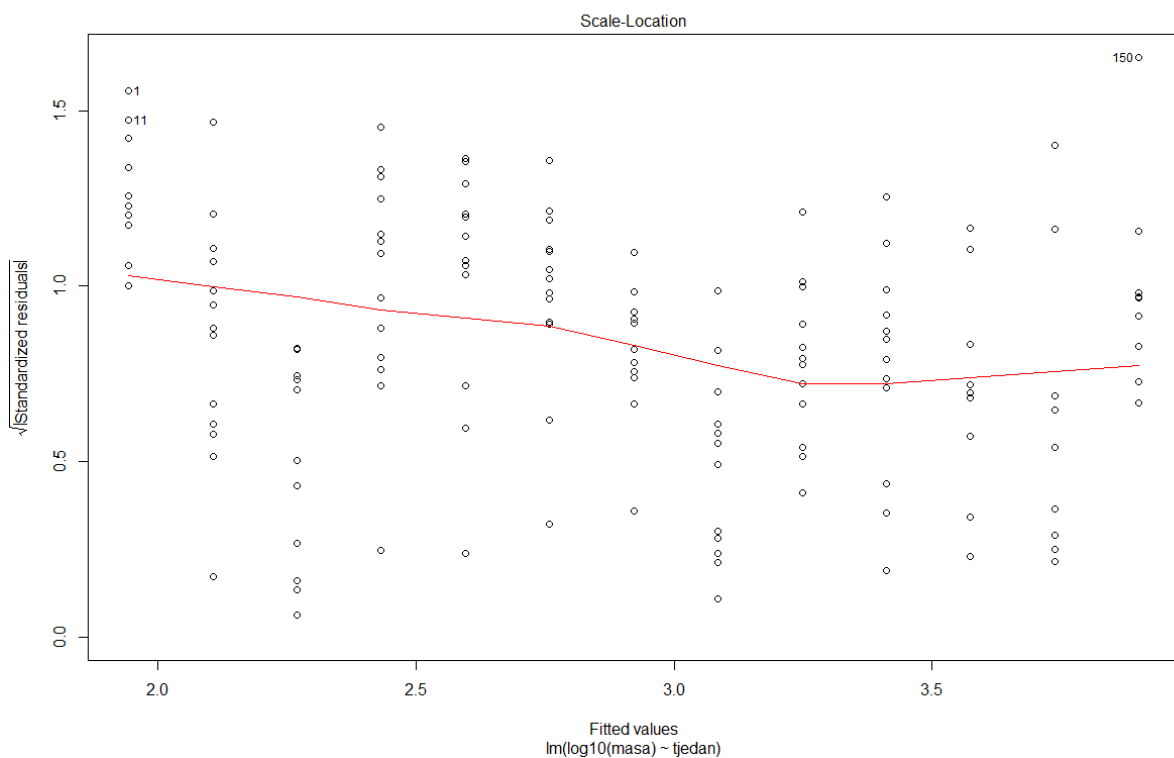
```
shapiro.test(residuals(purani.lm2))
```

Shapiro-wilk normality test

```
data: residuals(purani.lm2)  
W = 0.98734, p-value = 0.1601
```

Test sada nije signifikantan, što znači da je logaritmiranjem izvornih podataka ispunjena ova pretpostavka. Homogenost varijanci se može provjeriti na trećem dijagnostičkom grafikonu koji prikazuje disperziju reziduala:

```
plot(purani.lm2)
```



Slika 7. Disperzija pogrešaka u analizi prirasta logaritama tjelesnih masa purana

Potpunu jednakost varijanci označavala bi ravna crvena crta u sredini grafikona. U ovom je primjeru ona blago zakrivljena, pa iako pretpostavka o homogenosti varijanci nije u potpunosti zadovoljena, ova se transformacija ipak može smatrati dovoljno

dobrim rješenjem u ovom primjeru. Procijenjene vrijednosti konstantnog člana i regresijskog koeficijenta u analizi transformiranih podataka su:

```
coefficients(purani.lm2)
(Intercept)      tjedan
    1.780809     0.163037
```

Obzirom na to da su analizirane logaritmirane vrijednosti zavisne varijable (tjelesne mase) vrijednost regresijskog koeficijenta predstavlja postotni prirast vrijednosti zavisne varijable uz jediničnu promjenu nezavisne, što u ovom primjeru znači da tjedni prirast tjelesne mase purana iznosi 16.3% (u odnosu na vrijednost iz prethodnog tjedna).

## Primjer 1.3

U istraživanju utjecaja tri različite gnojidbe na svojstva pelargonija prikupljeni su podatci o nekoliko gospodarski važnih svojstava. U ovom primjeru analizirati će se podatci o visini biljaka (po deset biljaka za svaku gnojidbu).

### *1. Učitavanje podataka i analiza učinaka gnojidbi*

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
pe1 <- read.csv("Primjer1_3.csv")
```

Zatim se provodi analiza varijance za usporedbu učinaka gnojidbi na visinu biljke, primjenom funkcije `lm`:

```
pe1.lm1 <- lm(visina~gnojidba,data=pe1)
```

Tablica analize varijance može se ispisati primjenom generičke funkcije `anova`:

```
anova(pe1.lm1)
Analysis of Variance Table
```

Response: visina

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gnojidba	2	303.8	151.900	8.1748	0.001675 **
Residuals	27	501.7	18.581		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

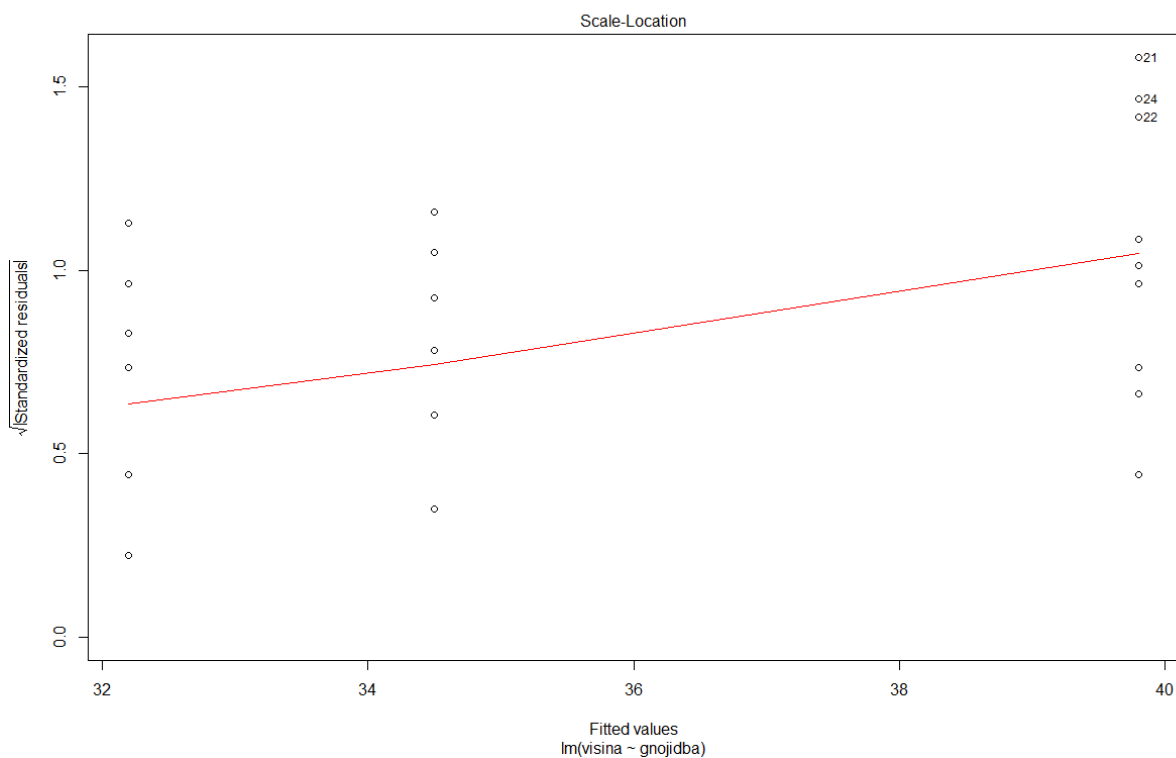
Na temelju ishoda F-testa može se zaključiti da se gnojidbe značajno razlikuju u učincima na visinu biljke. Ovaj se zaključak ne može smatrati valjanim prije provjere pretpostavki.

## 2. Provjera pretpostavki

Dijagnostički grafikoni generiraju se primjenom generičke funkcije `plot`:

```
plot(pe1.lm1)
```

U ovom će primjeru biti prikazan samo treći od četiri dijagnostička grafikona:



Slika 8. Disperzija pogrešaka u analizi učinaka gnojidbi na visinu pelargonija

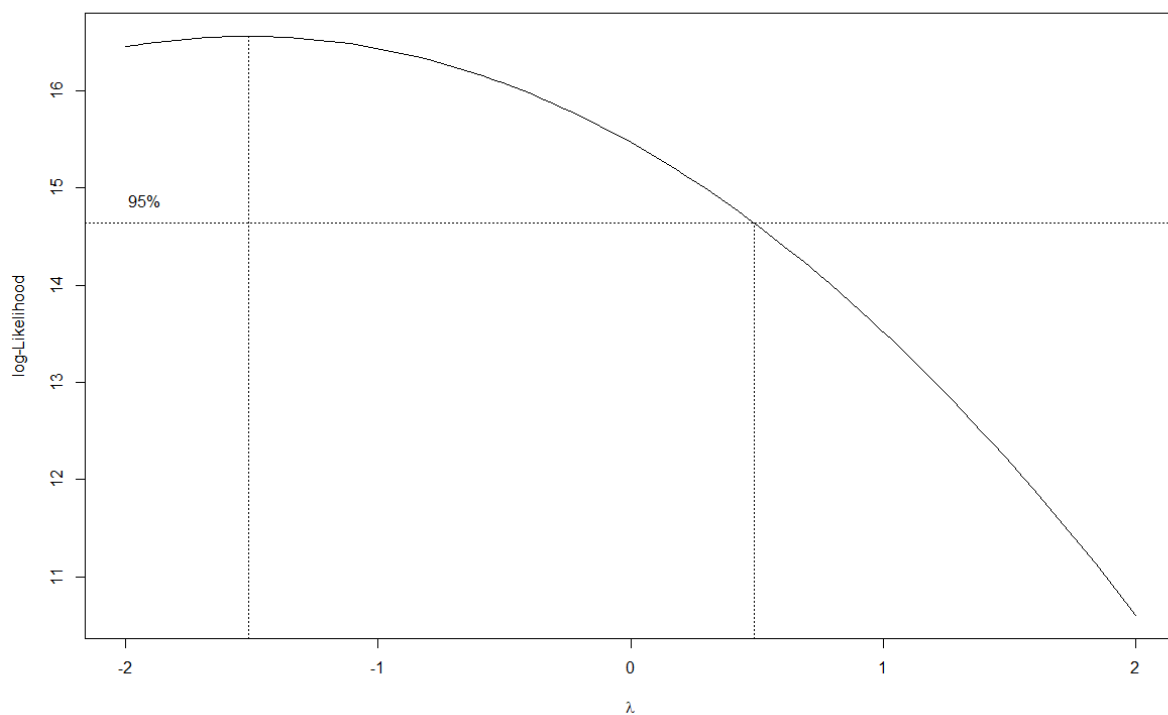
Za model analize varijance (zavisna varijabla kontinuirana, nezavisna kategorijska) karakteristično je da je broj procjena jednak broju kategorija, odnosno, procjene su za sve jednako tretirane varijante iste, jer su jednake prosjeku kategorije (tretiranja). U ovom primjeru su procjene prosječne visine jednako tretiranih biljaka, pa su pogreške grupirane u tri stupca. Nagib crvene crte nije promjenjiv kao u prethodnom primjeru, nego je konstantan, što ukazuje na nehomogenost varijanci pri kojoj je trend rasta varijanci pogrešaka proporcionalan trendu rasta vrijednosti procjena. U ovakvim se slučajevima primjenjuju transformacije za stabilizaciju varijanci, a vrijednost parametra  $\lambda$  trebala bi odgovarati nagibu crvene crte. Ta se vrijednost može pronaći primjenom funkcije `boxcox`.

### *3. Pronalaženje odgovarajuće transformacije*

Nakon aktiviranja paketa `MASS`, primjenom funkcije `boxcox` na objekt u kojem su spremljeni rezultati inicijalne analize traži se odgovarajuća transformacija:

```
library(MASS)
boxcox(pe1.lm1)
```

Rezultat primjene funkcije je grafikon s logaritmima vjerodostojnosti za različite vrijednosti parametra  $\lambda$ :



Slika 9. Vjerodostojnosti transformacija uz različite vrijednosti parametra  $\lambda$  za analizu učinaka gnojidbi na visinu pelargonija

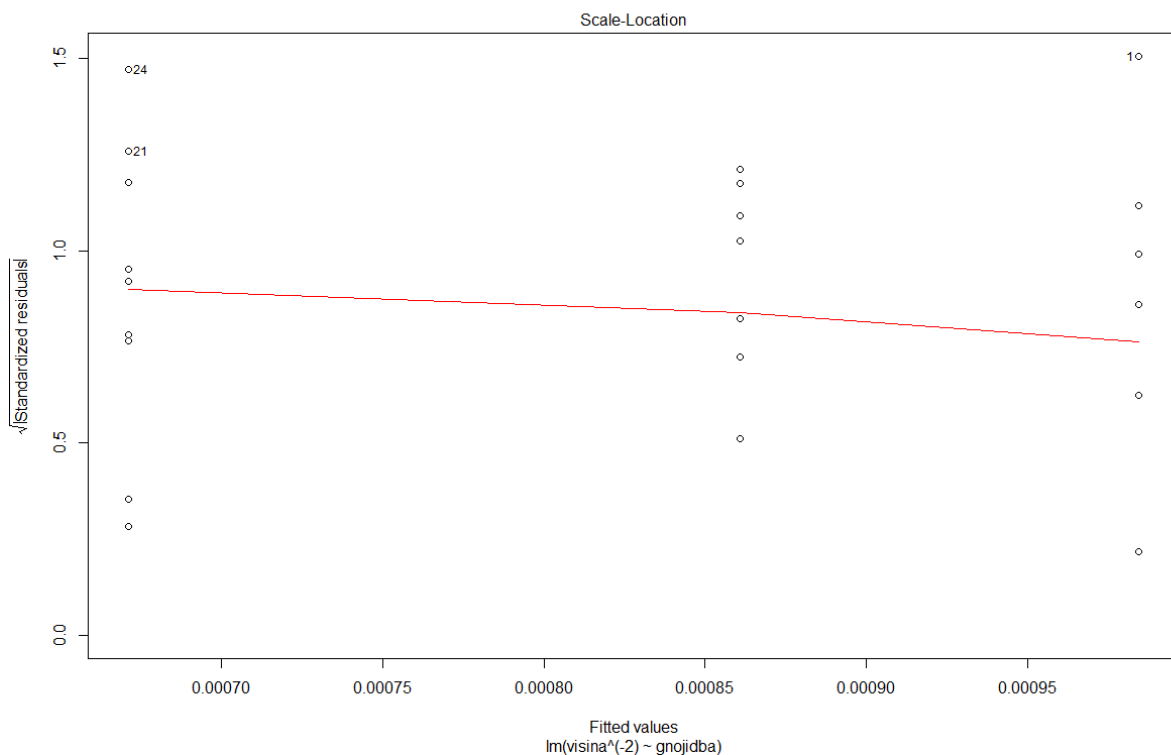
Područje najveće vjerodostojnosti obuhvaća relativno široki interval vrijednosti parametra  $\lambda$ . Primjena bilo koje vrijednosti iz tog intervala rezultirati će zadovoljavajućim rješenjem, a u ovom će se primjeru primijeniti transformacija za vrijednost  $\lambda = -2$ .

#### 4. Analiza transformiranih vrijednosti zavisne varijable

Ponovna se analiza provodi tako da se u model umjesto izvornih uvrste transformirane vrijednosti zavisne varijable:

```
pel.lm2 <- lm(visina**(-2)~gnojidba,data=pel)
```

Treći dijagnostički grafikon sada izgleda ovako:



Slika 10. Disperzija pogrešaka u analizi učinaka gnojidbi na transformirane vrijednosti visina pelargonija

Proporcionalan odnos varijanci pogrešaka i veličine procjena sa Slike 8. na ovom se grafikonu više ne može uočiti. Idealna ujednačenost varijanci nije postignuta, ali se može smatrati da je primijenjena transformacija ( $1/\text{visina}^2$ ) uspješno stabilizirala varijance. Rezultat ponovljene analize na transformiranim podacima je:

`anova(pe1.lm2)`

Analysis of Variance Table

Response: visina<sup>(-2)</sup>

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gnojidba	2	4.9827e-07	2.4913e-07	7.6877	0.002277 **
Residuals	27	8.7499e-07	3.2407e-08		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dakle, razlike u visinama pelargonija postignute primjenom različitih gnojidbi su značajne.

## Literatura

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.
2. Shapiro S.S., Wilk M.B. (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52: 591–611.
3. Tukey J.W. (1977) Exploratory Data Analysis. Addison-Wesley, Reading, MA.
4. Box G.E.P., Cox D.R. (1964) An analysis of transformations. *Journal of the Royal Statistical Society B* 26: 211-252.
5. Janječić Z., Mužić S., Bedeković D., Gunjača J., Pintar J., Kos I. (2009) Production results and fatty acid composition in breast muscle of Zagorje turkeys on free range rearing. Akbay, R. (ur.): Proceedings of 2nd Mediterranean Summit of WPSA, Antalya, Turkey. pp. 375-378.
6. Cook R.D. (1977) Detection of Influential Observations in Linear Regression. *Technometrics* 19: 15-18.
7. Venables W.N., Ripley B.D. (2002) Modern Applied Statistics with S. Springer, New York.

## Višestruke usporedbe

Kada je cilj istraživanja ispitati razlike između učinaka više od dvaju tretiranja, statistička se analiza sastoji od niza t-testova. Kako ti t-testovi međusobno nisu neovisni, s povećanjem broja testova povećava se i rizik od počinjenja pogreške tipa I (proglašavanja učinaka tretiranja različitima kada oni to nisu). Ako je vjerojatnost pogreške tipa I u svakom pojedinačnom testu jednaka nekoj vrijednosti  $\alpha_C$ , onda se rizik ( $\alpha_E$ ) da će se pri provedbi n testova za višestruke usporedbe počiniti barem jedna pogreška tipa I može procijeniti pomoću ovog izraza:

$$\alpha_E = 1 - (1 - \alpha_C)^n \quad (3)$$

Problem povećanja rizika pogreške tipa I može se ublažiti nekom od velikog broja metoda korekcije višestrukih testova koje su razvili različiti autori, od kojih će ovdje biti navedeno samo nekoliko.

**Fisherov LSD test** – je metoda u kojoj se ne primijenjuje nikakva korekcija testova, nego se kontrola pogreške osigurava uvjetom da se test provodi samo za one efekte za koje je u prethodnoj analizi varijance F-test bio signifikantan. Ovo je test s najmanjom strogošću u odnosu na sve ostale, odnosno, uz njegovu će primjenu najveći broj testova biti signifikantan.

**Bonferronijeva korekcija** – zasniva se na ograničavanju rizika činjenja barem jedne pogreške tipa I provođenjem strožih pojedinačnih testova. Ukoliko se rizik činjenja pogreške želi ograničiti na neku zadanu vrijednost  $\alpha_E$ , onda se pogreška  $\alpha_C$  u svakom od n pojedinačnih testova mora svesti na vrijednost  $\alpha_E / n$ . Ovo je najstroži mogući test, te će uz njegovu primjenu najmanji broj testova biti signifikantan.

**False discovery rate (FDR)** – skupni je naziv za nekoliko metoda kontrole pogreške tipa I, koje su umjesto na ograničavanje rizika činjenja barem jedne pogreške tipa I usmjerene na ograničavanje udjela pogrešnih odluka („stope lažnih otkrića“). Ove se metode odlikuju nešto manjom strogošću, a uslijed toga i većom snagom testova.

**Dunnettov test** – je metoda koja se koristi za analizu pokusa u kojima je cilj usporediti sva učinak svih preostalih u odnosu na jedno kontrolno tretiranje.



## Primjer 2.

U uzgoju vinove loze vrlo važnu ulogu ima kompatibilnost plemke i podloge. Dobra, odnosno loša kompatibilnost može, pored ostalog, utjecati i na kemijski sastav grožđa i vina, a time i na njihovu kvalitetu. U istraživanju kompatibilnosti kultivara Chardonnay s različitim podlogama, plemke Chardonnaya cijepljene su na četiri podloge: 41B, Fercal, Kober 5BB i Paulsen 1103 (Maletić 1999). U ovom primjeru analizirati će se utjecaj podloga na sadržaj ukupnih kiselina (g/l) u 1997. godini.

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
pr.2 <- read.csv("Primjer2.csv")
```

### 2. Analiza varijance

Zatim se provodi analiza varijance za usporedbu učinaka podloga na sadržaj ukupnih kiselina, primjenom funkcije `lm`:

```
pr2.anova <- lm(ukk~podloga,data=pr.2)
```

Tablica analize varijance može se ispisati primjenom generičke funkcije `anova`:

```
anova(pr2.anova)
```

Analysis of Variance Table

Response: ukk

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
podloga	3	2.7219	0.90729	4.4575	0.02528 *
Residuals	12	2.4425	0.20354		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Na temelju ishoda F-testa može se općenito zaključiti da se podloge značajno razlikuju u učincima na sadržaj ukupnih kiselina, odnosno da je bar jedna od razlika

između podloga signifikantan. Za konkretne zaključke o razlikama između sve četiri podloge potrebno je provesti višestruke usporedbe.

### 3. Višestruke usporedbe

U R-u se postupak višestrukih usporedbi može provesti primjenom nekoliko funkcija iz paketa `multcomp`, koji je najprije potrebno aktivirati (a prethodno instalirati):

```
library(multcomp)
```

U prvom se koraku primijenjuje funkcija `glht`, uz pomoć koje se za sve planirane usporedbe izračunaju razlike i njihove standardne pogreške, koje se spremne u novi objekt:

```
pr2.glht <- glht(pr2.anova, linfct=mcp(podloga="Tukey"))
```

Prvi argument funkcije `glht` je objekt u kojem su spremljeni rezultati analize varijance, na temelju kojih se procjenjuju prosjeci podloga, njihove razlike i pogreške razlika. Drugim se argumentom (`linfct`) zadaju se usporedbe koje će se provesti, koje se mogu definirati pomoću matrice koeficijenata, pomoću simboličkih opisa ili pomoću funkcije `mcp`. U ovom se primjeru koristi funkcija `mcp`, čiji je jedini argument efekt u modelu za koji se provode usporedbe (`podloga`), s ključnom riječi `Tukey` kojom se generira matrica koeficijenata za usporedbe svih podloga međusobno. Alternativno, ova se matrica može generirati i izravno, upisom opisa usporedbi i koeficijenata u odgovarajuće retke i stupce matrice, te spremiti u novi objekt `cm`:

```
cm <- rbind("Fercal - 41B" = c(-1,1,0,0),  
           "Kober - 41B" = c(-1,0,1,0),  
           "Paulsen - 41B" = c(-1,0,0,1),  
           "Kober - Fercal" = c(0,-1,1,0),  
           "Paulsen - Fercal" = c(0,-1,0,1),  
           "Paulsen - Kober" = c(0,0,-1,1))
```

Umjesto ključne riječi `Tukey` sada se za definiranje istih usporedbi može koristiti matrica `cm`, ali je naravno primjena ključne riječi znatno jednostavniji i brži način za definiranje usporedbi svih podloga međusobno. Rezultati testova se mogu ispisati primjenom generičke funkcije `summary`, čiji je prvi argument ime `glht` objekta u kojem su spremljene nazivi usporedbi, razlike i pogreške razlika. Drugi argument određuje

metodu korekcije za višestruke testove, pa se primjenom ključne riječi `none` provodi Fisherov LSD test:

```
summary(pr2.glm, test=adjusted("none"))
```

```
Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: lm(formula = ukk ~ podloga, data = pr.2)

Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
Fercal - 41B == 0    -0.950     0.319  -2.978  0.01153 *
Kober - 41B == 0    -1.050     0.319  -3.291  0.00644 **
Paulsen - 41B == 0   -0.775     0.319  -2.429  0.03177 *
Kober - Fercal == 0   -0.100     0.319  -0.313  0.75931
Paulsen - Fercal == 0  0.175     0.319   0.549  0.59336
Paulsen - Kober == 0  0.275     0.319   0.862  0.40557
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- none method)
```

Jedine signifikantne razlike su između podloge 41B i tri preostale podloge, a prema negativnom se predznaku uz procjene tih razlika može zaključiti da je sadržaj ukupnih kiselina kultivara Chardonnay na podlozi 41B viši nego na tri preostale podloge, između kojih nema značajnih razlika. Iako zbog malog broja usporedbi korekcija nije nužno potrebna, usporedbe radi, zamjenom ključne riječi `none` ključnom riječi `bon` mogu se provesti testovi uz Bonferronijevu metodu korekcije:

```
summary(pr2.glm, test=adjusted("bon"))
```

```
Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: lm(formula = ukk ~ podloga, data = pr.2)

Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
Fercal - 41B == 0    -0.950     0.319  -2.978  0.0692 .
Kober - 41B == 0    -1.050     0.319  -3.291  0.0387 *
Paulsen - 41B == 0   -0.775     0.319  -2.429  0.1906
Kober - Fercal == 0   -0.100     0.319  -0.313  1.0000
Paulsen - Fercal == 0  0.175     0.319   0.549  1.0000
Paulsen - Kober == 0  0.275     0.319   0.862  1.0000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- bonferroni method)
```

Nakon primjene najstrože varijante korekcije preostaje samo jedna signifikantna razlika, i to između podloga Kober i 41B. FDR metoda manje je stroga od Bonferronijeve, što se u ovom primjeru može provjeriti primjenom ključne riječi **BH**, kojom se provode testovi prema Benjamini-Hochbergovoj metodi:

```
summary(pr2.glm, test=adjusted("BH"))
```

```

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: lm(formula = ukk ~ podloga, data = pr.2)

Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
Fercal - 41B == 0      -0.950     0.319  -2.978  0.0346 *
Kober - 41B == 0      -1.050     0.319  -3.291  0.0346 *
Paulsen - 41B == 0    -0.775     0.319  -2.429  0.0635 .
Kober - Fercal == 0   -0.100     0.319  -0.313  0.7593
Paulsen - Fercal == 0  0.175     0.319   0.549  0.7120
Paulsen - Kober == 0  0.275     0.319   0.862  0.6084
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- BH method)

```

Primjenom ove metode signifikantni su ishodi dvaju testova, točno jednog više nego kod najstrože (Bonferroni), odnosno točno jednog manje nego kod najliberalnije metode (Fisherov LSD).

Osim ovakvih tabličnih ispisa, rezultati testova mogu se prikazati i u obliku grafikona, što ih čini znatno preglednije. Najprije će se zbog praktičnosti rezultati Fisherovog LSD testa spremi u novi objekt:

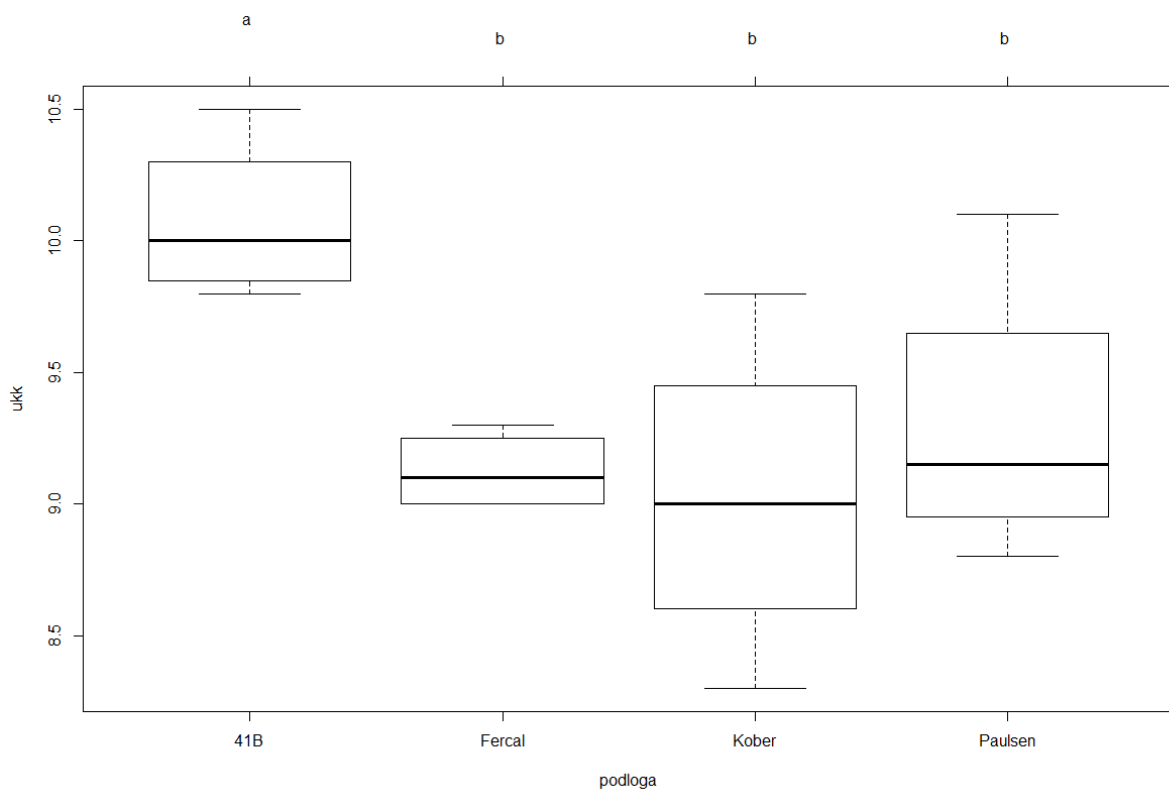
```
pr2.mcp <- summary(pr2.glm, test=adjusted("none"))
```

Vrlo pregledan način prikaza značajnih razlika između prosjeka postiže se dodjeljivanjem slovnih oznaka prema načelu da se prosjeci označeni istim slovom ne razlikuju, a prosjeci označeni različitim slovima razlikuju. Paket **multcomp** sadrži i funkciju **clld** koja služi upravo za dodjeljivanje slovnih oznaka (cld = „compact letter display“). Ta se funkcija primijeni na objekt u kojem su spremljeni rezultati testova (argumentom **decreasing=T** se slova dodjeljuju prosjecima od najvećeg prema najmanjem):

```
cld(pr2.mcp,decreasing=T)
  41B Fercal Kober Paulsen
  "a"  "b"  "b"  "b"
```

Na temelju ovakvog prikaza razlika između znatno je lakše doći do prethodnog zaključka da se podloga 41B razlikuje od ostalih, koje se ne razlikuju međusobno. Primjenom generičke funkcije `plot` na rezultate funkcije `cld`, slova će se prikazati uklopiti u grafikon distribucija sadržaja kiselina po podlogama (box-plot):

```
plot(cld(pr2.mcp,decreasing=T))
```



Slika 11. Grafički prikaz Fisherovog LSD testa razlika između podloga u sadržaju ukupnih kiselina (podloge označene istim slovom se ne razlikuju)

## Literatura

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.

2. Maletić E. (1999) Utjecaj podloga na biološka i gospodarska svojstva kultivara Chardonnay (*Vitis vinifera* L.) (doktorska disertacija). Sveučilište u Zagrebu, Agronomski fakultet.
3. Bretz F., Hothorn T., Westfall P. (2010) Multiple Comparisons Using R. Chapman and Hall/CRC, Boca Raton, FL.

## Višefaktorijski pokusi

Osnovni model analize varijance podrazumijeva da na zavisnu (kontinuiranu) varijablu djeluje jedna nezavisna (kategorijska) varijabla, odnosno faktor. U višefaktorijskim se pokusima istražuje djelovanje dvaju ili više faktora (kategorijskih varijabli) na zavisnu varijablu. U ovom se kontekstu kategorije nazivaju razinama faktora, a tretiranja predstavljaju kombinacije razina faktora. Na temelju međusobnog odnosa faktora razlikuju se dva glavna tipa modela za analizu višefaktorijskih pokusa: unakrsni i hijerarhijski model.

Unakrsni model pretpostavlja ravnopravan odnos faktora u kojem se svaka razina jednog faktora može kombinirati s bilo kojom razinom drugog faktora. Ukupni se učinak tretiranja u ovom modelu rastavlja na pojedinačne učinke faktora (glavne efekte) i učinke njihovih interakcija. Prema tome, model za analizu pokusa u kojem se istražuju učinci dvaju faktora A i B na zavisnu varijablu izgleda ovako:

$$Y = A + B + A \cdot B \quad (4)$$

U hijerarhijskom modelu razine „podređenog“ faktora su „ugniježdene“ (nested) u razinama „nadređenog“ faktora, odnosno podijeljene su u skupine od kojih se svaka kombinira samo s jednom razinom „nadređenog“ faktora. Ukupni se učinak tretiranja u ovom modelu rastavlja na učinak nadređenog faktora i učinak podređenog faktora ugniježđenog u nadređenom faktoru. Osnovna razlika u zapisu u odnosu na unakrsni model je u tome što je ovdje ispušten samostalni učinak „podređenog“ faktora:

$$Y = A + A \cdot B \quad (5)$$

Osim na temelju odnosa faktora, modeli za analizu višefaktorijskih pokusa mogu se razlikovati i prema tipu učinaka, odnosno načinu na koji se tretiraju učinci u modelu. Ako su predmet istraživanja samo one razine faktora koje se pojavljuju u pokusu onda se učinci tretiraju kao fiksni, a ako razine faktora koje se pojavljuju u pokusu predstavljaju samo slučajni uzorak iz populacije koja je predmet istraživanja, onda se učinci tretiraju kao slučajni. Analiza i interpretacija rezultata analize su za fiksne učinke usmjerene na njihove prosječne vrijednosti, a za slučajne učinke na

njihove varijance. Model u kojem se pojavljuju samo slučajni efekti naziva se slučajnim, a model koji sadrži i fiksne i slučajne efekte naziva se mješovitim (mixed). U fiksnim se modelima ipak pojavljuje jedan slučajni efekt – učinak pogreške.

U praktičnom smislu opisane razlike između modela najviše se odražavaju na formiranje F-testova u analizi varijance, odnosno na odabir djelitelja (nazivnika) za testiranje pojedinih učinaka modela. U tom se kontekstu može navesti nekoliko načelnih pravila:

1. F-testovi se provode samo za fiksne efekte, a slučajni efekti se mogu koristiti kao samo kao djelitelji;
2. u fiksnom se modelu testovi za sve efekte provode u odnosu na pogrešku;
3. ukoliko je bilo koji od efekata uključenih u interakciju slučajan, onda i interakcijski efekt mora biti slučajan;
4. ako je interakcijski efekt slučajan, onda se on koristi za test svih efekata uključenih u interakciju (bez obzira na to jesu li fiksni ili slučajni);
5. u hijerarhijskom se modelu F-test za nadređeni faktor provodi u odnosu na u njemu ugniježđeni faktor.

Slijedeći navedena pravila, u dvofaktorijelnom pokusu s jednim fiksnim i jednim slučajnim efektom, učinak interakcije će se također tretirati kao slučajni efekt. U sustavu zapisa koji se koristi u ovom priručniku, u zapisu modelu se najprije se navode fiksni efekti, a zatim slučajni, a za oznaku kojom se međusobno razdvajaju izvorno je predložen simbol „:“ (Piepho i sur. 2003). Pošto se u R-u taj simbol koristi za označavanje interakcije, kako bi se izbjegla mogućnost zabune, u ovom će se priručniku za razdvajanje fiksnih i slučajnih efekata u mješovitom modelu koristiti oznaka „::“. Zato model za opisani dvofaktorijelni pokus izgleda ovako:

$$Y = A :: B + A \cdot B \quad (6)$$

U kompleksnijim modelima može se pojaviti i potreba za formiranjem nazivnika F omjera kao linearne kombinacije više efekata. Sve postaje još kompleksnije u situacijama kada dio podataka nedostaje (missing values). Primjeri koji se koriste u ovom priručniku biti će relativno jednostavni, pa će za njihovo rješavanje biti dovoljno slijediti navedena pravila.



## Primjer 3.1

U istraživanju utjecaja podloga na neka svojstva kultivara vinove loze Sauvignon bijeli (Vaupotič, 2008), proveden je višegodišnji pokus u kojem je kultivar Sauvignon uzgojen na devet različitih podloga: 196/17, Boerner, V. Riparia (klon), te dva klona podloge Kober 5BB i četiri klona podloge SO4. U ovom primjeru analizirati će se podatci o urodu grožđa (kg po trsu), za 30 trsova po podlozi, iz dvije uzastopne godine (2000. i 2001.).

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
vin <-  
  read.csv("Primjer3_1.csv", colClasses = c("factor", "factor", "numeric"))
```

Prve dvije varijable su godine i podloge, pa se zbog toga što su godine označene brojevima pri učitavanju koristi argument `colClasses` kako se ta varijabla ne bi učitala kao numerička, nego kao faktor.

### 2. Analiza varijance

U ovom se primjeru radi o unakrsnom modelu, pa se provodi analiza učinaka godina, podloga i njihove interakcije na urod grožđa, primjenom funkcije `lm`:

```
vin.lm1 <- lm(urod~godina*podloga,data=vin)
```

Simbol `*` u sintaksi R-a se koristi za skraćeni način definiranja modela. U ovom primjeru `godina*podloga` se u analizi razvija u izraz `godina+podloga+godina:podloga`, pri čemu se simbol `:` koristi za definiranje interakcije faktora. Provjera pretpostavki (opisana u Primjerima 1.1, 1.2 i 1.3) otkrila bi jedan atipičan podatak. Taj se podatak nalazi u 390. retku objekta `vin`:

```
vin[390,]
      godina podloga urod
390    2001   SO4_31     0
```

Dakle, sa jednog od trsova na podlozi SO4 – klon 31 nije ubran niti jedan grozd, pa će se taj podatak izostaviti, a analiza ponovo provesti na reduciranom setu podataka:

```
vin.corr <- vin[-390,]
vin.lm2 <- lm(urod~godina*podloga,data=vin.corr)
```

### 3. Interpretacija rezultata

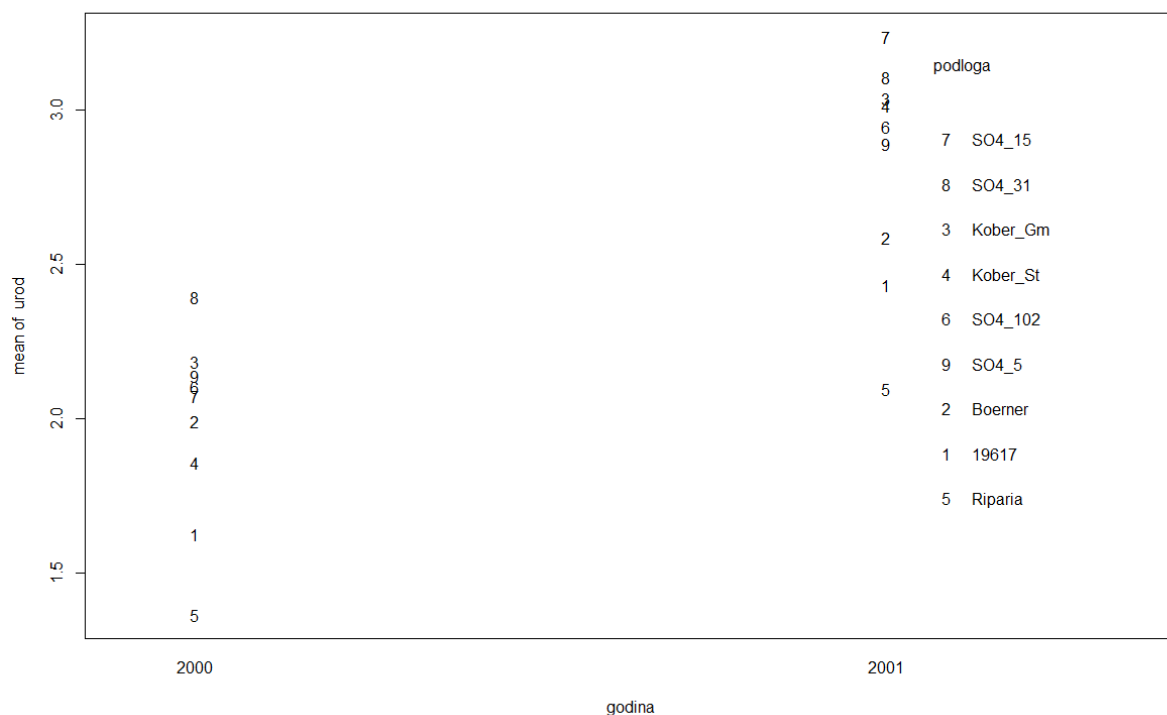
Tablica analize varijance ispiše se primjenom generičke funkcije `anova`:

```
anova(vin.lm2)
Analysis of Variance Table

Response: urod
      Df Sum Sq Mean Sq  F value    Pr(>F)
godina    1  96.44   96.440 111.5706 < 2.2e-16 ***
podloga    8   51.65    6.456   7.4692 1.891e-09 ***
godina:podloga  8    4.58    0.573    0.6624  0.7248
Residuals 521 450.34    0.864
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Prikazani F-testovi podrazumijevaju fiksni model, odnosno tretiranje učinaka obaju faktora i njihove interakcije kao fiksnih efekata. Testovi za oba faktora su signifikantni, što znači da postoje značajne razlike u urodu između godina, kao i između podloga. Učinak interakcije nije signifikantan, pa se može zaključiti da je odnos između podloga isti u obje godine, odnosno da je odnos između godina isti za sve podloge. Ove se interpretacije odnosa mogu provjeriti i grafičkim prikazom interakcije uz pomoć funkcije :

```
with(vin.corr, interaction.plot(godina, podloga, urod, type="p"))
```



Slika 12. Grafički prikaz učinka interakcije godina i podloga na urod grožđa kultivara Sauvignon

Na Slici 12. vidljivo je da je urod na svim podlogama bio veći u 2001. godini. Poredak je većine podloga isti u obje godine, a iznimke su podloge SO4 kl. 15 i Kober (St.) koje su napredovale za tri mjesta u poretku, te podloge SO4 kl. 5 i SO4 kl. 102 koje su zamijenile mjesta. Međutim, učinak je svih tih promjena u poretku očito zanemariv, jer razlike između podloga koje su zamijenile mjesta vjerojatno nisu značajne ni u jednoj ni u drugoj godini, na što ukazuje nesignifikantan test za interakciju u analizi varijance.

### Primjer 3.2

Istraživanje varijabilnosti svojstava lista i cvijeta divljih trešnja obuhvatilo je populacije s različitih lokaliteta u Bosni i Hercegovini i Hrvatskoj (Ballian i sur. 2012). U ovom će se primjeru analizirati podatci o širini lista populacija iz okolice Gospića, Karlovca i Tomislavgrada. Svaka je populacija predstavljena sa pet stabala, a svako stablo sa pet listova.

## 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
dt <-  
  read.csv("Primjer3_2.csv", colClasses = c("factor", "factor", "numeric"))
```

Argument `colClasses` koristi se kako bi se varijabla `stablo` učitala kao faktor, jer su stabla označena brojevima.

## 2. Analiza varijance

Faktor `stablo` ugniježđen je u faktoru `lokalitet`, jer se naravno radi o pet različitih stabala na svakom lokalitetu, odnosno ukupno 15 stabala. Zbog toga se analiza provodi prema hijerarhijskom modelu, uz uporabu specijalne oznake `%in%` koja se koristi za označavanje ugniježđenih efekata:

```
dt.lm1 <- lm(sirina_lista~lokalitet+stablo%in%lokalitet,data=dt)
```

Tablica analize varijance može se ispisati primjenom generičke funkcije `anova`:

```
anova(dt.lm1)
```

Analysis of Variance Table

Response: sirina\_lista

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
lokalitet	2	351.0	175.498	7.0109	0.0018354	**
lokalitet:stablo	12	1035.1	86.260	3.4460	0.0006956	***
Residuals	60	1501.9	25.032			

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Iako se u modelu koristi specijalna oznaka `,` u tablici analize varijance učinak stabla unutar lokaliteta naveden kao interakcija između lokaliteta i stabala. Upravo je interakcija s nadređenim faktorom bez glavnog efekta ugniježđenog faktora alternativni način njegovog zapisa. Stoga je u praktičnom smislu jedina razlika između unakrsnog i hijerarhijskog modela u tome što se u hijerarhijskom modelu ne

pojavljuje glavni efekt ugniježđenog faktora. Za skraćeni zapis hijerarhijskog modela koristi se znak /, pa se ista analiza može provesti i na ovaj način:

```
dt.lm2 <- lm(sirina_lista~lokalitet/stablo,data=dt)
anova(dt.lm2)
Analysis of Variance Table

Response: sirina_lista
          Df Sum Sq Mean Sq F value    Pr(>F)
lokalitet  2   351.0  175.498   7.0109 0.0018354 **
lokalitet:stablo 12 1035.1   86.260   3.4460 0.0006956 ***
Residuals   60 1501.9   25.032
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ova je tablica analize varijance identična prethodnoj, pa se i u njoj pojavljuje isti problem: neispravan F-test za faktor `lokalitet`. Prema načelu 5. navedenom u uvodnom dijelu F statistika za učinak lokaliteta bi se trebala izračunati kao omjer između prosječnog kvadrata lokaliteta i stabala unutar lokaliteta ( $175.498 / 86.260 = 2.035$ ), a ne u odnosu na pogrešku kao što je navedeno u tablici ( $175.498 / 25.032 = 7.011$ ). Ispravan test se može provesti zamjenom funkcije `lm` funkcijom `aov` i definiranjem efekta stabala unutar lokaliteta kao druge pogreške:

```
dt.lm3 <- aov(sirina_lista~lokalitet+Error(lokalitet:stablo),data=dt)
Warning message:
In aov(sirina_lista ~ lokalitet + Error(lokalitet:stablo), data = dt) :
  Error() model is singular
```

Iako se pojavljuje upozorenje o pogrešci, rezultati analize pohranjeni su u objekt `dt.lm3`. U ovom se slučaju ne mogu ispisati primjenom generičke funkcije `anova`, ali se umjesto nje može primijeniti generička funkcija `summary`:

```
summary(dt.lm3)

Error: lokalitet:stablo
          Df Sum Sq Mean Sq F value Pr(>F)
lokalitet  2   351   175.50   2.035 0.173
Residuals 12  1035    86.26

Error: within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 60  1502    25.03
```

Definiranje nekog efekta kao druge pogreške u kontekstu R-a automatski podrazumijeva i da se taj efekt tretira kao slučajan, pa je zato u ovom ispisu naveden F-test samo za lokalitete, ali ne i za stabla unutar lokaliteta (prema načelu 1. iz

uvodnog dijela). Nedostatak ovog testa zapravo i nije nedostatak, jer je primarni cilj ove analize ispitivanje razlika između lokaliteta, dok bi informacija o razlikama između stabala imala vrlo malu praktičnu vrijednost. Prema tome, na temelju ishoda analize može se zaključiti da se divlje trešnje s ova tri lokaliteta ne razlikuju u širini lista.

Tretiranjem jednog od efekata u modelu kao slučajnog, izvorni se fiksni model pretvara u mješoviti, a za analize prema mješovitom modelu se u R-u koriste specijalizirani paketi, od kojih je najpoznatiji `lme4`. Stoga sljedeća varijanta ove analize uključuje primjenu jedne funkcije iz tog paketa.

### 3. Analiza prema mješovitom modelu

Prethodno instalirani paket `lme4` najprije se aktivira:

```
library(lme4)
```

Analiza se provodi primjenom funkcije `lmer`, a slučajni efekti se označavaju zagradama, unutar kojih se ispred imena efekta upisuju broj 1 i okomita crta:

```
dt.lm4 <- lmer(sirina_lista~lokalitet+(1|lokalitet:stablo),data=dt)
```

Tablica analize varijance (za fiksne efekte) može se ispisati primjenom generičke funkcije `anova`:

```
anova(dt.lm4)
Analysis of Variance Table
      Df Sum Sq Mean Sq F value
lokalitet  2 101.86  50.928  2.0345
```

Vrijednost F statistike je identična kao u ranijoj varijante analize, ali joj pripadajuća P vrijednost nije navedena. Razlog tome je što u ovom kontekstu F-test nije egzaktn, nego samo aproksimativan, pa su autori ovog paketa odlučili izostaviti P vrijednost. No, postoji dodatni paket `lmerTest` koji za objekt s rezultatima primjene funkcije `lmer` generira nekoliko različitih testova. Instalirani paket se najprije aktivira:

```
library(lmerTest)
```

F-test za fiksne efekte, u ovom slučaju lokalitet, ispiše se primjenom generičke funkcije `anova`:

```
anova(dt.lm4)
```

```
Analysis of Variance Table of type III with Satterthwaite
approximation for degrees of freedom
```

```
      Sum Sq Mean Sq NumDF DenDF F.value Pr(>F)
lokalitet 101.86  50.928     2    12  2.0345 0.1734
```

Ovaj paket također pruža mogućnost provedbe testa za slučajne efekte, koji se izvodi kao test omjera vjerodostojnosti modela sa i bez efekta za koji se provodi test:

```
rand(dt.lm4)
```

```
Analysis of Random effects Table:
```

```
      Chi.sq Chi.DF p.value
lokalitet:stablo  9.77     1  0.002 **
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test za slučajne efekte nije ekvivalentan testu za fiksne efekte, nego njegov signifikantan ishod znači da je model koji uključuje testirani efekt učinkovitiji od modela bez tog efekta.

## Literatura

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.
2. Piepho H.P., Büchse A., Emrich K. (2003) A Hitchhiker's Guide to Mixed Models for Randomized Experiments. Journal of Agronomy & Crop Science, 189: 310-322.
3. Vaupotić T. (2008) Istraživanja ampelografskih značajki cv. Sauvignon bijeli (*Vitis vinifera* L.) na različitim podlogama (magistarski rad). Sveučilište u Zagrebu, Agronomski fakultet.
4. Ballian D., Bogunić F., Čabaravdić A., Pekeč S., Franjić J. (2012) Population differentiation in the wild cherry (*Prunus avium* L.) in Bosnia and Herzegovina. Periodicum biologorum 114: 43-54.
5. Bates D., Maechler M., Bolker B., Walker S. (2015) Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67: 1-48.

6. Kuznetsova A., Brockhoff P.B., Christensen R.H.B. (2017) lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26



## Dizajni s potpunim blokovima

Osnovni je razlog postavljanja pokusa procjena i usporedba učinaka tretiranja koja se primijenjuju na pokusnim jedinicama. Dakle, razlike između različito tretiranih pokusnih jedinica predmet su istraživanja, a razlike jednako tretiranih pokusnih jedinica predstavljaju smetnju (tj. pogrešku pokusa). One nastaju uslijed djelovanja niza faktora koji se nalaze izvan kontrole eksperimentatora, a jedan od načina kako se može ublažiti njihov utjecaj, odnosno postići bolja kontrola pogreške pokusa, grupiranje je pokusnih jedinica u blokove. U analizi pokusa se tada u model može uvrstiti i učinak blokova, pomoću kojeg se može objasniti bar dio varijabilnosti koja nastaje djelovanjem faktora izvan kontrole eksperimentatora, te na taj način smanjiti pogreška. Blokovi unutar kojih su primijenjena sva tretiranja, na način da je svako od njih primijenjeno na samo jednoj pokusnoj jedinici, nazivaju se potpunim blokovima ili repeticijama.

**Potpuno slučajni raspored** (Completely Randomized Design - CRD). Ovo je najjednostavniji mogući dizajn, koji podrazumijeva u potpunosti slučajnu dodjelu tretiranja pokusnim jedinicama, bez ikakvih ograničenja u randomizaciji. U ovom dizajnu nema grupiranja pokusnih jedinica, zbog čega nema ni kontrole pogreške, pa se više koristi kao model za analizu, jer se radi o modelu istovjetnom osnovnom modelu analize varijance.

**Slučajni blokni raspored** (Randomized Complete Block Design – RCBD). Ovaj je dizajn strukturiran tako da se sastoji od zadanog broja potpunih blokova ili repeticija. U svakoj repeticiji broj pokusnih jedinica mora biti jednak broju tretiranja, pa je ukupan broj pokusnih jedinica jednak produktu broja tretiranja i broja repeticija. Analiza podataka iz pokusa postavljenih prema ovom dizajnu provodi se prema modelu koji nastaje dodavanjem učinka repeticija u osnovni model analize varijance:

$$Y = TRET + REP \quad (7)$$

**Latinski kvadrat** (Latin Square – LS). Grupiranje pokusnih jedinica se provodi prema dva različita kriterija ili u dva različita smjera. U modelu prema kojem se provodi analiza se stoga pojavljuju dva učinka blokova (koji se uvjetno mogu nazvati „vodoravnim“ i „okomitim“ repeticijama:

$$Y = TRET + V.REP + O.REP \quad (8)$$

Učinkovitost dizajna mjeri se relativnim smanjenjem pogreške razlike prosjeka tretiranja u analizi prema modelu za taj dizajn u odnosu na pogrešku u analizi prema nekom jednostavnijem modelu (dizajnu). Pogreška razlike se procjenjuje pomoću prosječnog kvadrata pogreške u analizi varijance i broja repeticija (odnosno broja pokusnih jedinica na kojima je primijenjeno svako od tretiranja):

$$s_D = \sqrt{\frac{2MSE}{r}} \quad (9)$$

Relativna učinkovitost slučajnog blokno dizajna određuje se u odnosu na pogrešku razlike procijenjenu na temelju analize istih podataka prema modelu potpuno slučajnog rasporeda, te predstavlja omjer kvadratnih vrijednosti (varijanci) pogrešaka:

$$E_{R(RCBD)} = \frac{\frac{2MSE_{CRD}}{r}}{\frac{2MSE_{RCBD}}{r}} = \frac{MSE_{CRD}}{MSE_{RCBD}} \quad (10)$$

Dakle, omjer varijanci pogrešaka usporedbi jednak je omjeru prosječnih kvadrata pogrešaka dvaju modela. Na isti se način relativna učinkovitost latinskog kvadrata može odrediti u odnosu na slučajni blokni raspored:

$$E_{R(LS)} = \frac{MSE_{RCBD}}{MSE_{LS}} \quad (11)$$

Analiza prema reduciranom modelu (slučajnom bloknom rasporedu) se u ovom slučaju može provesti na dva načina: izostavljanjem jednog ili drugog tipa blokova. Dakle, za latinski kvadrat postoje dvije procjene učinkovitosti, a svaka se odnosi na jedan tip blokova (onaj koji je izostavljen iz reduciranog modela).

## Primjer 4.1

Prodaja sjemena i sadnog materijala dopuštena je samo za one sorte koje se nalaze na službenoj sortnoj listi. Preduvjet za uvrštavanje na sortnu listu za svaku je sortu uspješno završen postupak priznavanja, u kojem se gospodarski važna svojstva

prijavljenih sorata uspoređuju sa svojstvima sorata koje zauzimaju najveći udio u proizvodnji (koj se nazivaju standardi). Tijekom tog postupka provode se poljski pokusi na nekoliko lokacija, ponovljeni u više godina. U sklopu postupka priznavanja sorata jare zobi, u Osijeku je 1997. godine proveden pokus s četiri sorte, dizajniran kao slučajni blokni raspored s četiri repeticije. U ovom će se primjeru analizirati podatci o urodu (dt/ha).

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
jz97 <-  
  read.csv("Primjer4_1.csv", colClasses = c("factor", "factor", "numeric"))
```

Argument `colClasses` koristi se kako bi se varijable `sorta` i `rep` (repeticija) učitale kao faktori, jer su obje označene brojevima (sorte zato jer se zbog nepristranosti ime same sorte i njenog tvorca ne otkriva do završetka postupka).

### 2. Analiza varijance

Analiza se provodi primjenom modela slučajnog blokno rasporeda:

```
jz97.rcbd <- lm(urod~sorta+rep,data=jz97)
```

Generička funkcija `anova` koristi se za ispis tablice analize varijance:

```
anova(jz97.rcbd)  
Analysis of Variance Table  
  
Response: urod  
          Df Sum Sq Mean Sq F value Pr(>F)  
sorta      3  287.98   95.992   4.3640 0.02692 *  
rep        4  161.42   40.356   1.8346 0.18707  
Residuals 12  263.96   21.997  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test za repeticije se u načelu ne razmatra, jer blokovi nisu predmet istraživanja nego samo sredstvo koji služi u svrhu smanjenja pogreške. U ovom primjeru, iako test za

blokove nije signifikantan, učinak blokova nije zanemariv, jer je njegov udio u ukupnoj varijabilnosti:

```
161.42/(287.98+161.42+263.96)  
[1] 0.2262813
```

To je nešto više od petine, pa se može smatrati da su u ovom primjeru blokovi ispunili svoju primarnu svrhu, a to je smanjenje pogreške. Primarni interes u ovom pokusu su razlike između sorata, pa pošto je test za učinak sorata signifikantan, može se provesti i test za višestruke usporedbe, kojim će se usporediti sorte koje se ispituju sa standardnom sortom.

### 3. Usporedbe sorata sa standardom

Najprije se aktivira paket `multcomp`:

```
library(multcomp)
```

Od četiri sorte u pokusu dvije su bile novostvorene, a dvije standardne. Kako se za jaru zob u postupku priznavanja koristila samo jedna standardna sorta (Šampionka), pokus je izvorno trebao uključivati samo tri sorte. Zbog toga je povećan broj parcela u svakoj repeticiji dodavanjem još jedne parcele sa standardnom sortom, pa su brojevi 1 i 4 označavali standardnu sortu. Kako bi se mogla provesti usporedba dviju sorata sa standardom potrebno je dodjeliti jedinstvenu oznaku za oba uzorka standarda, te ponoviti prethodne korake. Međutim, postoji i jednostavnije rješenje, usporedba s prosjekom uzoraka standarda, za koju je potrebno konstruirati matricu kontrasta. U načelu se za Dunnettove usporedbe (s kontrolom, tj. standardom) ta matrica može jednostavno konstruirati uz pomoć funkcije `contrMat`, međutim se u ovom primjeru zbog potrebe definiranja kontrasta s prosjekom standarda mora definirati izravno:

```
cm <- rbind("2 - pr.st." = c(-0.5, 1, 0, -0.5),  
           "3 - pr.st." = c(-0.5, 0, 1, -0.5))
```

Svaka od dviju usporedbi definirana je u zasebnom retku. U prvom retku definirana je usporedba sorte br. 2 s prosjekom standarda ("2 - pr.st."), kao vektor koeficijenata koji definiraju linearnu kombinaciju četiri prosjeka (koji se doslovno može pročitati kao: „od prosjeka sorte br. 2 oduzmi polovicu prosjeka sorte br. 1 i

polovicu prosjeka sorte br. 4“). Na analogan način definirana je i usporedba sorte br. 3 s prosjekom standarda u drugom retku, a zatim su dva retka spojena u matricu primjenom funkcije `rbind`. Razlike i njihove standardne pogreške izračunaju se primjenom funkcije `glht`:

```
jz97.glht <- glht(jz97.rcbd,linfct=mcp(sorta=cm))
```

Pošto se provode samo dva testa, korekcija nije potrebna, pa se rezultati ispišu primjenom generičke funkcije `summary` s ključnom riječi `none`:

```
summary(jz97.glht,test=adjusted("none"))
```

```
Simultaneous Tests for General Linear Hypotheses
Multiple Comparisons of Means: User-defined Contrasts

Fit: lm(formula = urod ~ sorta + rep, data = jz97)
Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
2 - pr.st. == 0    7.476      2.569   2.910  0.0131 *
3 - pr.st. == 0    7.105      2.569   2.766  0.0171 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- none method)
```

Testovi su signifikantni, pa se može zaključiti da su obje sorte u Osijeku 1997. ostvarile veći urod od standarda. Ukoliko sorte i u svim ostalim pokusima ostvare ovakav rezultat, ili barem budu na razini standarda, na kraju postupka priznavanja se predlože za uvrštavanje na sortnu listu.

#### 4. Procjena učinkovitosti

Relativna učinkovitost slučajnog bloknog rasporeda određuje se u odnosu na potpuno slučajni raspored. Zbog toga je najprije potrebno provesti i analizu prema modelu potpuno slučajnog rasporeda:

```
jz97.crd <- lm(urod~sorta,data=jz97)
```

Prosječni se kvadrati za oba modela mogu ispisati primjenom generičke funkcije `anova` s dodatkom  `'$'Mean Sq'`:

```
anova(jz97.crd)$'Mean Sq'  
[1] 95.99205 26.58631  
> anova(jz97.rcbd)$'Mean Sq'  
[1] 95.99205 40.35561 21.99655
```

Prosječni kvadrat pogreške je uvijek zadnja vrijednost u nizu:

```
mse.crd <- anova(jz97.crd)$'Mean Sq'[2]  
mse.rcbd <- anova(jz97.rcbd)$'Mean Sq'[3]
```

Relativna učinkovitost slučajnog blokno­g rasporeda omjer je dvaju prosječnih kvadrata:

```
Er.rcbd <- mse.crd/mse.rcbd  
Er.rcbd  
[1] 1.208658
```

Pogreška usporedbe prosjeka tretiranja bi u potpuno slučajnom rasporedu bila bi 1.21 puta veća, pa relativno povećanje učinkovitosti slučajnog blokno­g rasporeda u odnosu na potpuno slučajni raspored iznosi 21%. Dakle, slučajni blokni raspored s četiri repeticije jednako je učinkovit kao što bi bio potpuno slučajni raspored s 1.21 puta više ponavljanja svakog tretiranja, odnosno u kojem bi svako tretiranje bilo primijenjeno na pet pokusnih jedinica. Zbog toga primjena slučajnog blokno­g rasporeda donosi i uštedu (u smislu smanjenih troškova pokusa), jer bi za postizanje iste informativnosti pokusa primjena potpuno slučajnog rasporeda zahtjevala pet parcela više.

## Primjer 4.2

U postupku priznavanja sorata (opisanom u Primjeru 4.2) jare zobi, u Osijeku je 1998. godine proveden sortni pokus s pet sorata, dizajniran kao latinski kvadrat. U ovom će se primjeru analizirati podatci o urodu (dt/ha).

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
jz98 <- read.csv("Primjer4_2.csv",  
                colClasses = c("factor","factor","factor","numeric"))
```

Argument `colClasses` koristi se kako bi se varijable `sorta`, te `repV` i `rep0` („vodoravne“ i „okomite“ repeticije) učitale kao faktori, jer su sve tri označene brojevima.

### 2. Analiza varijance

Analiza se provodi primjenom modela latinskog kvadrata:

```
jz98.ls <- lm(urod~sorta+repV+rep0,data=jz98)
```

Tablica analize varijance se ispiše primjenom generičke funkcije :

```
anova(jz98.ls)
```

Analysis of Variance Table

Response: urod

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
sorta	4	526.22	131.555	2.6137	0.08837 .
repV	4	333.75	83.438	1.6577	0.22387
rep0	4	747.27	186.817	3.7116	0.03444 *
Residuals	12	604.00	50.333		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Razlike između sorata nisu značajne, što znači da su četiri nove sorte na razini standarda (Šampionke), te u ovom primjeru nema potrebe za provođenjem višestrukih usporedbi. Rezultati ovog pokusa, prema tome, ne idu prilogu odluci za niti odluci protiv uvrštavanja novih sorata na sortnu listu, nego to ovisi o rezultatima ostalih pokusa.

Od dva tipa repeticija, značajne razlike u uvjetima za rast i razvoj biljaka jare zobi postoje samo između repeticija položenih u „okomitom“ smjeru. Opravdanost uvođenja jednog i drugog tipa repeticija može se dodatno provjeriti usporedbom ove analize s analizama prema modelima u kojima su ispušteni učinci jednog ili drugog tipa repeticija.

### 3. Procjena učinkovitosti

Relativna učinkovitost „vodoravnih“ repeticija u latinskom kvadratu određuje se u odnosu na analizu prema modelu slučajnog bloknoeg rasporeda koji sadrži samo jedan učinak repeticija, i to „okomitih“:

```
jz98.rcbdo <- lm(urod~sorta+rep0,data=jz98)
```

Za procjenu učinkovitosti „okomitih“ repeticija služi analiza prema modelu koji sadrži samo učinak „vodoravnih“ repeticija:

```
jz98.rcbdv <- lm(urod~sorta+repv,data=jz98)
```

Tri prosječna kvadrata potrebna za procjenu učinkovitosti su:

```
mse.ls <- anova(jz98.ls)$'Mean Sq'[4]  
mse.rcbdo <- anova(jz98.rcbdo)$'Mean Sq'[3]  
mse.rcbdv <- anova(jz98.rcbdv)$'Mean Sq'[3]
```

Relativna učinkovitost „vodoravnih“ repeticija predstavlja omjer prosječnih kvadrata pogrešaka u modelu latinskog kvadrata i modelu slučajnog bloknoeg rasporeda s „okomitim“ repeticijama:

```
Er.v <- mse.rcbdo/mse.ls
```



Za „okomite“ se repeticije učinkovitost procjenjuje u odnosu na model s „vodoravnim“ repeticijama:

```
Er.o <- mse.rcbdv/mse.1s
```

Dvije procjene učinkovitosti iznose:

```
Er.v  
[1] 1.164425  
Er.o  
[1] 1.677899
```

U skladu se prethodnim zaključcima na temelju F-testova u analizi varijanci, doprinos „vodoravnih“ repeticija povećanju učinkovitosti pokusa je znatno manji, iznosi tek 16%. Za učinkovitost je pokusa znatno važnije bilo koristiti „okomite“ repeticije, čije povećanje učinkovitosti u odnosu na slučajni blokni raspored iznosi čak 68%. To znači da bi hipotetski pokus, postavljen prema slučajnom bloknom rasporedu s „vodoravnim“ repeticijama, morao uključivati tri dodatne repeticije kako bi postigao istu razinu informativnosti.

## Literatura

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.
2. Bretz F., Hothorn T., Westfall P. (2010) Multiple Comparisons Using R. Chapman and Hall/CRC, Boca Raton, FL.

## Dizajni s nepotpunim blokovima

U dizajnim iz prethodnog poglavlja potpuni blokovi mogu ispuniti svoju svrhu smanjenja pogreške samo ako je zadovoljena pretpostavka da su uvjeti koji vladaju unutar repeticija ujednačeni. S porastom broja tretiranja raste i vjerojatnost da će ta pretpostavka biti narušena, pa se u pokusima u kojima se ispituje učinak većeg broja tretiranja javlja potreba za primjenom manjih blokova koji nužno ne sadrže sva tretiranja. Primarni kriterij u klasifikaciji dizajna s nepotpunim blokovima je mogućnost da se više nepotpunih blokova kombinira tako da tvore potpuni blok, odnosno repeticiju. Dizajni od čijih se blokova mogu tvoriti repeticije nazivaju se rješivima (resolvable), a oni kod kojih to nije moguće nerješivima (non-resolvable). U primjerima u ovom priručniku će se isključivo koristiti rješivi dizajni.

**Alfa dizajn** ( $\alpha$  Design). U ovom rješivom dizajnu s cikličkom strukturom svaka se repeticija sastoji od  $s$  nepotpunih blokova veličine  $k$ , pri čemu je optimalno rješenje kada je  $s \approx k$ . Čak i kada ne postoje  $s$  i  $k$  čiji bi produkt bio jednak broju tretiranja ( $s \times k = t$ ), postoji mogućnost generiranja dizajna s dvije veličine blokova  $k$  i  $k - 1$  (veličine blokova se smiju razlikovati samo za jednu pokusnu jedinicu). Model za analizu podataka pokusa postavljenog prema alfa dizajnu je hijerarhijski, jer je učinak nepotpunih blokova ugniježđen unutar učinka repeticija. Osim toga, učinak nepotpunih blokova se tretira kao slučajni efekt; njihovi prosjeci nisu usporedivi jer su sastavljeni od različito tretiranih pokusnih jedinica. Prema tome, podatci iz pokusa postavljenog prema alfa dizajnu analiziraju se prema mješovitom modelu:

$$Y = TRET + REP :: REP \cdot BLOK \quad (12)$$

Relativna učinkovitost alfa dizajna procjenjuje se u odnosu na slučajni blokni raspored i predstavlja relativno smanjenje progredne razlike tretiranja koje se postiže uvođenjem nepotpunih blokova. Razlike tretiranja u alfa dizajnu su opterećene različitim pogreškama, ovisno o tome pojavljuju li se dva tretiranja u istom nepotpunom bloku ili ne. Zbog toga se procjena učinkovitosti alfa dizajna temelji na prosječnoj pogrešci svih usporedbi tretiranja:

$$E_{R(\alpha)} = \frac{S_{D_{RCBD}}^2}{\bar{S}_{D_\alpha}^2} \quad (13)$$

**Redno-stupčani dizajn** (Row-Column Design). Strukturni elementi dizajna su potpuni blokovi (repeticije) i dva tipa nepotpunih blokova: redovi i stupci. Repeticije su pravokutnog oblika, sastoje se od redova i stupaca, pa je primjena ovog dizajna ograničena samo na one brojeve tretiranja koji se mogu rastaviti na produkt dvaju cijelih brojeva (broja redova i broja stupaca). Konstrukcija redno-stupčanog dizajna može se temeljiti na alfa dizajnu, a i model za analizu se može izvesti na analogan način. Redovi i stupci su ugniježđeni unutar repeticija, a tretiraju se kao slučajni efekti:

$$Y = TRET + REP :: REP \cdot RED + REP \cdot STUPAC \quad (14)$$

Reducirani model za procjenu učinkovitosti redno-stupčanog dizajna može biti model u kojem je ispušten učinak stupaca ili model u kojem je ispušten učinak redova. Na taj se način može procijeniti učinkovitost dodavanja efekta stupaca, odnosno dodavanja efekta redova u jednodimenzionalni model koji sadrži samo jedan tip blokova, i koji se može interpretirati kao alfa dizajn. Relativna učinkovitost kvadratna je vrijednost omjer prosječnih pogrešaka razlika tretiranja dva modela:

$$E_{R(RC)} = \frac{\bar{S}_{D\alpha}^2}{\bar{S}_{DRC}^2} \quad (15)$$

**Latinizacija** podrazumijeva provođenje radnomizacije tako da se u susjednim blokovima različitih repeticija izbjegava ponavljanje tretiranja. Na taj način se formiraju „dugi“ blokovi koji se protežu kroz više repeticija, u kojima se nalazi većina ili čak sva tretiranja. Latinizacija se može provoditi u oba prethodno opisana dizajna, a u oba slučaja će se analiza provesti prema osnovnom modelu u koji je dodan učinak „dugih“ blokova, koji se tretiraju kao fiksni efekt (jer sadrže većinu ili sva tretiranja pa se mogu tretirati kao potpuni blokovi). Tako npr. model za analizu latiniziranog alfa dizajna izgleda ovako:

$$Y = TRET + REP + DBLOK :: REP \cdot BLOK \quad (16)$$

## Primjer 5.1

U postupku priznavanja sorata (opisanom u Primjeru 4.2) suncokreta, u Osijeku je 2001. godine proveden sortni pokus s dvadeset hibrida. Pokus je izveden prema alfa dizajnu s četiri repeticije, od kojih je svaka bila podijeljena na četiri bloka, a svaki blok se sastojao od pet parcela (Gunjača i sur. 2005). U ovom će se primjeru analizirati podatci o urodu (dt/ha).

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
sun <- read.csv("Primjer5_1.csv",
               colClasses = c("factor", "factor", "factor", "numeric"))
```

Argument `colClasses` koristi se kako bi se varijable `hibrid`, te `rep` i `blok` (repeticije i nepotpuni blokovi) učitale kao faktori, jer su sve tri označene brojevima.

### 2. Analiza prema mješovitom modelu

Najprije je potrebno aktivirati pakete `lme4` i `lmerTest`:

```
library(lme4)
library(lmerTest)
```

Analiza se provodi primjenom funkcije `lmer`, a slučajni efekti blokova unutar repeticija se označavaju na način opisan u Primjeru 3.2:

```
sun.alpha <- lmer(urod~hibrid+rep+(1|rep:blok),data=sun)
```

Nakon toga se tablica analize varijance ispiše primjenom generičke funkcije `anova`:

```
anova(sun.alpha)
Analysis of Variance Table of type III with Satterthwaite
approximation for degrees of freedom
      Sum Sq Mean Sq NumDF  DenDF F.value Pr(>F)
hibrid 1194.3  62.857    19 47.452 10.0679 7e-11 ***
rep      11.4   3.801     3 10.486  0.6089 0.6237
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Razlike između hibrida su značajne, pa će uz pomoć višestrukih usporedbi istražiti razlike u urodu između novih hibrida i standarda koji je označen brojem 1. Varijance slučajnih efekata mogu se ispisati primjenom generičke funkcije `print` s argumentom `comp="Variance"` na rezultat primjene funkcije `varCorr` (iz paketa `lme4`):

```
print(VarCorr(sun.alpha), comp="Variance")
Groups   Name          Variance
rep:blok (Intercept) 7.4655
Residual                6.2433
```

Varijanca blokova veća je od varijance pogreške, pa se već sada može naslutiti da je uvođenje blokova rezultiralo određenim povećanjem učinkovitosti analize. Test omjera vjerodostojnosti modela sa i bez blokova može se provesti primjenom funkcije `rand` (iz paketa `lmerTest`):

```
rand(sun.alpha)
Analysis of Random effects Table:
      Chi.sq Chi.DF p.value
rep:blok  15.7     1 8e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ovaj test potvrđuje da je uvođenje učinka blokova rezultiralo značajnim povećanjem učinkovitosti modela. Intenzitet tog povećanja će se procijeniti uz pomoć pogrešaka razlika učinaka hibrida dvaju modela, a za procjenu prosječne pogreške u modelu alfa dizajna potrebno je provesti višestruke usporedbe.

### 3. Višestruke usporedbe

Najprije je potrebno aktivirati paket `multcomp`:

```
library(multcomp)
```

Višestruke se usporedbe provode u odnosu na standardni hibrid označen brojem 1, pa se usporedbe generiraju primjenom ključne riječi `Dunnnett` (koja se koristi za usporedbe s kontrolnim tretiranjem):

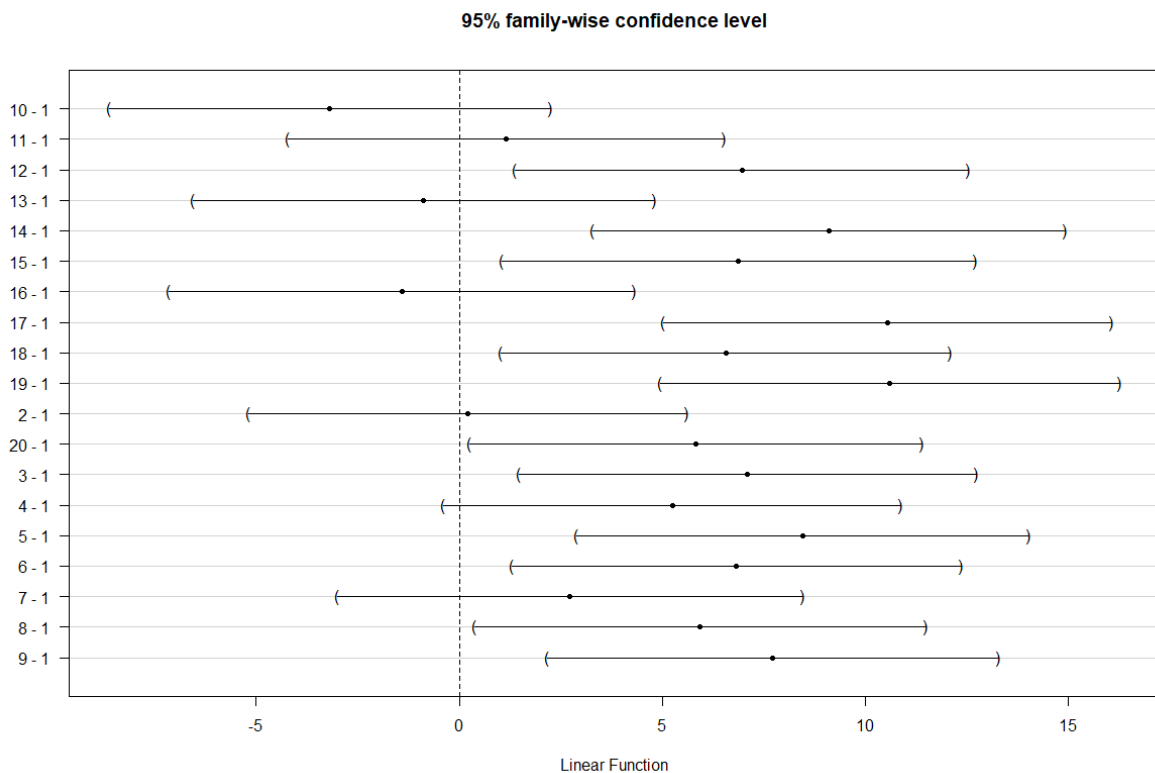
```
sun.glht <- glht(sun.alpha, linfct=mcp(hibrid="Dunnnett"))
```

Devetnaest zadanih testova provodi se primjenom generičke funkcije uz Bonferronijevu korekciju i sprema u novi objekt za kasniju uporabu:

```
sun.mc <- summary(sun.glht, test=adjusted(type="bon"))
```

Umjesto ispisa tablice, rezultate testova je bolje prikazati grafički:

```
plot(sun.mc)
```



Slika 13. Grafički prikaz Dunnettovog testa razlika uroda hibrida (razlike čiji intervali pouzdanosti uključuju vrijednost 0 nisu signifikantne)

Intervali pouzdanosti razlike između standarda i hibrida br. 2, 4, 7, 10, 11, 13 i 16 uključuju nulu, pa se urod navedenih hibrida ne razlikuje značajno od hibrida standarda. Razlike između svih ostalih hibrida i standarda su signifikantne, a svi se intervali pouzdanosti razlika nalaze desno od nule, što znači da je urod tih hibrida značajno veći u odnosu na urod standarda.

Prosječna pogreška svih usporedbi potrebna je za procjenu učinkovitosti, a može se izračunati primjenom funkcije `mean` na odgovarajući dio objekta `sun.mc` u kojem su spremljene razlike i njihove pogreške:

```
mean_sed_alpha <- mean(sun.mc$test$sigma)
```

#### 4. Procjena učinkovitosti alfa dizajna

Relativna učinkovitost alfa dizajna određuje se u odnosu na slučajni blokni raspored, pa je potrebno provesti i analizu prema modelu slučajnog blokno rasporeda:

```
sun.rcbd <- lm(urod~hibrid+rep,data=sun)
```

Pogreška razlike prosjeka hibrida u slučajnom je bloknom rasporedu ista za sve usporedbe, a njena je varijanca jednaka omjeru dvostruke vrijednosti prosječnog kvadrata pogreške i broja repeticija:

```
mse.rcbd <- anova(sun.rcbd)$'Mean Sq'[3]  
sq_sed_rcbd <- 2*mse.rcbd/4
```

Relativna učinkovitost omjer je kvadratnih vrijednosti dviju pogrešaka:

```
Er_alpha <- sq_sed_rcbd/mean_sed_alpha**2  
Er_alpha  
[1] 1.564138
```

Relativno povećanje učinkovitosti u odnosu na slučajni blokni raspored je 56%. To povećanje posljedica je podjele repeticija na manje i ujedno homogenije blokove. Alfa dizajn s četiri repeticije jednako je učinkovit kao što bi bio slučajni blokni raspored s 1.56 puta više repeticija, odnosno približno šest repeticija. Zbog toga je primjenom alfa dizajna ostvarena ušteda od dvije repeticije u odnosu na slučajni blokni raspored.

## Primjer 5.2

Pokus s dvadeset inbred-linija kukuruza postavljen je u Rugvici 2008. godine, prema latiniziranom redno-stupčanom dizajnu u 4 repeticije. Svaka se repeticija sastojala od

četiri reda i pet stupca, latinizacija je provedena tako da su susjedni redovi svih repeticija tvorili jedan „dugi“ red (koji je sadržavao svih dvadeset linija). U ovom će se primjeru analizirati podatci o urodu (dt/ha).

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
kuk <- read.csv("Primjer5_2.csv",
               colClasses = c("factor", "factor", "factor", "factor", "numeric"))
```

Argument `colClasses` koristi se kako bi se varijable `linija`, te `rep`, `red` i `stupac` (repeticije i nepotpuni blokovi) učitale kao faktori, jer su sve četiri označene brojevima.

### 2. Analiza prema mješovitom modelu

Najprije je potrebno aktivirati pakete `lme4` i `lmerTest`:

```
library(lme4)
library(lmerTest)
```

Analiza se provodi primjenom funkcije `lmer`, a ista se varijabla `red` može koristiti za „dugih“ redova (fiksni efekt), odnosno redova kao nepotpunih blokova (slučajni efekt redova unutar repeticija):

```
kuk.lrc <- lmer(urod~linija+rep+red+(1|rep:red)+(1|rep:stupac),data=kuk)
```

Tablica analize varijance s F-testovima za fiksne efekte može se ispisati uz pomoć generičke funkcije `anova`:

```
anova(kuk.lrc)
Analysis of Variance Table of type III with Satterthwaite
approximation for degrees of freedom
```

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
linija	2120.18	111.588	19	46.125	4.7240	8.258e-06	***
rep	700.54	233.512	3	8.556	9.8856	0.003797	**
red	96.50	32.166	3	8.556	1.3617	0.318231	



```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Između linija postoje značajne razlike koje se mogu ispitati uz pomoć višestrukih usporedbi, a na taj način će se ujedno dobiti i prosječna pogreška razlike koja je potrebna za procjenu relativne učinkovitosti dizajna. Repeticije se značajno razlikuju, ali „dugi“ redovi ne, pa se njihov učinak u ovom primjeru može izostaviti iz modela. Varijance slučajnih efekata ispišu se na način opisan u Primjeru 5.1:

```
print(VarCorr(kuk.lrc),comp="Variance")
Groups      Name      Variance
rep:stupac (Intercept)  0.000
rep:red     (Intercept) 85.614
Residual                    23.621
```

Kada se analiza prema mješovitim modelima provodi primjenom metode REML, procjenama komponenta varijance koje imaju negativne vrijednosti automatski se dodjeljuje vrijednost nula, što je vjerojatno razlog zašto je varijanca stupaca jednaka nuli. U svakom slučaju učinak stupac unutar repeticija je zanemariv, pa se može izostaviti iz model. Naprotiv, učinak redova unutar repeticija je znantan, njegova varijanca je skoro četverostruko veća od varijance pogreške.

### 3. Ponovljena analiza prema reduciranom modelu

Prema rezultatima inicijalne analize, model se može reducirati izostavljanjem učinaka „dugih“ redova i stupaca unutar repeticija. Prema tako reduciranom modelu ponovno se provede analiza:

```
kuk.lrc2 <- lmer(urod~linija+rep+(1|rep:red),data=kuk)
```

Ispiše se nova tablica analize varijance:

```
anova(kuk.lrc2)
Analysis of Variance Table of type III with Satterthwaite
approximation for degrees of freedom
      Sum Sq Mean Sq NumDF DenDF F.value Pr(>F)
linija 2111.88  111.15   19 46.082  4.7072 8.681e-06 ***
rep      640.35  213.45    3 11.626  9.0394  0.00228 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Test za oba preostala efekta je i u ponovljenoj analizi signifikantan, pa se ispišu varijance slučajnih efekata:

```
print(VarCorr(kuk.lrc2), comp="variance")
  Groups   Name      Variance
  rep:red (Intercept) 94.072
  Residual                23.613
```

Odnos preostalih slučajnih efekta se također nije bitno mijenjao. Sada se može provesti i test za usporedbu modela sa i bez nepotpunih blokova (redova unutar repeticija):

```
rand(kuk.lrc2)
Analysis of Random effects Table:
      Chi.sq Chi.DF p.value
rep:red  49.8      1 2e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

I ovaj test je signifikantan, što potvrđuje opravdanost primjene nepotpunih blokova, barem u jednom smjeru.

#### 4. Višestruke usporedbe

Početni korak je aktivacija paket `multcomp`:

```
library(multcomp)
```

U ovom primjeru ispitivanje razlika između linija nije primarni cilj analize, nego je više sredstvo za procjenu učinkovitosti dizajna. Usporediti će se sve linije međusobno, primjenom ključne riječi :

```
kuk.glht <- glht(kuk.lrc2, linfct=mcp(linija="Tukey"))
```

Ukupan broj testova je  $20 \cdot (20-1) / 2 = 190$ , pa je umjesto ispisivanja rezultate bolje spremite u novi objekt:

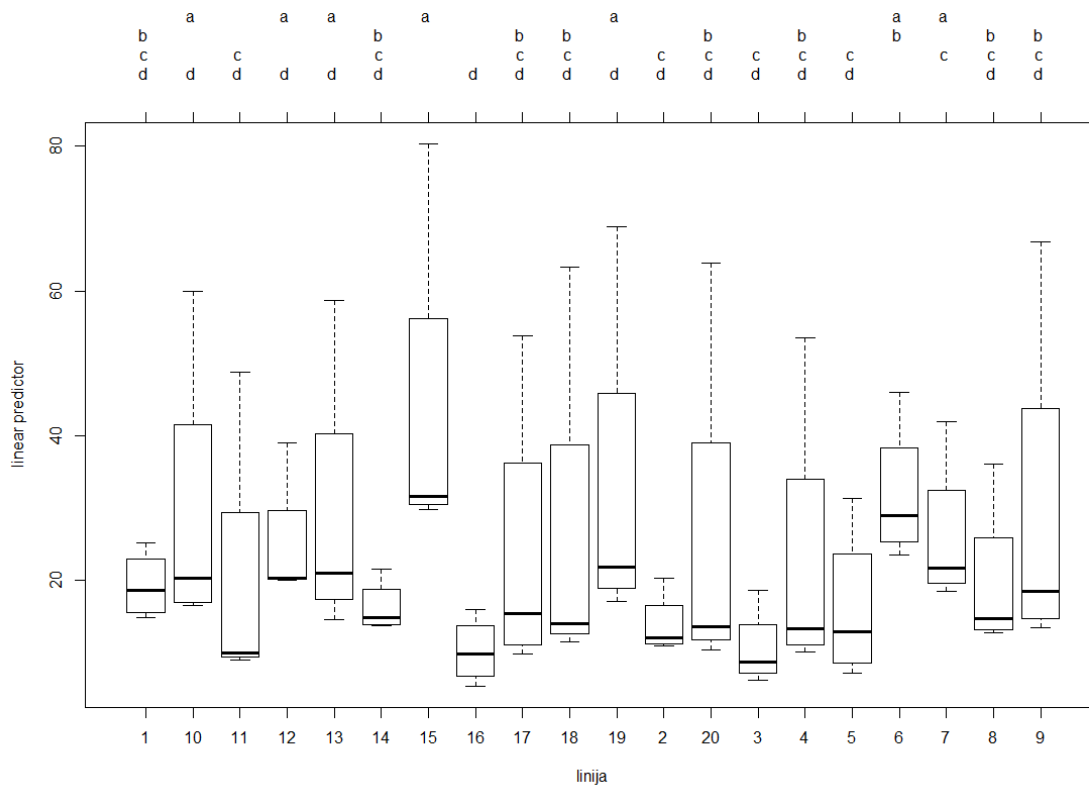
```
kuk.mc <- summary(kuk.glht, test=adjusted(type="bon"))
```

Slovne oznake na grafikonu distribucija (box-plot) su vjerojatno najpogodniji način ispisa u ovom primjeru. Prethodno je potrebno podesiti gornju marginu:

```
old.par <- par(mai=c(1,1,1.25,1), no.readonly = TRUE)
```

Nakon toga se može kreirati grafikon:

```
plot(cld(kuk.mc,decreasing = T))
par(old.par)
```



Slika 14. Grafički prikaz Bonferronijevog testa razlika uroda (dt/ha) inbred-linija kukuruza (linije označene istim slovom se ne razlikuju)

Prosječna pogreška svih usporedbi potrebna je za procjenu učinkovitosti, a može se izračunati primjenom na način opisan u Primjeru 5.1:

```
mean_sed_1rc2 <- mean(kuk.mc$test$sigma)
```

### 5. Procjena učinkovitosti

Reduciranjem modela izbacivanjem učinaka „dugih“ redova i stupaca unutar repeticija, izvorni se latinizirani redno-stupačni dizajn sveo na jednodimenzionalni

dizajn s nepotpunim blokovima, odnosno alfa dizajn. Procjena učinkovitosti se stoga u ovom primjeru provodi na isti način kao i u prethodnom primjeru, pa je reducirani model za procjenu učinkovitosti slučajni blokni raspored:

```
kuk.rcbd <- lm(urod~linija+rep,data=kuk)
```

Zatim se na temelju rezultata ove analize procijeni pogreška razlike prosjeka linija:

```
mse.rcbd <- anova(kuk.rcbd)$'Mean Sq'[3]  
sq_sed_rcbd <- 2*mse.rcbd/4
```

Relativna učinkovitost omjer je kvadratnih vrijednosti dviju pogrešaka:

```
Er_1rc2 <- sq_sed_rcbd/mean_sed_1rc2**2  
Er_1rc2  
[1] 3.471743
```

U ovom je primjeru relativno povećanje učinkovitosti u odnosu na slučajni blokni raspored iznimno veliko, jer je dizajn nastao podjelom repeticija na redove 3.47 puta učinkovitiji u odnosu na slučajni blokni raspored.

## Literatura

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.
2. Gunjača J., Renka H., Jambrešić I., Šindrak Z., Safner T., Liović I., Pecina M. (2005) Efficiency of Alpha Designs in Croatian Variety Trials. Lužar-Stiffler V., Hljuz Dobrić V. (ur.): Proceedings of the 27<sup>th</sup> International Conference on Information Technology Interfaces ITI 2005, Cavtat, Hrvatska. pp. 207-210.
3. Bates D., Maechler M., Bolker B., Walker S. (2015) Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67: 1-48.
4. Kuznetsova A., Brockhoff P.B., Christensen R.H.B. (2017) lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software 82: 1–26
5. Bretz F., Hothorn T., Westfall P. (2010) Multiple Comparisons Using R. Chapman and Hall/CRC, Boca Raton, FL.

## Dizajni s razdijeljenim parcelama

U višefaktorijskim se pokusima zbog različitih razloga tretiranja razinama nekog od faktora mogu primijeniti samo na pokusnim jedinicama određene minimalne veličine, ali se to ograničenje ne mora odnositi i na ostale faktore. U tim se situacijama primjenjuje neka od varijanti dizajna s razdijeljenim parcelama, od kojih je najjednostavnija split-plot, dizajn za dvofaktorijske pokuse. U tom se dizajnu pojavljuju dvije veličine pokusnih jedinica – glavne parcele (whole plots) i podparcele (subplots). Svaka se repeticija sastoji od onoliko glavnih parcela koliko ima razina glavnog faktora, a svaka se glavna parcela zatim dijeli na onoliko podparcela koliko ima razina podfaktora. Analiza pokusa postavljenog prema ovom dizajnu može se provesti prema fiksnom modelu, pri čemu se odvija u dvije faze: prva faza je analiza podataka s glavnih parcela, u kojoj se provode testovi za učinke glavnog faktora i repeticija s pogreškom za glavne parcele; druga faza je analiza preostale varijabilnosti podparcela u kojoj se provode testovi za učinke podfaktora i interakcije s pogreškom za podparcele (Vasilj 2000). Jednostavnija varijanta je analiza prema mješovitom modelu, u kojem se pogreška za glavne parcele definira kao slučajni efekt interakcije repeticija i glavnog faktora:

$$Y = A + REP + B + A \cdot B :: A \cdot REP \quad (17)$$

Višestruke se usporedbe razina faktora provode na uobičajen način, opisan u Primjeru 2. Za različite kombinacije razina faktora (interakcija) provode se dva tipa testova: za usporedbe razina podfaktora na istoj razini glavnog faktora i za usporedbe razina glavnog faktora na istoj ili različitim razinama podfaktora. Način procjene pogrešaka za ova dva tipa testova navodi Kuehl (2000) na str. 478-479.

### Primjer 6.

U svrhu istraživanja utjecaja četiri razine gnojidbe dušikom (83, 124, 133 i 151 kg/ha) na urod i kvalitetu zrna deset sorata jare zobi, proveden je pokus na pokusnom polju Bc Instituta u Botincu 2014. godine (Bukan i sur. 2015). Pokus je postavljen kao split-

plot s četiri repeticije, s gnojidbom kao glavnim i sortama kao podfaktorom. U ovom će se primjeru analizirati podatci o visini biljke (cm).

### 1. Učitavanje podataka

U prvom koraku učitati će se podatci iz priložene datoteke u csv formatu (uz pretpostavku da je kao radni direktorij prethodno postavljen direktorij u kojem se nalazi datoteka s podacima):

```
pr6 <- read.csv("Primjer6.csv",  
               colClasses = c("factor", "factor", "factor", "numeric"))
```

Argument `colClasses` koristi se kako bi se varijable `gnojdba`, `rep` i `sorta` učitale kao faktori, jer su sve tri označene brojevima.

### 2. Analiza prema mješovitom modelu

Najprije je potrebno aktivirati pakete `lme4` i `lmerTest`:

```
library(lme4)  
library(lmerTest)
```

Analiza se provodi primjenom funkcije `lmer`, a model uključuje jedan slučajni efekt, interakciju gnojidbi i repeticija, koja služi kao pogreška za test učinaka gnojidbi i repeticija:

```
visina.lme <-  
  lmer(visina~gnojdba+rep+(1|gnojdba:rep)+sorta+gnojdba:sorta,data=pr6)
```

U ovom se primjeru ne koristi skraćeni (uz pomoć simbola `*`), nego puni zapis kojim se definiraju učinci faktora i njihove interakcije. Slučajni efekt je postavljen u sredinu modela jer predstavlja pogreška za fiksne efekte koji se nalaze ispred njega. Navedeni će se redoslijed efekata zadržati i u ispisu rezultata dobivenom primjenom generičke funkcije `anova`, pa će se jasno razdvojiti testovi provedeni s različitim pogreškama:

```
anova(visina.lme)  
Analysis of Variance Table of type III with Satterthwaite  
approximation for degrees of freedom
```

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
gnojidba	1915.5	638.49	3	9	12.2396	0.001579	**
rep	4626.3	1542.11	3	9	29.5615	5.440e-05	***
sorta	4039.9	448.88	9	108	8.6049	1.238e-09	***
gnojidba:sorta	1835.3	67.97	27	108	1.3030	0.171171	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Svi su efekti u modelu signifikantni, osim učinka interakcije. Zbog toga su u ovom primjeru učinci dvaju faktora aditivni, pa se višestruke usporedbe mogu provesti zasebno za učinke gnojidbi i sorata.

### 3. Višestruke usporedbe

Početni korak je aktivacija paket `multcomp`:

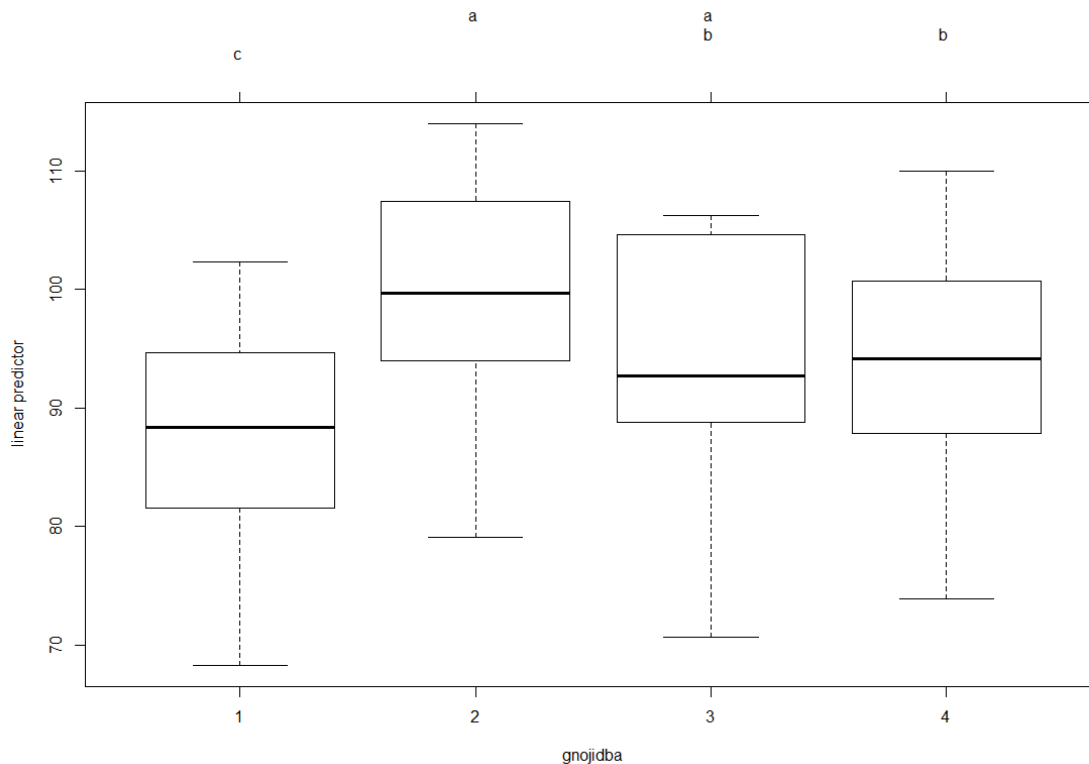
```
library(multcomp)
```

Razlike između četiri razine gnojidbe, kao i pogreške tih razlika procjenjuju se primjenom funkcije `glht`:

```
gnojidba.glht <- glht(visina.lme, linfct=mcp(gnojidba="Tukey",
interaction_average = T, covariate_average = T))
```

Izmjena dviju zadanih postavki (`interaction_average`, `covariate_average`) nužna je zbog ispravne procjene razlika. Ukupno se provodi šest usporedbi, pa primjena korekcije neće imati velikog utjecaja na ishode testova. Zbog se grafički može prikazati izravno rezultat primjene funkcije `glht`:

```
old.par <- par(mai=c(1,1,1.25,1), no.readonly = TRUE)
plot(cld(gnojidba.glht, decreasing=T))
par(old.par)
```



Slika 15. Grafički prikaz testa razlika u visinama biljaka (cm) jare zobi ostvarenim uz četiri razine gnojidbe dušikom

Biljke su bile najniže uz najnižu razinu gnojidbe, a najviše uz dodanih 124, odnosno 133 kg/ha. Isti se postupak može primijeniti i za testove razlika između sorata:

```
sorta.glht <- glht(visina.lme, linfct=mcp(sorta="Tukey",
  interaction_average = T, covariate_average = T))
```

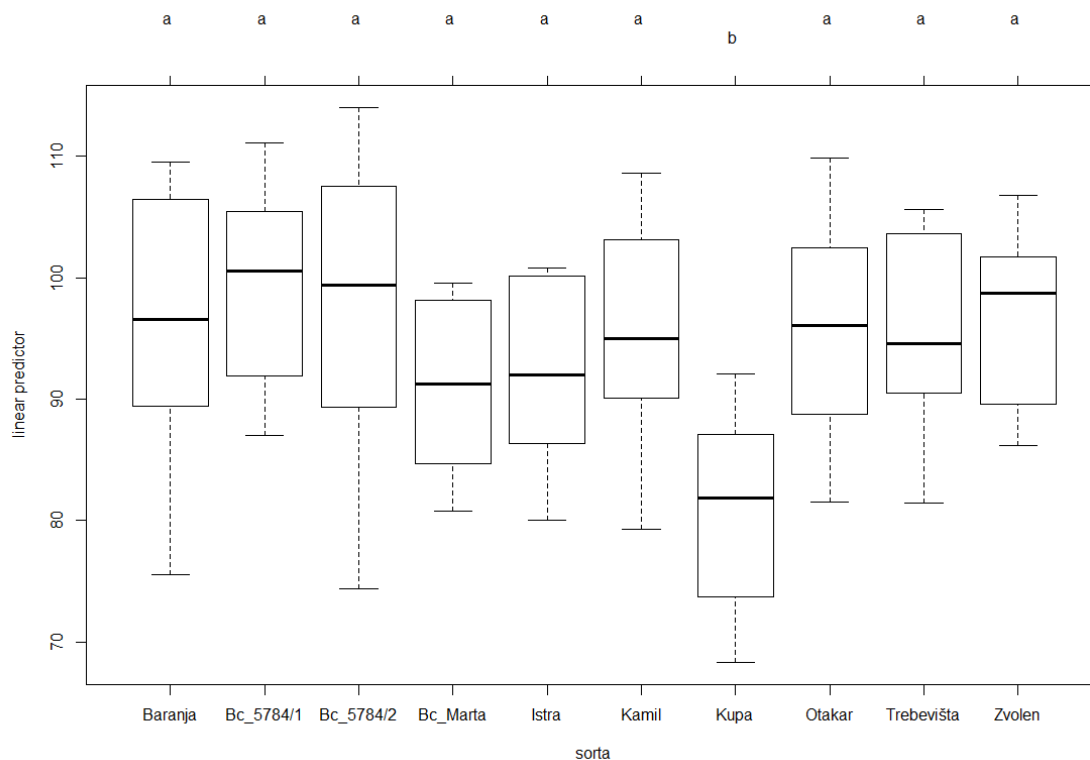
Osim što je zbog velikog broja testova potrebna korekcija:

```
sorta.mc <- summary(sorta.glht, test=adjusted("bon"))
```

Grafički prikaz rezultata korigiranih testova:

```
old.par <- par(mai=c(1,1,1.1,1), no.readonly = TRUE)
plot(cld(sorta.mc, decreasing=T))
par(old.par)
```





Slika 16. Grafički prikaz testa razlika u visinama biljaka (cm) deset sorata jare zobi

Sorta Kupa niža je od svih ostalih sorata, između kojih nema razlika u visini. U ovom primjeru učinak interakcije nije signifikantan, pa nema potrebe za provođenjem testova opisanih u uvodnom dijelu. Kada je to potrebno, ove testove je umjesto uz pomoć funkcije `glht` jednostavnije provesti primjenom funkcije `emmeans` iz istoimenog paketa.

## Literatura

1. Kuehl R.O. (2000) Design of Experiments: Statistical Principles of Research Design and Analysis. Duxbury Press, Pacific Grove, CA.
2. Vasilj Đ. (2000) Biometrika i eksperimentiranje u bilinogojstvu. Hrvatsko agronomsko društvo, Zagreb.
3. Bukan M., Maričević M., Ikić I., Mlinar R., Lovrić A., Gunjača J., Šarčević H. (2015) Utjecaj gnojidbe dušikom na prinos i kvalitetu zrna jare zobi pljevičastog i golog zrna. Poljoprivreda 21: 15-21.

4. Bates D., Maechler M., Bolker B., Walker S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1-48.
5. Kuznetsova A., Brockhoff P.B., Christensen R.H.B. (2017) lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26
6. Bretz F., Hothorn T., Westfall P. (2010) *Multiple Comparisons Using R*. Chapman and Hall/CRC, Boca Raton, FL.

## Dodatak A: Modeli

Model	Jednadžba		R kod	
	broj	izraz	puni	skraćeni
Potpuno slučajni raspored	( 2 )	$Y=TRET$		$Y\sim TRET$
Dvofaktorijelni unakrsni fiksni model	( 4 )	$Y=A+B+A\cdot B$	$Y\sim A+B+A:B$	$Y\sim A*B$
Dvofaktorijelni hijerarhijski model	( 5 )	$Y=A+A\cdot B$	$Y\sim A+A:B$	$Y\sim A/B$
Dvofaktorijelni unakrsni mješoviti model	( 6 )	$Y=A::B+A\cdot B$		$Y\sim A+(1/B)+(1/A:B)$
Slučajni blokni raspored	( 7 )	$Y=TRET+REP$		$Y\sim TRET+REP$
Latinski kvadrat	( 8 )	$Y=TRET+VREP+O.REP$		$Y\sim TRET+V.REP+O.REP$
Alfa dizajn	( 12 )	$Y=TRET+REP::REP:BLOK$		$Y\sim TRET+REP+(1 REP:BLOK)$
Redno-stupčani dizajn	( 14 )	$Y=TRET+REP::REP:RED+REP:STUPAC$		$Y\sim TRET+REP+(1 REP:RED)+(1 REP:STUPAC)$
Latinizirani alfa dizajn	( 16 )	$Y=TRET+REP+DBLOK::REP:BLOK$		$Y\sim TRET+REP+DBLOK+(1 REP:BLOK)$
Split-plot	( 17 )	$Y=A+REP+B+A\cdot B::A\cdot REP$	$Y\sim A+REP+B+A:B+(1 A:REP)$	$Y\sim A*B+REP+(1 A:REP)$

# Dodatak B: Primjeri

## Primjer 1.1

broj mahuna	broj zrna								
37,90	40,93	22,54	26,58	25,56	11,27	26,66	15,27	35,88	
22,52	25,56	34,91	47,98	9,16	38,85	21,31	13,33	13,30	
13,33	51,121	47,116	24,47	12,28	24,48	43,98	25,53	30,60	
22,53	20,48	39,95	29,61	29,68	17,43	22,56	19,43	17,39	
25,57	46,93	24,59	53,125	10,24	14,34	22,50	12,32	26,58	
28,68	26,50	19,51	40,82	9,20	21,45	19,45	59,127	34,80	
13,33	24,53	48,118	24,49	18,40	35,85	33,85	37,71	34,76	
61,148	24,49	52,123	50,108	7,15	23,59	31,72	28,63	72,51	
21,52	76,187	59,150	55,129	11,18	18,46	17,36	41,89	51,130	
43,114	49,124	52,130	18,40	28,56	37,99	28,68	47,113	24,53	
29,74	51,119	32,80	54,125	17,36	10,23	35,81	38,91	44,89	
8,16	15,37	28,65	50,100	12,24	32,80	36,93	36,92	22,44	
47,109	18,41	22,52	30,69	37,70	23,54	30,72	30,62	34,84	
27,56	33,80	45,125	64,149	19,45	24,59	7,13	26,66	38,74	
22,53	17,37	18,40	44,104	43,91	13,32	20,53	36,86	34,82	
17,33	34,82	35,87	27,49	7,9	15,32	50,107	41,101	16,35	
29,61	37,84	85,205	76,193	9,17	24,50	26,63	32,74	17,38	
23,55	12,30	17,36	38,79	14,31	30,64	22,56	33,73	27,67	
33,84	20,42	35,77	30,57	19,50	18,25	36,95	36,80	7,17	
28,68	46,116	24,56	23,54	73,189	37,98	41,103	25,57	48,111	
15,36	38,90	49,110	33,72	26,61	12,28	65,152	20,37	37,86	
10,20	39,89	30,66	29,56	54,132	23,50	25,53	53,113	21,57	
26,67	43,100	34,55	63,141	16,33	21,47	22,51	21,49	17,39	
19,47	39,83	8,14	86,207	44,89	18,47	30,57	26,48	31,62	
21,58	40,87	30,45	30,64	69,157	23,49	36,90	21,52	28,70	
34,74	28,65	34,82	24,58	36,83	26,58	19,49	19,47	19,47	
29,74	41,106	30,70	30,75	25,57	23,58	18,33	28,70	48,124	
30,76	31,82	53,132	23,52	11,26	47,104	19,47	37,75	21,49	
31,77	26,62	55,138	29,64	20,49	19,44	56,127	25,68	37,64	
33,75	15,35	50,111	37,82	26,57	29,68	25,58	35,73	28,53	
17,45	18,44	24,45	37,72	16,37	27,64	17,46	23,52	25,54	
20,48	33,84	19,38	14,25	10,16	34,81	56,120	16,28	40,97	
16,39	24,60	39,81	37,86	24,61	16,41	45,108	25,57	31,71	
11,19	18,48	26,55	53,137	33,82	30,68	29,51	35,86	15,27	
60,130	39,94	17,28	27,62	32,70	30,62	36,91	25,62	9,19	
25,58	21,50	31,67	27,60	45,113	13,26	26,55	25,53	34,81	
34,76	39,95	16,39	24,57	29,72	30,71	26,64	15,37	11,26	
28,67	23,53	36,97	21,50	43,113	36,86	18,43	35,87	34,52	
35,97	33,70	10,21	23,46	13,30	54,141	33,79	26,60	22,55	
10,20	23,54	32,78	27,54	26,61	32,79	8,20	33,71	16,39	
12,30	19,41	32,78	17,39	28,73	16,42	20,40	29,69	35,78	
14,31	26,49	49,126	9,14	56,140	21,44	34,77	27,58	37,86	
17,35	14,29	17,36	21,40	25,57	23,58	35,81	34,79	28,59	
14,37	38,95	17,39	18,45	42,90	31,72	27,63	27,63	29,67	
21,50	26,61	58,148	35,67	31,82	23,51	49,130	20,44	22,55	
23,57	20,48	18,43	48,102	41,104	36,93	81,190	12,28	18,45	
10,23	45,117	30,74	21,55	13,27	9,19	15,36	25,60	29,69	
25,49	40,90	28,68	59,156	16,40	43,99	17,31	21,50	19,43	
42,	67,174	31,72	27,61	28,70	48,119	30,73	40,98	22,60	
31,72	32,55	38,82	25,52	24,51	11,25	22,51	32,71	42,96	
6,13	30,76	40,90	15,33	15,30	19,37	12,29	41,89	20,54	
18,32	25,59	21,54	32,78	24,58	23,50	31,65	36,78	37,83	
24,51	37,54	27,75	33,83	23,59	18,51	34,97	14,35	30,57	
60,150	28,70	29,67	22,51	22,47	31,73	21,47	60,164	18,44	
38,110	34,84	58,148	37,86	29,69	31,65	31,59	17,41	24,46	
10,24	22,55	37,111	40,86	61,150	26,67	18,43	28,65	41,101	
13,31	37,86	15,30	43,108	39,94	33,85	18,37	47,110	21,42	
24,42	14,35	44,96	30,68	32,75	30,64	25,55	32,78		
25,69	12,29	74,152	17,39	33,81	15,38	27,58	15,34		
21,49	71,168	28,56	30,60	21,52	41,109	31,68	29,64		
27,59	33,76	20,44	27,54	19,50	22,52	19,46	20,46		
15,26	46,124	35,50	24,64	29,73	43,95	11,21	35,92		
30,68	39,99	34,73	33,65	26,56	33,81	35,79	19,47		

## Primjer 1.2

tjedan ,masa						
1,49	2,107	4,373	6,750	8,1430	9,1390	11,2790
1,60	2,95	4,396	6,720	8,1330	9,1520	12,3400
1,61	2,97	4,340	6,810	8,1190	10,1760	12,6120
1,54	3,158	4,312	6,740	8,1230	10,2700	12,6040
1,62	3,183	4,307	6,773	8,1130	10,3100	12,5860
1,63	3,178	4,362	6,697	8,1320	10,3070	12,5400
1,63	3,187	5,389	6,726	8,1240	10,2600	12,5350
1,69	3,183	5,617	6,630	8,1290	10,2500	12,5400
1,67	3,198	5,620	6,770	8,1200	10,3000	12,5640
1,67	3,185	5,511	7,710	8,1220	10,3160	12,5380
1,52	3,210	5,447	7,1120	8,1080	10,3270	12,3940
1,57	3,219	5,542	7,1120	8,1190	10,2940	13,4120
1,54	3,213	5,559	7,954	9,1240	10,1900	13,7000
2,76	3,212	5,542	7,930	9,1390	10,2280	13,7140
2,103	3,183	5,592	7,960	9,1660	11,2700	13,6740
2,106	3,186	5,522	7,1020	9,1500	11,4220	13,6300
2,101	4,275	5,518	7,1030	9,1650	11,4440	13,6340
2,90	4,453	5,430	7,1058	9,1380	11,4250	13,6500
2,117	4,412	5,562	7,970	9,1700	11,4060	13,6350
2,115	4,316	6,560	7,1015	9,1560	11,3800	13,6300
2,127	4,327	6,900	7,810	9,1590	11,3800	13,5760
2,118	4,369	6,822	7,970	9,1460	11,4200	
2,120	4,417	6,698	8,960	9,1530	11,3860	

## Primjer 1.3

gnojidba,visina			
G1,27	G1,30	G2,35	G3,42
G1,32	G1,30	G2,36	G3,35
G1,36	G2,30	G2,35	G3,38
G1,32	G2,35	G2,40	G3,39
G1,35	G2,32	G3,50	G3,36
G1,32	G2,32	G3,48	G3,44
G1,33	G2,39	G3,35	
G1,35	G2,31	G3,31	

## Primjer 2.

podloga,ukk			
Paulsen,9.2	41B,9.8	Kober,9.8	Fercal,9
Paulsen,10.1	41B,9.9	Kober,8.3	Fercal,9
Paulsen,9.1	41B,10.5	Kober,9.1	Fercal,9.2
Paulsen,8.8	41B,10.1	Kober,8.9	Fercal,9.3

## Primjer 3.1

godina, podloga, urod

2000, Kober_St, 3	2000, SO4_15, 2.3	2000, SO4_102, 2.4	2000, Riparia, 0.5
2000, Kober_St, 1.3	2000, SO4_15, 1.6	2000, SO4_102, 1.2	2000, Riparia, 0.4
2000, Kober_St, 3.5	2000, SO4_15, 2.3	2000, SO4_102, 1.7	2000, Riparia, 1.2
2000, Kober_St, 1.6	2000, SO4_15, 1.6	2000, SO4_102, 2.1	2000, Riparia, 2.3
2000, Kober_St, 0.6	2000, SO4_15, 3	2000, SO4_102, 1.5	2000, Riparia, 0.1
2000, Kober_St, 1.6	2000, SO4_15, 0.9	2000, SO4_102, 3.2	2000, Riparia, 1.8
2000, Kober_St, 2.2	2000, SO4_15, 1.5	2000, SO4_102, 1.2	2000, Riparia, 0.9
2000, Kober_St, 3	2000, SO4_15, 3.1	2000, SO4_102, 2.4	2000, Riparia, 0.6
2000, Kober_St, 2.2	2000, SO4_15, 2.1	2000, SO4_102, 2.5	2000, Riparia, 1.2
2000, Kober_St, 1.9	2000, SO4_15, 1.4	2000, SO4_102, 1.3	2000, Riparia, 0.4
2000, Kober_St, 3.2	2000, SO4_15, 0.9	2000, Kober_Gm, 2.1	2000, Riparia, 0.7
2000, Kober_St, 1.9	2000, SO4_15, 1.8	2000, Kober_Gm, 2.8	2000, Riparia, 1.6
2000, Kober_St, 1.9	2000, SO4_15, 2.1	2000, Kober_Gm, 2.9	2000, Riparia, 1.5
2000, Kober_St, 1.8	2000, SO4_15, 2.1	2000, Kober_Gm, 1.8	2000, Riparia, 3.9
2000, Kober_St, 1.5	2000, SO4_15, 1.7	2000, Kober_Gm, 1.3	2000, Riparia, 1
2000, Kober_St, 0.4	2000, SO4_15, 1	2000, Kober_Gm, 3.7	2000, Riparia, 1
2000, Kober_St, 2.1	2000, SO4_15, 2	2000, Kober_Gm, 2.2	2000, Riparia, 2.4
2000, Kober_St, 2.9	2000, SO4_15, 2.7	2000, Kober_Gm, 3.5	2000, Riparia, 1.2
2000, Kober_St, 0.4	2000, SO4_15, 1.6	2000, Kober_Gm, 2.1	2000, Riparia, 2.1
2000, Kober_St, 1.2	2000, SO4_15, 1.9	2000, Kober_Gm, 1.1	2000, Riparia, 2.3
2000, Kober_St, 1.4	2000, SO4_31, 3.3	2000, Kober_Gm, 2.9	2000, Riparia, 1.4
2000, Kober_St, 2.2	2000, SO4_31, 1.2	2000, Kober_Gm, 2.7	2000, Riparia, 1.6
2000, Kober_St, 1.7	2000, SO4_31, 1.5	2000, Kober_Gm, 1.4	2000, Riparia, 1.3
2000, Kober_St, 1	2000, SO4_31, 1.5	2000, Kober_Gm, 2.9	2000, Riparia, 1.2
2000, Kober_St, 1	2000, SO4_31, 3.6	2000, Kober_Gm, 1.8	2000, Riparia, 2.9
2000, Kober_St, 2.1	2000, SO4_31, 1.5	2000, Kober_Gm, 0.6	2000, Riparia, 1.9
2000, Kober_St, 2.5	2000, SO4_31, 1.8	2000, Kober_Gm, 1.4	2000, Riparia, 1.4
2000, Kober_St, 1.4	2000, SO4_31, 3.1	2000, Kober_Gm, 2	2000, Riparia, 0.4
2000, Kober_St, 1.3	2000, SO4_31, 1.9	2000, Kober_Gm, 2.2	2000, Riparia, 0.6
2000, Kober_St, 2.9	2000, SO4_31, 1.8	2000, Kober_Gm, 1.7	2000, Riparia, 1.1
2000, SO4_5, 1.2	2000, SO4_31, 2.1	2000, Kober_Gm, 2	2000, 19617, 3
2000, SO4_5, 2.4	2000, SO4_31, 4	2000, Kober_Gm, 0.5	2000, 19617, 2
2000, SO4_5, 2.7	2000, SO4_31, 3.3	2000, Kober_Gm, 1.8	2000, 19617, 1.7
2000, SO4_5, 2.4	2000, SO4_31, 3.6	2000, Kober_Gm, 1.2	2000, 19617, 2.2
2000, SO4_5, 3.1	2000, SO4_31, 1.4	2000, Kober_Gm, 2.2	2000, 19617, 1.5
2000, SO4_5, 2.4	2000, SO4_31, 2.3	2000, Kober_Gm, 3.1	2000, 19617, 2
2000, SO4_5, 0.8	2000, SO4_31, 2.4	2000, Kober_Gm, 2.2	2000, 19617, 2
2000, SO4_5, 2.2	2000, SO4_31, 3.9	2000, Kober_Gm, 4.1	2000, 19617, 2.3
2000, SO4_5, 1.6	2000, SO4_31, 3.1	2000, Kober_Gm, 2.9	2000, 19617, 2
2000, SO4_5, 3.1	2000, SO4_31, 2.6	2000, Kober_Gm, 2.5	2000, 19617, 0.8
2000, SO4_5, 2.1	2000, SO4_31, 2.2	2000, Boerner, 2.3	2000, 19617, 1.4
2000, SO4_5, 2.4	2000, SO4_31, 3.8	2000, Boerner, 0.7	2000, 19617, 0.5
2000, SO4_5, 1.6	2000, SO4_31, 2.4	2000, Boerner, 1.9	2000, 19617, 1.5
2000, SO4_5, 2.4	2000, SO4_31, 2.4	2000, Boerner, 2.8	2000, 19617, 0.4
2000, SO4_5, 2.1	2000, SO4_31, 1.8	2000, Boerner, 2.4	2000, 19617, 1
2000, SO4_5, 0.8	2000, SO4_31, 0.6	2000, Boerner, 1.2	2000, 19617, 2.2
2000, SO4_5, 2.9	2000, SO4_31, 2.6	2000, Boerner, 1.5	2000, 19617, 1.8
2000, SO4_5, 1.4	2000, SO4_31, 1.5	2000, Boerner, 3.5	2000, 19617, 2.2
2000, SO4_5, 1.3	2000, SO4_31, 2.4	2000, Boerner, 2.7	2000, 19617, 1.1
2000, SO4_5, 4.5	2000, SO4_31, 2.2	2000, Boerner, 2.9	2000, 19617, 3
2000, SO4_5, 1.6	2000, SO4_102, 1.2	2000, Boerner, 2.7	2000, 19617, 0.7
2000, SO4_5, 0.3	2000, SO4_102, 3.2	2000, Boerner, 1.8	2000, 19617, 1.9
2000, SO4_5, 3.8	2000, SO4_102, 2.2	2000, Boerner, 2	2000, 19617, 2.3
2000, SO4_5, 3.4	2000, SO4_102, 1.5	2000, Boerner, 3.4	2000, 19617, 1.9
2000, SO4_5, 2.5	2000, SO4_102, 3	2000, Boerner, 2.4	2000, 19617, 1.2
2000, SO4_5, 2.7	2000, SO4_102, 2	2000, Boerner, 2.1	2000, 19617, 1.7
2000, SO4_5, 1.4	2000, SO4_102, 1.4	2000, Boerner, 1.6	2000, 19617, 0.9
2000, SO4_5, 1.2	2000, SO4_102, 1.5	2000, Boerner, 0.5	2000, 19617, 1.1
2000, SO4_5, 3.3	2000, SO4_102, 2.1	2000, Boerner, 1.8	2000, 19617, 2.1
2000, SO4_5, 0.6	2000, SO4_102, 2.4	2000, Boerner, 0.6	2000, 19617, 0.4
2000, SO4_15, 1.8	2000, SO4_102, 3.4	2000, Boerner, 1.7	2001, Kober_St, 4.9
2000, SO4_15, 2.1	2000, SO4_102, 2.6	2000, Boerner, 2.3	2001, Kober_St, 1.4
2000, SO4_15, 4.4	2000, SO4_102, 1.1	2000, Boerner, 1.9	2001, Kober_St, 4.5
2000, SO4_15, 2.9	2000, SO4_102, 0.7	2000, Boerner, 0.2	2001, Kober_St, 2
2000, SO4_15, 2.9	2000, SO4_102, 2	2000, Boerner, 1.9	2001, Kober_St, 1.3
2000, SO4_15, 0.5	2000, SO4_102, 1.7	2000, Boerner, 2.7	2001, Kober_St, 4
2000, SO4_15, 2.5	2000, SO4_102, 2.1	2000, Boerner, 2.1	2001, Kober_St, 3.1
2000, SO4_15, 1.9	2000, SO4_102, 3.6	2000, Boerner, 2.7	2001, Kober_St, 3
2000, SO4_15, 3.6	2000, SO4_102, 2.2	2000, Boerner, 1.3	2001, Kober_St, 2.1
2000, SO4_15, 2	2000, SO4_102, 3.7	2000, Boerner, 2.2	2001, Kober_St, 2.8

2001,Kober_St,3.5	2001,SO4_15,3.2	2001,SO4_102,3.4	2001,Boerner,3.6
2001,Kober_St,3.4	2001,SO4_15,2.6	2001,SO4_102,2.1	2001,Boerner,2.5
2001,Kober_St,4.4	2001,SO4_15,3.5	2001,SO4_102,2.7	2001,Boerner,1.8
2001,Kober_St,2.9	2001,SO4_15,2.7	2001,SO4_102,1.3	2001,Boerner,2.3
2001,Kober_St,2.8	2001,SO4_15,3.2	2001,SO4_102,3.2	2001,Boerner,1.7
2001,Kober_St,2.9	2001,SO4_15,2.3	2001,SO4_102,4	2001,Riparia,2
2001,Kober_St,2.5	2001,SO4_15,2.9	2001,SO4_102,1.5	2001,Riparia,0.9
2001,Kober_St,2.4	2001,SO4_15,3.2	2001,SO4_102,2.4	2001,Riparia,1.4
2001,Kober_St,2	2001,SO4_15,3.2	2001,SO4_102,4.5	2001,Riparia,1.8
2001,Kober_St,3	2001,SO4_15,3.7	2001,SO4_102,2.4	2001,Riparia,2.3
2001,Kober_St,1.2	2001,SO4_15,3.6	2001,Kober_Gm,3.8	2001,Riparia,2.5
2001,Kober_St,3.2	2001,SO4_15,2.7	2001,Kober_Gm,3.8	2001,Riparia,2
2001,Kober_St,4.3	2001,SO4_15,2	2001,Kober_Gm,3	2001,Riparia,2
2001,Kober_St,2.9	2001,SO4_15,1.9	2001,Kober_Gm,1.5	2001,Riparia,2.4
2001,Kober_St,2.9	2001,SO4_15,2.3	2001,Kober_Gm,3.6	2001,Riparia,2
2001,Kober_St,3.4	2001,SO4_31,4.1	2001,Kober_Gm,2.7	2001,Riparia,2
2001,Kober_St,1.9	2001,SO4_31,1.5	2001,Kober_Gm,2.5	2001,Riparia,1
2001,Kober_St,4.4	2001,SO4_31,3.6	2001,Kober_Gm,3.3	2001,Riparia,2
2001,Kober_St,2.9	2001,SO4_31,3.5	2001,Kober_Gm,2.6	2001,Riparia,2.7
2001,Kober_St,4.5	2001,SO4_31,4.8	2001,Kober_Gm,1.5	2001,Riparia,1.2
2001,SO4_5,0.5	2001,SO4_31,2.6	2001,Kober_Gm,3.7	2001,Riparia,3.6
2001,SO4_5,1.6	2001,SO4_31,1.2	2001,Kober_Gm,4.5	2001,Riparia,3.6
2001,SO4_5,4.1	2001,SO4_31,3.1	2001,Kober_Gm,3.9	2001,Riparia,1.6
2001,SO4_5,4	2001,SO4_31,2.9	2001,Kober_Gm,3.7	2001,Riparia,2.5
2001,SO4_5,3.1	2001,SO4_31,2.8	2001,Kober_Gm,2.7	2001,Riparia,3.1
2001,SO4_5,2.2	2001,SO4_31,3	2001,Kober_Gm,3	2001,Riparia,2
2001,SO4_5,4.2	2001,SO4_31,4.3	2001,Kober_Gm,4.3	2001,Riparia,0.7
2001,SO4_5,2.3	2001,SO4_31,2.4	2001,Kober_Gm,2.7	2001,Riparia,2.5
2001,SO4_5,2.8	2001,SO4_31,3	2001,Kober_Gm,2.1	2001,Riparia,2.5
2001,SO4_5,2.8	2001,SO4_31,3.6	2001,Kober_Gm,0.9	2001,Riparia,1.9
2001,SO4_5,3	2001,SO4_31,3.7	2001,Kober_Gm,1.6	2001,Riparia,4
2001,SO4_5,4.4	2001,SO4_31,3	2001,Kober_Gm,1.9	2001,Riparia,2
2001,SO4_5,3	2001,SO4_31,4.1	2001,Kober_Gm,1.7	2001,Riparia,2.2
2001,SO4_5,5	2001,SO4_31,5	2001,Kober_Gm,3.4	2001,Riparia,0.5
2001,SO4_5,2.9	2001,SO4_31,5.7	2001,Kober_Gm,3	2001,Riparia,2
2001,SO4_5,4	2001,SO4_31,1.8	2001,Kober_Gm,4.3	2001,19617,2.6
2001,SO4_5,2.6	2001,SO4_31,3.6	2001,Kober_Gm,3	2001,19617,1.6
2001,SO4_5,0.9	2001,SO4_31,3	2001,Kober_Gm,4	2001,19617,2.4
2001,SO4_5,1.8	2001,SO4_31,3	2001,Kober_Gm,3.6	2001,19617,3.7
2001,SO4_5,3	2001,SO4_31,2	2001,Kober_Gm,4.9	2001,19617,1.8
2001,SO4_5,1.7	2001,SO4_31,1.2	2001,Boerner,1.3	2001,19617,1.7
2001,SO4_5,2.9	2001,SO4_31,2.3	2001,Boerner,0.3	2001,19617,3
2001,SO4_5,4.2	2001,SO4_31,2.3	2001,Boerner,1.7	2001,19617,3.1
2001,SO4_5,2.2	2001,SO4_31,3	2001,Boerner,3.5	2001,19617,1.6
2001,SO4_5,4.3	2001,SO4_31,0	2001,Boerner,3	2001,19617,2.4
2001,SO4_5,3.4	2001,SO4_102,3.1	2001,Boerner,1.9	2001,19617,2.8
2001,SO4_5,1.5	2001,SO4_102,3.8	2001,Boerner,1.2	2001,19617,2.4
2001,SO4_5,2.8	2001,SO4_102,2.1	2001,Boerner,2.3	2001,19617,1.8
2001,SO4_5,4.2	2001,SO4_102,1.1	2001,Boerner,3.1	2001,19617,1.3
2001,SO4_5,1.4	2001,SO4_102,3.7	2001,Boerner,4.4	2001,19617,2.1
2001,SO4_15,2.2	2001,SO4_102,1.9	2001,Boerner,2	2001,19617,3.4
2001,SO4_15,3.2	2001,SO4_102,4	2001,Boerner,3.1	2001,19617,1.9
2001,SO4_15,4.3	2001,SO4_102,1.6	2001,Boerner,1	2001,19617,3.1
2001,SO4_15,3.9	2001,SO4_102,4.3	2001,Boerner,4.1	2001,19617,3.4
2001,SO4_15,4.6	2001,SO4_102,3.3	2001,Boerner,2.8	2001,19617,3.7
2001,SO4_15,3.2	2001,SO4_102,4.8	2001,Boerner,1.8	2001,19617,2.4
2001,SO4_15,3.2	2001,SO4_102,4.6	2001,Boerner,2.8	2001,19617,2
2001,SO4_15,6.2	2001,SO4_102,3.1	2001,Boerner,1.7	2001,19617,4.1
2001,SO4_15,4.2	2001,SO4_102,3	2001,Boerner,1.8	2001,19617,1.2
2001,SO4_15,3.5	2001,SO4_102,2.7	2001,Boerner,3.5	2001,19617,1.4
2001,SO4_15,4.5	2001,SO4_102,1.4	2001,Boerner,3.9	2001,19617,2.7
2001,SO4_15,1.4	2001,SO4_102,2.9	2001,Boerner,4.1	2001,19617,3.4
2001,SO4_15,2	2001,SO4_102,3.1	2001,Boerner,3	2001,19617,3.4
2001,SO4_15,4.3	2001,SO4_102,3.4	2001,Boerner,2	2001,19617,1.9
2001,SO4_15,3.5	2001,SO4_102,3	2001,Boerner,5.4	

## Primjer 3.2

lokalitet, stablo, sirina\_lista

Karlovac,1,54.5	Tomislavgrad,1,47.55	Gospic,1,42.13
Karlovac,1,57.33	Tomislavgrad,1,40.17	Gospic,1,37.2
Karlovac,1,62.41	Tomislavgrad,1,51.77	Gospic,1,41.62
Karlovac,1,40.27	Tomislavgrad,1,53.02	Gospic,1,35.61
Karlovac,1,43.17	Tomislavgrad,1,42.22	Gospic,1,38.74
Karlovac,2,51.55	Tomislavgrad,2,47.47	Gospic,2,53.29
Karlovac,2,49.63	Tomislavgrad,2,54.87	Gospic,2,51.61
Karlovac,2,36.6	Tomislavgrad,2,39.84	Gospic,2,55.77
Karlovac,2,42.12	Tomislavgrad,2,44.09	Gospic,2,53.67
Karlovac,2,41.21	Tomislavgrad,2,60.86	Gospic,2,48.88
Karlovac,3,51.28	Tomislavgrad,7,54.01	Gospic,3,41.15
Karlovac,3,45.16	Tomislavgrad,7,53.54	Gospic,3,35.23
Karlovac,3,48.62	Tomislavgrad,7,43.16	Gospic,3,40.25
Karlovac,3,44.08	Tomislavgrad,7,44.04	Gospic,3,41.47
Karlovac,3,43.41	Tomislavgrad,7,51.56	Gospic,3,43.22
Karlovac,4,46.4	Tomislavgrad,5,46.65	Gospic,4,40.49
Karlovac,4,43.8	Tomislavgrad,5,47.23	Gospic,4,47.12
Karlovac,4,46.28	Tomislavgrad,5,46.41	Gospic,4,47.68
Karlovac,4,43.64	Tomislavgrad,5,48.79	Gospic,4,47.89
Karlovac,4,42.91	Tomislavgrad,5,50.18	Gospic,4,47.31
Karlovac,5,54.16	Tomislavgrad,6,62.36	Gospic,6,37.78
Karlovac,5,55.01	Tomislavgrad,6,53.4	Gospic,6,38.63
Karlovac,5,52.32	Tomislavgrad,6,49.67	Gospic,6,43.24
Karlovac,5,51.88	Tomislavgrad,6,49.54	Gospic,6,52.13
Karlovac,5,49.04	Tomislavgrad,6,54.86	Gospic,6,45.67

## Primjer 4.1

sorta, rep, urod

1,1,68.86046512	2,2,70.65116279	3,3,70.48837209	4,4,65.94418605
2,1,74.68837209	3,2,74.51627907	4,3,73.70232558	1,5,71.81162791
3,1,76.53023256	4,2,57.21627907	1,4,63.94186047	2,5,73.67906977
4,1,73.70232558	1,3,60.99069767	2,4,81.75348837	3,5,71.49534884
1,2,59.02325581	2,3,68.63255814	3,4,74.51627907	4,5,68.85348837

## Primjer 4.2

sorta, repV, repO, urod

1,1,1,60.55813953	3,2,4,56.95069767	5,3,2,60.64186047	2,5,3,60.76046512
2,1,2,70.8872093	4,2,1,70.24418605	1,4,3,60.55813953	3,5,1,56.95069767
3,1,3,71.18837209	5,2,3,56.59906977	2,4,4,70.8872093	4,5,2,60.20930233
4,1,4,70.24418605	1,3,4,50.46511628	3,4,2,77.29023256	5,5,4,64.68465116
5,1,5,80.85581395	2,3,1,70.8872093	4,4,5,84.29302326	
1,2,2,50.46511628	3,3,5,61.01860465	5,4,1,60.64186047	
2,2,5,85.06465116	4,3,3,70.24418605	1,5,5,74.68837209	



## Primjer 5.1

hibrid, rep, blok, urod

1,1,3,18.89	5,1,3,20.78	9,1,3,25.44	13,1,4,19.92	17,1,4,33.92
1,2,2,25.45	5,2,4,35.17	9,2,3,28.63	13,2,4,19.75	17,2,4,34.18
1,3,3,19.85	5,3,1,36.74	9,3,4,31.08	13,3,2,29.23	17,3,3,29
1,4,3,17.77	5,4,1,29.11	9,4,4,26.14	13,4,4,21.62	17,4,2,35.49
2,1,1,23.49	6,1,4,29.65	10,1,3,11.62	14,1,2,29.61	18,1,1,29.86
2,2,2,26.36	6,2,2,29.77	10,2,4,20.26	14,2,3,29.07	18,2,2,27.44
2,3,3,19.94	6,3,2,31.75	10,3,3,17.15	14,3,1,35.1	18,3,1,35.42
2,4,2,23.83	6,4,2,33.63	10,4,4,20.73	14,4,1,30.23	18,4,2,31.45
3,1,2,21.82	7,1,1,29.82	11,1,3,18.73	15,1,2,27.87	19,1,1,33.55
3,2,1,34.64	7,2,3,20.38	11,2,2,27.98	15,2,3,24.86	19,2,4,33.31
3,3,2,34.97	7,3,2,28.09	11,3,4,19.19	15,3,1,29.43	19,3,4,33.07
3,4,3,27.4	7,4,1,25.18	11,4,2,25.84	15,4,1,32.92	19,4,4,33.1
4,1,2,24.64	8,1,4,29.38	12,1,1,31.65	16,1,2,21.51	20,1,4,27.76
4,2,1,29.34	8,2,1,28.61	12,2,1,28.31	16,2,1,20.98	20,2,3,26.9
4,3,4,26.94	8,3,4,25.98	12,3,2,33.43	16,3,1,21.27	20,3,3,25.94
4,4,3,25.47	8,4,3,27.56	12,4,3,28.76	16,4,4,21.32	20,4,1,27.53

## Primjer 5.2

hibrid, rep, duži\_red, red, stupac, visina

12,1,4,8,1,264	4,3,4,7,21,262	18,3,3,6,23,229	15,1,2,4,2,261
22,1,4,8,2,228	41,3,4,7,20,199	12,3,3,6,24,243	8,1,2,4,3,245
19,1,4,8,3,209	27,3,4,7,19,245	21,4,3,6,25,167	47,1,2,4,4,247
32,1,4,8,4,235	49,3,4,7,18,257	44,4,3,6,26,200	9,1,2,4,5,261
34,1,4,8,5,243	15,3,4,7,17,267	27,4,3,6,27,240	63,1,2,4,6,268
60,1,4,8,6,247	48,2,4,7,16,242	14,4,3,6,28,256	14,1,2,4,7,252
40,1,4,8,7,196	8,2,4,7,15,248	54,4,3,6,29,250	52,1,2,4,8,259
10,1,4,8,8,248	18,2,4,7,14,235	61,4,3,6,30,254	49,2,2,4,9,251
9,2,4,8,9,238	29,2,4,7,13,232	33,4,3,6,31,210	38,2,2,4,10,205
28,2,4,8,10,199	36,2,4,7,12,201	22,4,3,6,32,196	21,2,2,4,11,161
64,2,4,8,11,249	50,2,4,7,11,257	53,4,3,5,32,241	32,2,2,4,12,230
44,2,4,8,12,187	14,2,4,7,10,250	46,4,3,5,31,234	59,2,4,13,246
20,2,4,8,13,216	46,2,4,7,9,239	8,4,3,5,30,255	7,2,2,4,14,240
5,2,4,8,14,233	59,1,4,7,8,242	15,4,3,5,29,278	39,2,2,4,15,204
33,2,4,8,15,215	13,1,4,7,7,246	5,4,3,5,28,245	24,2,2,4,16,213
52,2,4,8,16,246	61,1,4,7,6,249	24,4,3,5,27,230	10,3,2,4,17,269
24,3,4,8,17,216	6,1,4,7,5,267	59,4,3,5,26,264	44,3,2,4,18,213
53,3,4,8,18,250	55,1,4,7,4,261	60,4,3,5,25,264	29,3,2,4,19,214
17,3,4,8,19,241	62,1,4,7,3,260	52,3,3,5,24,274	25,3,2,4,20,217
26,3,4,8,20,251	3,1,4,7,2,230	7,3,3,5,23,264	60,3,2,4,21,264
1,3,4,8,21,263	56,1,4,7,1,248	32,3,3,5,22,255	11,3,2,4,22,264
31,3,4,8,22,241	38,1,3,6,1,205	58,3,3,5,21,287	13,3,2,4,23,251
54,3,4,8,23,249	64,1,3,6,2,256	50,3,3,5,20,265	57,3,2,4,24,265
45,3,4,8,24,232	26,1,3,6,3,248	56,3,3,5,19,246	20,4,2,4,25,233
35,4,4,8,25,231	29,1,3,6,4,220	30,3,3,5,18,232	42,4,2,4,26,256
58,4,4,8,26,271	11,1,3,6,5,256	20,3,3,5,17,231	36,4,2,4,27,207
47,4,4,8,27,251	45,1,3,6,6,231	25,2,3,5,16,232	12,4,2,4,28,255
25,4,4,8,28,224	41,1,3,6,7,188	10,2,3,5,15,267	31,4,2,4,29,239
16,4,4,8,29,251	35,1,3,6,8,222	6,2,3,5,14,269	6,4,2,4,30,252
38,4,4,8,30,197	19,2,3,6,9,209	4,2,3,5,13,274	41,4,2,4,31,184
7,4,4,8,31,220	43,2,3,6,10,235	40,2,3,5,12,220	19,4,2,4,32,192
2,4,4,8,32,219	13,2,3,6,11,246	63,2,3,5,11,271	18,4,2,3,32,197
11,4,4,7,32,240	16,2,3,6,12,255	31,2,3,5,10,253	17,4,2,3,31,219
23,4,4,7,31,211	51,2,3,6,13,253	55,2,3,5,9,260	43,4,2,3,30,221
57,4,4,7,30,235	1,2,3,6,14,261	36,1,3,5,8,215	34,4,2,3,29,220
39,4,4,7,29,212	3,2,3,6,15,231	17,1,3,5,7,253	56,4,2,3,28,220
51,4,4,7,28,259	34,2,3,6,16,247	49,1,3,5,6,277	64,4,2,3,27,234
30,4,4,7,27,235	62,3,3,6,17,257	2,1,3,5,5,262	26,4,2,3,26,236
63,4,4,7,26,272	48,3,3,6,18,228	42,1,3,5,4,268	55,4,2,3,25,234
37,4,4,7,25,235	23,3,3,6,19,212	28,1,3,5,3,232	46,3,2,3,24,223
21,3,4,7,24,159	9,3,3,6,20,240	57,1,3,5,2,273	40,3,2,3,23,196
42,3,4,7,23,255	39,3,3,6,21,205	37,1,3,5,1,247	3,3,2,3,22,217
43,3,4,7,22,252	47,3,3,6,22,248	1,1,2,4,1,273	61,3,2,3,21,238

51,3,2,3,20,244	31,1,1,2,2,244	8,3,1,2,23,247	64,3,1,1,21,258
28,3,2,3,19,203	21,1,1,2,3,164	55,3,1,2,24,268	59,3,1,1,20,246
35,3,2,3,18,222	51,1,1,2,4,262	3,4,1,2,25,241	19,3,1,1,19,211
2,3,2,3,17,229	24,1,1,2,5,234	1,4,1,2,26,287	6,3,1,1,18,254
37,2,2,3,16,218	18,1,1,2,6,235	40,4,1,2,27,216	14,3,1,1,17,246
45,2,2,3,15,227	58,1,1,2,7,284	49,4,1,2,28,275	41,2,1,1,16,204
27,2,2,3,14,230	23,1,1,2,8,228	29,4,1,2,29,229	56,2,1,1,15,244
22,2,2,3,13,208	11,2,1,2,9,274	50,4,1,2,30,267	57,2,1,1,14,260
23,2,2,3,12,198	30,2,1,2,10,233	62,4,1,2,31,258	47,2,1,1,13,258
53,2,2,3,11,240	12,2,1,2,11,264	52,4,1,2,32,242	26,2,1,1,12,247
62,2,2,3,10,241	15,2,1,2,12,283	4,4,1,1,32,237	2,2,1,1,11,256
58,2,2,3,9,267	17,2,1,2,13,254	10,4,1,1,31,265	60,2,1,1,10,253
54,1,2,3,8,234	35,2,1,2,14,244	32,4,1,1,30,241	54,2,1,1,9,257
5,1,2,3,7,232	42,2,1,2,15,268	28,4,1,1,29,201	43,1,1,1,8,244
16,1,2,3,6,249	61,2,1,2,16,260	45,4,1,1,28,239	39,1,1,1,7,202
50,1,2,3,5,243	38,3,1,2,17,213	13,4,1,1,27,240	20,1,1,1,6,219
33,1,2,3,4,205	34,3,1,2,18,254	48,4,1,1,26,241	27,1,1,1,5,248
30,1,2,3,3,210	63,3,1,2,19,267	9,4,1,1,25,252	53,1,1,1,4,253
48,1,2,3,2,220	22,3,1,2,20,234	16,3,1,1,24,266	25,1,1,1,3,215
4,1,2,3,1,244	36,3,1,2,21,218	37,3,1,1,23,229	7,1,1,1,2,249
44,1,1,2,1,210	5,3,1,2,22,251	33,3,1,1,22,226	46,1,1,1,1,252

## Primjer 6.

gnojidba, rep, sorta, visina

3,1,Baranja,106	1,2,Kupa,71	2,3,Kamil,108.5	2,4,Zvolen,95
3,1,Istra,101.5	1,2,Bc_5784/2,52.5	2,3,Zvolen,105	2,4,Bc_5784/2,105
3,1,Otakar,106	1,2,Otakar,85	2,3,Baranja,108.5	2,4,Trebevišta,95
3,1,Bc_5784/2,113.5	1,2,Istra,89	2,3,Bc_Marta,90	2,4,Bc_5784/1,97.5
3,1,Kupa,89	1,2,Zvolen,95	2,3,Trebevišta,99	2,4,Kupa,82.5
3,1,Bc_Marta,99	1,2,Kamil,101	2,3,Bc_5784/2,107.5	2,4,Istra,87.5
3,1,Kamil,106	1,2,Trebevišta,81	2,3,Bc_5784/1,110	2,4,Otakar,97.5
3,1,Trebevišta,112.5	1,2,Bc_5784/1,92.5	2,3,Kupa,91	2,4,Baranja,92.5
3,1,Bc_5784/1,108.5	1,2,Bc_Marta,85	2,3,Otakar,108.5	2,4,Bc_Marta,92.5
3,1,Zvolen,105	1,2,Baranja,97.5	2,3,Istra,100	2,4,Kamil,95
2,1,Baranja,108.5	2,2,Istra,89	4,3,Kamil,100	1,4,Baranja,77.5
2,1,Bc_Marta,100	2,2,Kupa,85	4,3,Baranja,113.5	1,4,Otakar,77.5
2,1,Otakar,107.5	2,2,Trebevišta,95	4,3,Bc_Marta,95	1,4,Bc_5784/1,82.5
2,1,Kupa,86.5	2,2,Bc_5784/1,105	4,3,Otakar,105	1,4,Zvolen,80
2,1,Istra,102.5	2,2,Kamil,102.5	4,3,Bc_5784/2,115	1,4,Kamil,80
2,1,Trebevišta,105	2,2,Otakar,102.5	4,3,Trebevišta,105	1,4,Kupa,72.5
2,1,Bc_5784/2,115	2,2,Bc_Marta,87.5	4,3,Bc_5784/1,105	1,4,Bc_5784/2,85
2,1,Zvolen,108.5	2,2,Bc_5784/2,105	4,3,Kupa,82.5	1,4,Istra,82.5
2,1,Bc_5784/1,108.5	2,2,Baranja,105	4,3,Istra,105	1,4,Trebevišta,90
2,1,Kamil,105	2,2,Zvolen,95	4,3,Zvolen,105	1,4,Bc_Marta,82.5
1,1,Zvolen,106	4,2,Baranja,97.5	3,3,Zvolen,103.5	4,4,Otakar,83.5
1,1,Trebevišta,100	4,2,Istra,92.5	3,3,Trebevišta,105	4,4,Bc_Marta,90
1,1,Kupa,82.5	4,2,Bc_5784/1,100	3,3,Kupa,87.5	4,4,Kupa,75
1,1,Istra,92.5	4,2,Kupa,85	3,3,Baranja,106	4,4,Baranja,91
1,1,Bc_5784/2,101.5	4,2,Bc_5784/2,97.5	3,3,Bc_5784/2,102.5	4,4,Kamil,87.5
1,1,Otakar,95	4,2,Trebevišta,95	3,3,Otakar,108.5	4,4,Zvolen,85
1,1,Baranja,55	4,2,Otakar,85	3,3,Bc_Marta,97.5	4,4,Bc_5784/1,83.5
1,1,Kamil,101.5	4,2,Kamil,54	3,3,Kamil,112.5	4,4,Trebevišta,88.5
1,1,Bc_5784/1,103.5	4,2,Bc_Marta,85	3,3,Istra,102.5	4,4,Bc_5784/2,94
1,1,Bc_Marta,91	4,2,Zvolen,92.5	3,3,Bc_5784/1,102.5	4,4,Istra,81
4,1,Otakar,98.5	3,2,Kupa,67.5	1,3,Bc_5784/2,89	3,4,Baranja,92.5
4,1,Bc_5784/1,107.5	3,2,Bc_Marta,90	1,3,Istra,86.5	3,4,Otakar,90
4,1,Kamil,107.5	3,2,Bc_5784/1,87.5	1,3,Trebevišta,85	3,4,Zvolen,77.5
4,1,Bc_Marta,96.5	3,2,Bc_5784/2,90	1,3,Kupa,77.5	3,4,Bc_5784/1,95
4,1,Istra,93.5	3,2,Baranja,87.5	1,3,Bc_5784/1,100	3,4,Trebevišta,85
4,1,Trebevišta,96.5	3,2,Trebevišta,91	1,3,Bc_Marta,95	3,4,Kupa,72.5
4,1,Bc_5784/2,105	3,2,Zvolen,92.5	1,3,Kamil,102.5	3,4,Bc_Marta,82.5
4,1,Zvolen,99	3,2,Istra,87.5	1,3,Zvolen,99	3,4,Istra,82.5
4,1,Baranja,105	3,2,Otakar,87.5	1,3,Baranja,102.5	3,4,Kamil,80
4,1,Kupa,85	3,2,Kamil,90	1,3,Otakar,99	3,4,Bc_5784/2,90