

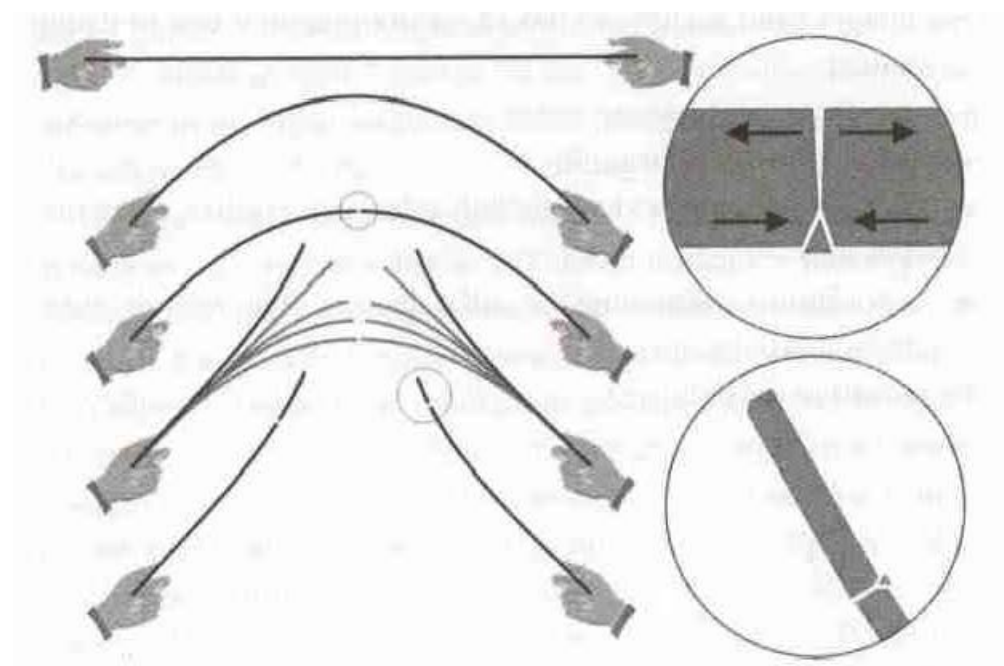
Kazalo

- 1 POSKUS V ŽIVO
- 2 KAJ JE POSKUS, KAJ JE OPAZOVANJE
- 3 NEKAJ ZGODOVINE
- 4 SIR RONALD AYLMER FISHER (1890-1962)
- 5 VLOGA STATISTIKE PRI ZNANSTVENO-RAZISKOVALNEM DELU
- 6 PRIMERI
- 7 NAMESTO ZAKLJUČKA

1 POSKUS V ŽIVO

Kvarkadabra v kuhinji, Znanstvene razlage kuhanja in prehrane, str.82

Naloga: primi špaget na konceh, razlomi ga na dva dela



http://www.lmm.jussieu.fr/spaghetti/audoly_neukirch_fragmentation.pdf

Fragmentation of rods by cascading cracks: why spaghetti do not break in half

Basile Audoly and Sébastien Neukirch

Lab. de Modélisation en Mécanique, CNRS/UPMC, 4 place Jussieu, Paris, France.

(Dated: December 22, 2004)

When thin brittle rods such as dry spaghetti pasta are bent beyond their limit curvature, they often break into more than two pieces, typically three or four. With the aim to understand these multiple breakings, we study the dynamics of a rod bent just below its limit curvature and suddenly released at one end. We find that the sudden relaxation of the curvature at the newly freed end leads to a burst of flexural waves, whose dynamics are described by a self-similar solution with no adjustable parameters. These flexural waves locally *increase* the curvature in the rod and we argue that this counter-intuitive mechanism is responsible for the fragmentation of brittle rods under bending. A simple experiment supporting the claim is presented.

PACS numbers: 62.20.Mk, 46.50.+a, 46.70.De

The physical process of fragmentation is relevant to several areas of science and technology. Because different physical phenomena are at work during the fragmentation of a solid body, it has mainly been studied from a statistical viewpoint [1–5]. Nevertheless a growing number of works have included physical considerations: surface energy contributions [6], nucleation and growth properties of the fracture process [7], elastic buckling [8, 9], and stress wave propagation [10]. Usually, in dynamic fragmentation, the abrupt application of fracturing forces

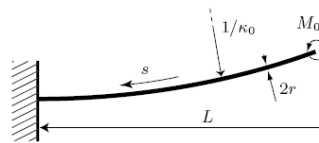


FIG. 1: The dynamics of a rod fragment following the initial breaking event in a brittle rod is modelled by releasing at time $t = 0$ a rod with fixed length L , initial curvature κ_0 and no initial velocity

.....

2006 sta dobila alternativno Nobelovo nagrado

- Uspešni poskusi imajo "zgodovino"
Naprej opazovanje, sledijo tipalni poskusi, nato načrtovani poskusi
- Bolj ko je domneva jasno opredeljena, boljši so statistični rezultati.
- Dober poskus sam po sebi ni dokaz, ampak vodi raziskovalce v ustrezno smer...
- Interdisciplinarnost je ključnega pomena

2 KAJ JE POSKUS, KAJ JE OPAZOVANJE

OPAZOVANJE, ANKETA, POPIS: raziskovalec popisuje, opisuje stanje.

NAČRTOVAN POSKUS: aktivno delovanje raziskovalca v smislu, da sam izbere vrednosti neodvisnih spremenljivk. Praviloma preverja specifično domnevo.

Dobro načrtovan poskus pogosto omogoča vrednotenje VZROČNO-POSLEDIČNIH povezav.

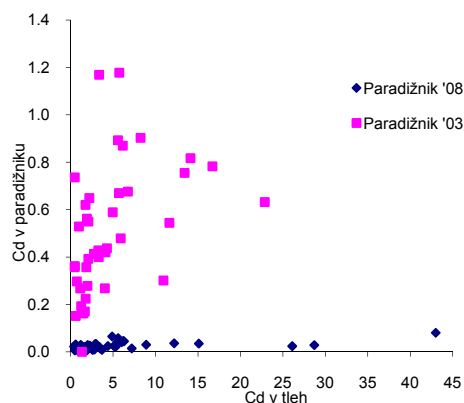
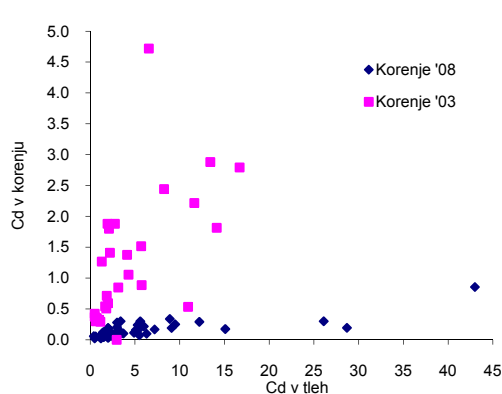
Opazovanje omogoča le vrednotenje POVEZAV.

2.1 Primer opazovanja: težke kovine v Celju

MO Celje financira raziskavo, vzorčijo tiste vrtničarje v Celju, ki se javijo na razpis (2003, 2008).

Namen: Kako količina Cd v zemlji vpliva na količino Cd v vrtninah? (M. Zupan)

Oznaka lokacije	Cd zlatotopka (mg/kg)	Analiza vzorcev rastlin (podatki veljajo za suho snov)						
		kumare (mg/kg)	paradižnik (mg/kg)	peteršil_listi (mg/kg)	petršil_koren (mg/kg)	korenje (mg/kg)	endivija (mg/kg)	radič (mg/kg)



- Pozor: različni skali na y-oseh!!!!
- Večinoma se vrtovi ne ponavljajo po letih.
- Različna laboratorija za določanje Cd v 2003 in v 2008.

Kaj vidimo?

- 2003 > 2008 za korenje in za paradižnik
- korenje > paradižnik v obeh letih

Kaj pa statistična analiza???

Za vsako leto posebej, primerjava po letih vprašljiva zaradi prepletanja vpliva leta, dodatnih dejavnikov (???) in laboratorija.

2.2 Načrtovan poskus: Žirovski vrh

Namen: poskusa je ugotoviti vpliv različnih koncentracij jalovine v treh tipih zemlje na privzem oziroma povečano vsebnost radioaktivnih elementov v treh vrstah rastlin. (M. Černe)

Dejavniki:

- vrsta rastline (A) → Korenje, Rukvica, Zelje
- tip zemlje (B) → Melj, Pesek, Glina
- % jalovine (C) → 0, 20, 40, 60

3 X 3 X 4 → 36 obravnavanj

M_0_K		M_0_R		M_0_Z	
M_20_K		M_20_R		M_20_Z	
M_40_K		M_40_R		M_40_Z	
M_60_K		M_60_R		M_60_Z	
P_0_K		P_0_R		P_0_Z	
P_20_K		P_20_R		P_20_Z	
P_40_K		P_40_R		P_40_Z	
P_60_K		P_60_R		P_60_Z	
G_0_K		G_0_R		G_0_Z	
G_20_K		G_20_R		G_20_Z	
G_40_K		G_40_R		G_40_Z	
G_60_K		G_60_R		G_60_Z	

Izid: vsebnost radionuklidov v rastlinah (U-238, Ra-226, Th-230 in Pb-210) v rastlini

Laboratorijski poskus, poskusna enota=lonec

Police: zgoraj levo, spodaj levo, zgoraj ob oknu, spodaj ob oknu (4 bloki)

Trifaktorski poskus v 4 blokih

3 NEKAJ ZGODOVINE

3.1 Dečki in deklice

1710 John Arbuthot pregleda zapise za City of London za 82 zaporednih let, in ugotovi, da se je vedno rodilo več dečkov kot deklic. (Vir: Daly, str. 369)

X je število let, ko se rodi več dečkov. $X \sim b(n=82, p)$

Arbuthot izračuna: če bi bilo razmerje p enako $\frac{1}{2}$, je verjetnost takega izida

$$P(X = 82) = \left(\frac{1}{2}\right)^{82} = 2 \cdot 10^{-25}$$

Sklep J. Arbuthot:

Observation constituted clear evidence for divine providence since, with wars and diseases resulting in higher death-rate for males, God had compensated by arranging more males to be produced, and hence arranged for every woman her proper husband.

Komentar:

- Statistična analiza danes enaka: test predznakov
- Vsebinska interpretacija rezultatov iz današnjega zornega kota

3.2 Kava ali čaj

Gustav III švedski kralj (1746-92)



Raziskovalna domneva: Pitje kave povzroča smrt.

Obnavljanji: pitje kave oz. čaja 3x na dan

Ponovitve: dva na smrt obsojena enojajčna dvojčka pomilosti, eden pije kavo, drugi čaj

Postopek: obsojenec pije kavo oz. čaj do smrti; le-to naj ugotovita dva neodvisna zdravnika.

....

Pivec čaja je umrl prvi, čeprav za tedanje čase pri častiljivi starosti 83 let.

Kraljevi zdravniki so sklepali, da je čaj bolj škodljiv kot kava.

This Article

Extract **FREE**

[Respond to this article](#)

[Alert me when this article is cited](#)

[Alert me when responses are posted](#)

[Alert me when a correction is posted](#)

Services

[Email this article to a friend](#)

[Find similar articles in BMJ](#)

[Add article to my folders](#)

[Download to citation manager](#)

[Request Permissions](#)

Citing Articles

[Citing Articles via Google Scholar](#)

BMJ 1996;312:1539 (15 June)

Letters

Coffee drinking was compared with tea drinking in monozygotic twins in 18th century

EDITOR,—One of the more peculiar attempts to throw light on the question of whether drinking coffee is bad for one's health¹ was carried out in the 18th century by King Gustaf III of Sweden. He is better known for instituting the Swedish Academy, the august body of 18 (18 because the king liked the sound of the Swedish word for that number, ardeton) whom Alfred Nobel later selected to award his prize in literature.

A pair of monozygotic twins had been sentenced to death for murder. Gustaf III commuted their death sentences to life imprisonment on the condition that one twin drank a large bowl of tea three times a day and that the other twin drank coffee. The twin who drank tea died first, aged 83—a remarkable age for the time. Thus the case was settled: coffee was the less dangerous of the two beverages. The king, on the other hand, was murdered at a masked ball in 1792 at the age of 45 and became the subject of an opera by Verdi.

Clinical lecturer Department of Chemical Pathology and Human Metabolism, Royal Free Hospital School of Medicine London NW3 2QG

Lars Breimer

Komentar:

- podobnost poskusnih enot maksimizirana
- tipalni poskus, $n=1$
- etični moment

3.3 Skorbut

Ladijski zdravnik James Lind: »Skorbut povzroči več smrti angleški kraljevi mornarici kot španska in francoska mornarica skupaj.« Izvede poskus leta **1747**, to je **prvi zabeleženi klinični poskus**

Raziskovalna domneva: ustrezen dodatek prehrani »zdravi« skorbut.

Obravnavanja (6): dodatki osnovni prehrani

- jabolčnik
- kapljice vitriol (na osnovi žveplove kisline H_2SO_4)
- pitje slane vode
- mešanica česen + gorčica + redkvice
- kis
- 2 pomaranči + 1 limona

Ponovitve: Izbere 12 podobnih mornarjev, ki so oboleli za skorbutom. Razdeli jih v 6 skupin, po 2 mornarja na obravnavanje.

Rezultati: po 1 tednu zmanjka citrusov, stanje teh dveh mornarjev se izjemno izboljša. Nekaj pozitivnega vpliva tudi pri jabolčniku.

Komentarji:

- Iz vidika današnje statistike manjka poskusu princip slučajnosti.
- Statistična analiza ni bila potrebna....
- Vsebinski del: koncept vitaminov ni bil znan.

4 SIR RONALD AYLMER FISHER (1890-1962)

Ronald Fisher je začetnik znanstvene discipline Načrtovanje in analiza poskusov .



1919: [Rothamsted Experimental Station](#)

Statistical Methods for Research Workers (1925)

The Design of Experiments (1935) podaja tako konceptualne kot tudi metodološke osnove za analizo poskusov.

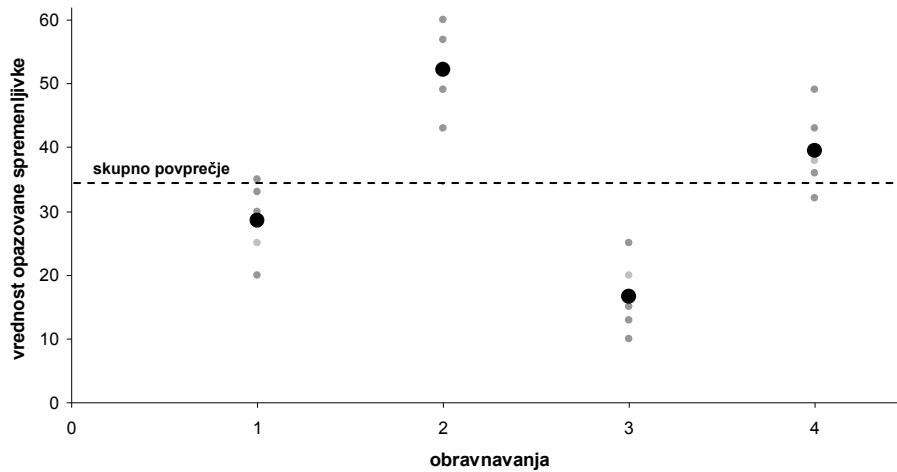
4.1 Analiza variance (ANOVA)

Y_i opisuje izid pri i -tem obravnavanju, število obravnavanj $K \geq 2$

Predpostavke: $Y_i \sim N(\mu_i, \sigma)$ $i = 1, 2, \dots, K$; podatki so neodvisni

Model: $Y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}$

Posledica: če je model OK, velja: $\varepsilon_{ij} \sim N(0, \sigma)$ iid



VSOTA KVADRATOV ODKLONOV preučevane spremenljivke se razbije po t. im. virih variabilnosti

$$\sum_{i=1}^K \sum_{j=1}^{n_i} (Y_{ij} - \mu)^2 = \sum_{i=1}^K n_i (\mu_i - \mu)^2 + \sum_{i=1}^K \sum_{j=1}^{n_i} (Y_{ij} - \mu_i)^2$$

skupna variabilnost = variabilnost obravnavj + nepojasnjena variabilnost

Isto velja za pripadajoče **STOPINJE PROSTOSTI**

$$N - 1 = K - 1 + N - K$$

...

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_K = \mu$$

ali

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_K = 0$$

Tabela ANOVA

Vir variabilnosti	VKO [SS]	SP[Df]	SKO [MS]	E(SKO)	F
Obravnavanja	VKO_{OBR}	$K-1$	$\frac{VKO_{OBR}}{SP_{OBR}}$	$\sigma^2 + \frac{1}{K-1} \sum_{i=1}^K n_i \alpha_i^2$	$\frac{SKO_{OBR}}{SKO_{OST}}$
Ostane	VKO_{OST}	$N-K$	$\frac{VKO_{OST}}{SP_{OST}}$	σ^2	
Skupaj	VKO_{SK}	$N-1$			

»experimental error« = napaka poskusa = SKO(ostane)

- Napaka poskusa je ocena za σ^2 .
- Ta ocena je nepristranska ocena variance, če je bil upoštevan **princip slučajnosti**. Sicer se ne ve....
- To oceno lahko izračunamo, če imamo **ponovitve obravnavanj**.
- **Boljša poskusna zasnova** zmanjša eksperimentalno napako in s tem poveča kakovost poskusa.

4.1.1 Ilustracija analize variance

Primer 1

skupina A: 1,2,3

povp A=2

skupina B: 5,6,7

povp B=6

skupno povprečje =4

skupna varianca=1

$$VKO_{skupaj} = (1-4)^2 + (2-4)^2 + (3-4)^2 + (5-4)^2 + (6-4)^2 + (7-4)^2 = 28$$

$$VKO_{obr} = 3(2-4)^2 + 3(6-4)^2 = 24$$

$$VKO_{ost} = (1-2)^2 + (2-2)^2 + (3-2)^2 + (5-6)^2 + (6-6)^2 + (7-6)^2 = 4$$

Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	24,0	1	24,0	24,00	0,0080
Within groups	4,0	4	1,0		
Total (Corr.)	28,0	5			

SKLEP: močno statistično značilni rezultati (p=0.008).

Primer 2

skupina A: 0,2,4

povp A=2

skupina B: 4,6,8

povp B=6

skupno povprečje =4
skupna varianca=4

Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	24,0	1	24,0	6,00	0,0705
Within groups	16,0	4	4,0		
Total (Corr.)	40,0	5			

SKLEP: mejno statistično značilni rezultati ($p=0.07$).

STATISTIČNI KOMENTAR: rezultati so mejni. Premajhno število enot.

Primer 3

skupina A: -2, 2, 6
skupina B: 2, 6, 10

povp A=2
povp B=6

skupno povprečje =4
skupna varianca=16

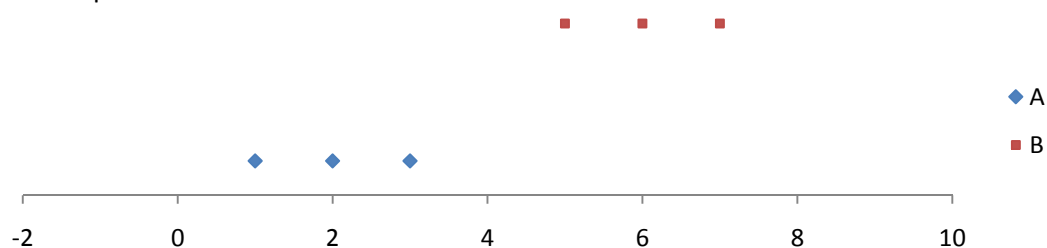
Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	24,0	1	24,0	1,50	0,2879
Within groups	64,0	4	16,0		
Total (Corr.)	88,0	5			

SKLEP: statistično neznačilni rezultati ($p=0.29$).

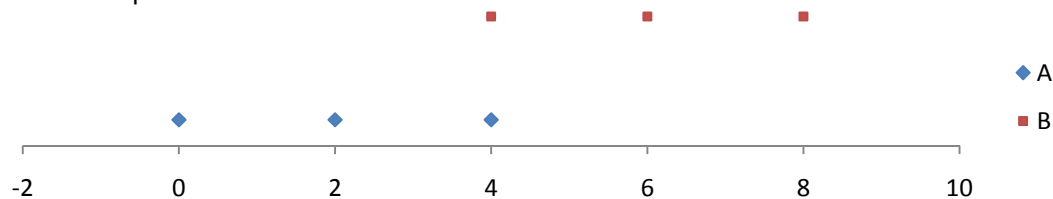
STATISTIČNI KOMENTAR: premajhna vzorca in velika variabilnost.

Primer 1

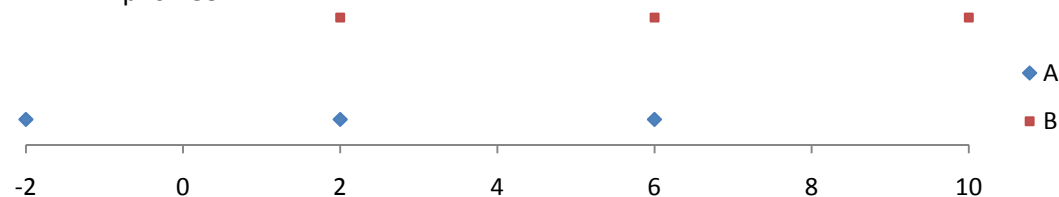
p=0.0000

**Primer 2**

p=0.07

**Primer 3**

p=0.288

**MIMOGREDE:****R. Mead posreduje pragmatičen nasvet za potrebno število enot v poskusu****Df za ostanek : $10 < Df < 20$; več kot 20 ni potrebno; manj kot 10 je premalo.****4.2 Fisherjevi koncepti za načrtovanje in analizo poskusov**

- **PRIMERJAVE MED OBRAVNAVANJI:** Kontrasti
- **PRINCIP SLUČAJNOSTI** je **varnostni ukrep**. S tem se izognemo potencialnim motnjam, ki jih povzročajo dejavniki, ki jih ne uspemo kontrolirati z ustrezno poskusno zasnovo.
 - izloči pristranost pri alokaciji enot (skušamo se izogniti možnosti, da ima določeno obravnavanja boljše/slabše pogoje od ostalih)
 - izloči vpliv raziskovalca, tehnikov, ocenjevalcev
 - omogoča uporabo matematične teorije.

- **PONOVITVE:** poskusne enote imajo svojo inherentno variabilnost. Če želimo dobiti oceno za varianco, je potrebno imeti ponovitve.
- **BLOKI:** poskusni material praviloma ni homogen. **Bloki so skupine enot, ki so glede na vrednosti motečih spremenljivk čim bolj homogene.** Bločna zasnova: v vsakem bloku vsako obravnavanje. S tem dosežemo boljšo oceno dejavnikov, ki jih preučujemo.
- **ORTOGONALNOST**
- **FAKTORSKA ZASNOVA**

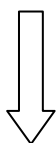
5 VLOGA STATISTIKE PRI ZNANSTVENO-RAZISKOVALNEM DELU

Interdisciplinarnost

- Namen in cilj raziskovalnega dela
- Jasno opredeljene raziskovalne domneve

Koraki ob sodelovanju statistika

- **Vrste spremenljivk**
 - **Odzivna spremenljivka (Y)** = Izid
 - **Preučevani dejavniki (X)**=pojasnjevalni dejavniki
 - Vrste: fiksni dejavniki, slučajni dejavniki
 - Relacije med njimi: križanje, hierarhija
 - **Moteči dejavniki**
 - znani (opisni, številski)
 - potencialni
 - neznani...
- **Ponovitve, psevdoponovitve**



- **Poskusna zasnova: KOMPROMIS**
 - upošteva namen in cilj poskusa.
 - Nevtralizira znane moteče dejavnike (npr. bločni poskus, latinski kvadrat), upoštevanje sospremenljivk v statističnem modelu.
 - Princip slučajnosti: varnostni ukrep proti potencialnim motečim dejavnikom
 - Upoštevati se mora tehnične, kadrovske, finančne omejitve.

Brez sodelovanja statistika

- Izvedba poskusa, pridobitev podatkov

Ob sodelovanju statistika

- Statistična analiza
- Interpretacija rezultatov
 - Statistična interpretacija
 - Vsebinska interpretacija

6 PRIMERI

6.1 Herbicidi

Namen poskusa je bilo primerjati učinkovitost štirih herbicidov A, A+, B, B+ in kontrole K, ki so jih uporabljali pri gojenju rdeče pese. Herbicidoma A in B je dodan isti dodatek +.

Za izvedbo poskusa so imeli 4 enake njive na različnih lokacijah.

Izid poskusa je bilo **število plevelov** na eksperimentalni parceli.

Raziskovalne domneve

- Domneva 1:** dodatek + poveča učinkovitost herbicida A.
- Domneva 2:** dodatek + poveča učinkovitost herbicida B.
- Domneva 3:** v povprečju imajo herbicidi večjo učinkovitost od kontrole.
- Domneva 4:** povprečje A-jev je enako povprečju B-jev

Obravnavanja: A=1, A+=2, B=3, B+=4, K=5

Moteči dejavnik = lokacija (njive). Vsako njivo razdelimo na 5 enakih parcel. Na vsaki njivi bo vsako obravnavanje po 1 x, določeno po principu slučajnosti

```
> for (i in 1:4) print(sample (1:5))
[1] 3 2 1 5 4
[1] 1 4 3 5 2
[1] 5 4 2 1 3
[1] 4 1 5 3 2
```

Shema poskusa:

Lokacija 1	B	A+	A	K	B+
Lokacija 2	A	B+	B	K	A+
Lokacija 3	K	B+	A+	A	B+
Lokacija 4	B+	A	K	B+	B

Poskusne enote: parcele znotraj bloka

Izid: število plevelov na parcelo

Poskusna zasnova: slučajni bloki

....

Tabela: *pleveli*= število plevelov na parceli

Herbicid	Njiva 1	Njiva 2	Njiva 3	Njiva 4
A+	2	7	12	18
A	4	12	23	22
B+	29	61	56	64
B	30	62	71	93
K	44	61	85	94

A. Preliminarna statistična analiza:

```
> tapply(herbicidi$pleveli, herbicidi$herbicid, mean)
  A  A+  B  B+  K
15.25 9.75 64.00 52.50 71.00
```

```
> tapply(herbicidi$pleveli, herbicidi$herbicid, var)
  A          A+          B          B+          K
 80.91667  46.91667 683.33333 256.33333 518.00000
```

Varianca so različne po obravnavanjih**Heteroskedastičnost!!!!**

Porazdelitev spremenljivke "število plevelov" je **diskretna**, verjetno bi jo lahko modelirali s Poissonova porazdelitvijo ; pri tej porazdelitvi je povprečje enako varianci = z večanjem povprečja se večja tudi varianca.

Transformacija: $pleveli1 = \sqrt{pleveli}$

```
> tapply(herbicidi$pleveli1, herbicidi$herbicid, mean)
  A  A+  B  B+  K
3.737587 2.941677 7.855258 7.169682 8.339601
```

```
> tapply(herbicidi$pleveli1, herbicidi$herbicid, var)
  A  A+  B  B+  K
1.462050 1.707256 3.059885 1.460874 1.934743
```

Test homogenosti varianc

```
> fligner.test(herbicidi$pleveli1, herbicidi$herbicid)
```

Fligner-Killeen test of homogeneity of variances

data: herbicidi\$plevelil and herbicidi\$herbicid
Fligner-Killeen:med chi-squared = 1.1131, df = 4, p-value = 0.8922

Uporaba ANOVA je legitimna.

B. ANOVA

Ničelna domneva:

$$H_0 : \mu_A = \mu_{A+} = \mu_B = \mu_{B+} = \mu_K = \mu$$

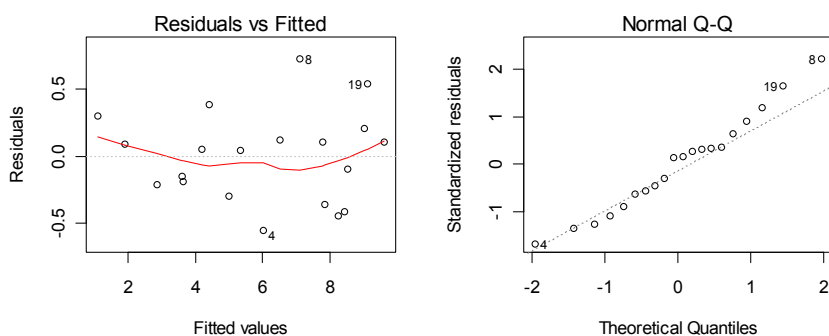
Povprečna zapleveljenost ni odvisna od obravnavanja.

```
> LinearModel.2 <- lm(plevelil ~ njiva +herbicid, data=herbicidi)
> summary.aov(LinearModel.2)
      Df Sum Sq Mean Sq F value    Pr(>F)
njiva   3  26.721   8.9069  49.626 4.886e-07 ***
herbicid 4  99.021  24.7553 137.929 6.395e-10 ***
Residuals 12  2.154   0.1795
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Ničelno domnevo zavrnamo. Obstajajo statično značilne razlike po herbicidih ($p=0.0000$).
- Ocena za varianco je 0.179.

C. Kontrola ustreznosti modela z grafičnimi prikazi

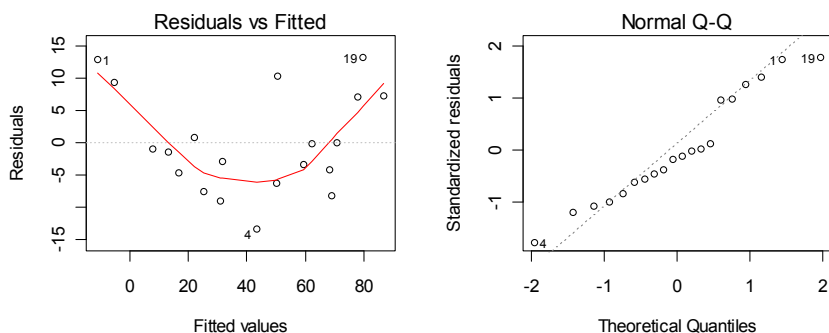
lm(pleveli1 ~ njiva + herbicid)



Zgleda OK.

**** Če korenske transformacije ne bi naredili, bi bili ostanki taki:

lm(pleveli ~ njiva + herbicid)



Ti ostanki niso ne ustrezajo zahtevam, model ni OK. Učenje na svojih napakah...

D. Analiza načrtovanih primerjav

Input: koeficienti za posamezni kontrast

Ime kontrasta:	.1	.2	.3	.4
A	1	0	-1	1
A+	-1	0	-1	1
B	0	1	-1	-1
B+	0	-1	-1	-1
K	0	0	4	0

To so 4 ortogonalni kontrasti.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	4.18197	0.18946	22.073	4.40e-11	***
njiva[T.2]	1.73890	0.26794	6.490	2.98e-05	***
njiva[T.3]	2.49582	0.26794	9.315	7.67e-07	***
njiva[T.4]	3.07244	0.26794	11.467	8.01e-08	***
herbicid.1	0.39796	0.14978	2.657	0.0209	*
herbicid.2	0.34279	0.14978	2.289	0.0410	*
herbicid.3	0.58271	0.04737	12.302	3.66e-08	***
herbicid.4	-2.08642	0.10591	-19.699	1.66e-10	***

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

Residual standard error: 0.4236 on 12 degrees of freedom
Multiple R-squared: 0.9832, Adjusted R-squared: 0.9733
F-statistic: 100.1 on 7 and 12 DF, p-value: 1.096e-09

Povprečja

A	A+	B	B+	K
3.737587	2.941677	7.855258	7.169682	8.339601

Odgovori na postavljena vprašanja

D1: obstoja statistično značilna razlika med A in A+ ($p=0.0209$); dodatek + povečuje učinkovitost herbicida A.

D2: obstoja statistično značilna razlika med B in B+ ($p=0.0410$); dodatek + povečuje učinkovitost herbicida B.

D3: povprečen efekt herbicidov se statistično značilno razlikuje od kontrole ($p=0.0000$); v povprečju herbicidi zmanjšujejo zapleveljenost v primerjavi s kontrolo.

D4: obstoja statistično značilna razlika med povp (A, A+) ter povp (B, B+) ($p=0.0000$); prvi je bolj učinkovit od drugega.

Nekaj alternativ za statistično analizo

E. Nimamo specificiranih vprašanj za primerjavo obravnavanj. Tedaj uporabimo enega od testov mnogoterih primerjav

```
> library(agricolae)
> comparison <- HSD.test(herbicidi$pleveli1,herbicidi$herbicid,12, 0.1795, group=TRUE,
main="Tukey")
```

Study: Tukey

HSD Test for herbicidi\$pleveli1

Mean Square Error: 0.1795

alpha: 0.05 ; Df Error: 12

Critical Value of Studentized Range: 4.50771

Honestly Significant Difference: **0.9549007**

Means with the same letter are not significantly different.

Groups, Treatments and means

a	K	8.339601
ab	B	7.855258
b	B+	7.169682
c	A	3.737587
c	A+	2.941677

Dobimo tri skupine: K in B, B in B+ ter A in A+, prvi dve skupini se prekrivata.

Opomba: Primerjava s kontrasti : A in A+ ($p=0.0209$), B in B+ ($p=0.0410$).

Nauk te zgodbe: Bolj ko so domneve jasno opredeljene, boljši so statistični rezultati.

F. Neparometrična alternativa (...ne najdemo ustrezne transformacije....)

```
> friedman.test(herbicidi$pleveli, herbicidi$herbicid, herbicidi$njiva)
```

```
Friedman rank sum test
```

```
data: herbicidi$pleveli, herbicidi$herbicid and herbicidi$njiva
Friedman chi-squared = 14.9367, df = 4, p-value = 0.004834
```

Test mnogoterih primerjav

Mejna razlika na povprečnih rangih:
$$D = z_{\alpha/K(K-1)} \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

```
> mean_rang
  A+      A      B+      B      K
3.625  5.375 12.625 15.250 15.625
```

```
> D=z_tab*d; D
[1] 11.74266
```

A+ (povprečni rang=3.625) se stat značilna razlikuje od K (15.625); razlika je 12, to je večje od mejne razlike 11.7.

Nauk te zgodbe: Moč neparometričnih testov je bistveno manjša v primerjavi s parametričnimi testi.

6.2 Jagode

Proučevali so vpliv sorte in časa prekrivanja na pridelek jagod. V poskus so vključili štiri sorte jagod (s1, s2, s3, s4) in tri termine prekrivanja (februar: koda "2", marec: koda "3", april: koda "4"). Poskus je bil zastavljen v poskusni zasnovi slučajni bloki, s štirimi bloki.

Zanimalo jih je tudi, ali obstaja medsebojni vpliv med dejavnikoma sorta in termin prekrivanja. Podatki so masa jagod (kg) na parcelo.

Zasnova: 4x3 faktorski poskus v 4 blokih

```
> jagode$obravnvanje=as.factor(paste(jagode$sorta, jagode$termin, sep=""))
> head(jagode)
  blok termin sorta pridelek obravnvanje
1     1      2   s1      10.2 s12
2     1      2   s2      11.1 s22
3     1      2   s3       6.8 s32
4     1      2   s4       5.3 s42
5     1      3   s1       8.0 s13
6     1      3   s2       9.7 s23
> levels(jagode$obravnvanje)
[1] "s12" "s13" "s14" "s22" "s23" "s24" "s32" "s33" "s34" "s42" "s43" "s44"
```

Analiza povprečij in varianc po obravnvanjih:

```
> numSummary(jagode[, "pridelek"], groups=jagode$obravnvanje,
+ statistics=c("mean", "sd"), quantiles=c(0,.25,.5,.75,1))
      mean      sd n
s12    11.175 1.1870833 4
s13     9.375 1.9465782 4
s14     4.900 2.0896571 4
s22     9.725 1.0436315 4
s23     9.775 1.3937360 4
s24     8.750 1.8266545 4
s32     9.025 1.5305228 4
s33     9.425 0.5909033 4
s34     3.775 1.1786291 4
s42     6.175 1.4453950 4
s43     5.625 2.1360009 4
s44     2.175 1.0242884 4
```

Ali so variance po obravnvanjih homogene?

```
> fligner.test(jagode$pridelek, jagode$obravnvanje)

      Fligner-Killeen test of homogeneity of variances

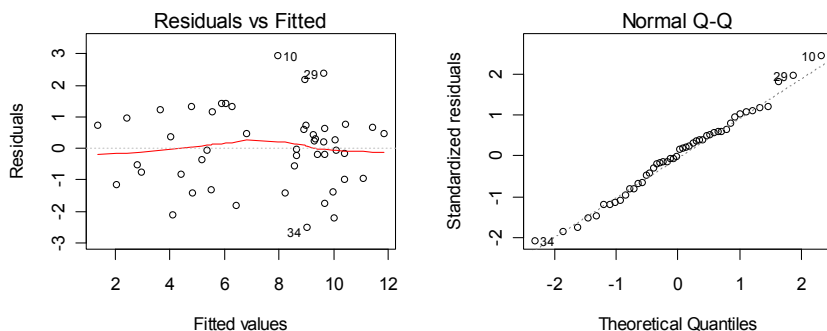
data:  jagode$pridelek and jagode$obravnvanje
Fligner-Killeen:med chi-squared = 11.0288, df = 11, p-value = 0.4409
```

ANOVA je legitimna

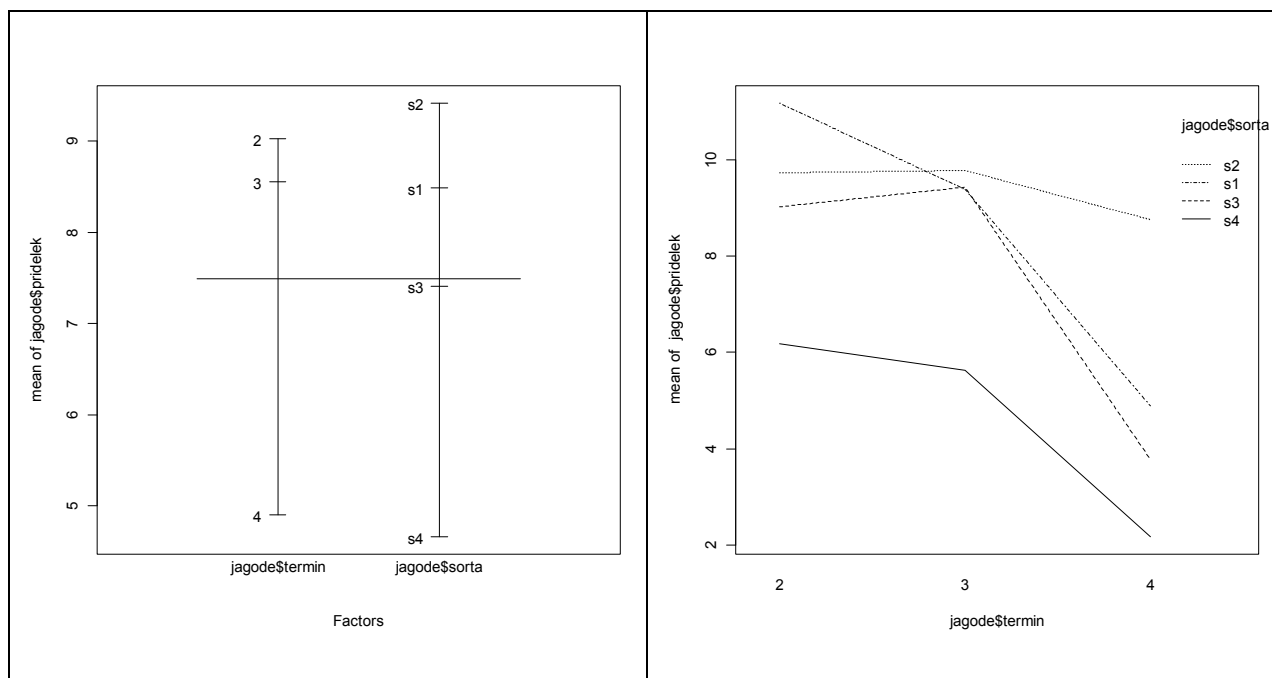
```
> LinearModel.1 <- lm(pridelek ~ blok+obravnvanje, data=jagode)
> summary.aov(LinearModel.1)
      Df Sum Sq Mean Sq F value    Pr(>F)
blok      3  13.69   4.564   2.1685   0.1104
obravnvanje 11 356.01  32.365  15.3777 5.524e-10 ***
Residuals  33   69.45   2.105
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ali je model primeren za podatke?

lm(pridelek ~ blok + obravnavanje)



Delni rezultat: med obravnavanji so statistično značilne razlike ($p=0.0000$). Nadaljevanje analize



```
> LinearModel.2 <- lm(pridelek ~ blok+sorta*termin, data=jagode)
```

```
> summary.aov(LinearModel.2)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
blok	3	13.692	4.564	2.1685	0.11039	
sorta	3	152.685	50.895	24.1822	1.835e-08	***
termin	2	163.007	81.503	38.7254	2.204e-09	***
sorta:termin	6	40.320	6.720	3.1929	0.01392	*
Residuals	33	69.453	2.105			

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

Interakcija je statistično značilna, ni pa močna v primerjavi z značilnostjo obeh glavnih vplivov.

Interpretacija interakcije: Kakšen je trend pri s1, s2, s3, s4? Linearni, kvadratni

Določite kontraste

Vnesite koeficiente kontrastov

Ime kontrasta:	.s1_L	.s1_Q	.s2_L	.s2_Q	.s3_L	.s3_Q	.s4_L	.s4_Q	.9	.10	.11
s12	-1	1	0	0	0	0	0	0	0	0	0
s13	0	-2	0	0	0	0	0	0	0	0	0
s14	1	1	0	0	0	0	0	0	0	0	0
s22	0	0	-1	1	0	0	0	0	0	0	0
s23	0	0	0	-2	0	0	0	0	0	0	0
s24	0	0	1	1	0	0	0	0	0	0	0
s32	0	0	0	0	-1	1	0	0	0	0	0
s33	0	0	0	0	0	-2	0	0	0	0	0
s34	0	0	0	0	1	1	0	0	0	0	0
s42	0	0	0	0	0	0	-1	1	0	0	0
s43	0	0	0	0	0	0	0	-2	0	0	0
s44	0	0	0	0	0	0	1	1	0	0	0

V redu Prekliči Pomoč

Opomba: tole je samo za ilustracijo koeficientov za kontraste in njihovih imen, sicer ne dela, ker RCmdr zahteva vse kontraste.

```

> contrasts(jagode$obravnavanje,8)

> contrasts(jagode$obravnavanje)=cbind(
+ c(-1,0,1,0,0,0,0,0,0,0,0,0),c(1,-2,1,0,0,0,0,0,0,0,0,0),
+ c(0,0,0,-1,0,1,0,0,0,0,0,0),c(0,0,0,1,-2,1,0,0,0,0,0,0),
+ c(0,0,0,0,0,-1,0,1,0,0,0),c(0,0,0,0,0,0,1,-2,1,0,0,0),
+ c(0,0,0,0,0,0,0,0,-1,0,1),c(0,0,0,0,0,0,0,0,0,1,-2,1))

> Kontrasti=contrasts(jagode$obravnavanje)

> ## poimenujem 8 kontrastov smiselno, ostala imena so brez veze

> colnames(Kontrasti)=c('.s1_L','.s1_Q','.s2_L','.s2_Q'
, '.s3_L','.s3_Q','.s4_L','.s4_Q','.bla','.bla','.bla')

> Kontrasti
      .s1_L .s1_Q .s2_L .s2_Q .s3_L .s3_Q .s4_L .s4_Q      .bla
s12      -1      1      0      0      0      0      0      0      -0.4957089
s13       0     -2      0      0      0      0      0      0      -0.4957089
s14       1      1      0      0      0      0      0      0      -0.4957089
s22       0      0     -1      1      0      0      0      0      0.2171194
s23       0      0      0     -2      0      0      0      0      0.2171194
s24       0      0      1      1      0      0      0      0      0.2171194
s32       0      0      0      0     -1      1      0      0      0.1104925
s33       0      0      0      0      0     -2      0      0      0.1104925
s34       0      0      0      0      1      1      0      0      0.1104925
s42       0      0      0      0      0      0     -1      1      0.1680970
s43       0      0      0      0      0      0      0     -2      0.1680970
s44       0      0      0      0      0      0      1      1      0.1680970

      .bla      .bla
s12     -0.04831632  0.04402522
s13     -0.04831632  0.04402522
s14     -0.04831632  0.04402522
s22     -0.44607142  0.06228541
s23     -0.44607142  0.06228541
s24     -0.44607142  0.06228541
s32      0.17715986 -0.45431904
s33      0.17715986 -0.45431904
s34      0.17715986 -0.45431904
s42      0.31722789  0.34800841
s43      0.31722789  0.34800841
s44      0.31722789  0.34800841

> contrasts(jagode$obravnavanje)=Kontrasti

> LinearModel.3 <- lm(pridelek ~ blok + obravnavanje, data=jagode)

> summary(LinearModel.3)

Call:
lm(formula = pridelek ~ blok + obravnavanje, data = jagode)

Residuals:
      Min       1Q   Median       3Q      Max
-2.50833 -0.86042  0.07917  0.72500  2.95000

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.6917    0.4188  15.978 < 2e-16 ***
blok2           0.6917    0.5923   1.168  0.251240
blok3           1.0583    0.5923   1.787  0.083138 .
blok4           1.4500    0.5923   2.448  0.019840 *
obravnavanje.s1_L -3.1375    0.5129  -6.117  6.84e-07 ***
obravnavanje.s1_Q -0.4458    0.2961  -1.506  0.141701
obravnavanje.s2_L -0.4875    0.5129  -0.950  0.348792
obravnavanje.s2_Q -0.1792    0.2961  -0.605  0.549301

```

```

obravnvanje.s3_L -2.6250    0.5129   -5.118  1.30e-05 ***
obravnvanje.s3_Q -1.0083    0.2961   -3.405  0.001754 **
obravnvanje.s4_L -2.0000    0.5129   -3.899  0.000447 ***
obravnvanje.s4_Q -0.4833    0.2961   -1.632  0.112153
obravnvanje.bla  -1.6773    0.7254   -2.312  0.027138 *
obravnvanje.bla  -5.4605    0.7254   -7.528  1.17e-08 ***
obravnvanje.bla  -2.3538    0.7254   -3.245  0.002693 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 1.451 on 33 degrees of freedom
Multiple R-squared:  0.8418,    Adjusted R-squared:  0.7748
F-statistic: 12.55 on 14 and 33 DF,  p-value: 2.013e-09

```

Interpretacija rezultatov:

- Sorta 1 ima linearni trend ($p=0.0000$), isto velja za sorto 4 ($p=0.0004$), največji pridelek je februarja.
- Sorta 2 je neobčutljiva na termin prekrivanja.
- Sorta 3 ima kvadratni trend ($p=0.0018$), najslabši pridelek je aprila.

Mučimo podatke dalje: Ali interakcijo povzroča sorta 2?

Ne smem izločiti iz podatkov sorto 2, ker potem testiranje ne bi bilo s pravo napako in Df.

```

> LinearModel.5 <- lm(pridelek ~ blok+sorta*termin, data=brez_s2)
> summary.aov(LinearModel.5)
      Df Sum Sq Mean Sq F value    Pr(>F)
blok    3  27.277   9.092   5.9356 0.003536 **
sorta    2  93.395  46.697  30.4853 2.578e-07 ***
termin   2 190.715  95.358  62.2517 3.175e-10 ***
sorta:termin 4   9.940   2.485   1.6223 0.201150
Residuals 24  36.763   1.532
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Naredim kontraste, ki ne vsebujejo sorte 2 in ugotavljam, kakšen je trend pri teh treh sortah. Spodaj navedeni kontrasti so ortogonalni.

$s_{H1} = c(1,1,1,0,0,0,-1,-1,-1,0,0,0)$: s1 napram s3 (Helmert)

$s_{H2} = c(1,1,1,0,0,0,1,1,1,-2,-2,-2)$: povprečje s1 in s3 napram s4 (Helmert)

$t_L = c(-1,0,1,0,0,0,-1,0,1,-1,0,1)$: linearni trend za sorte 1,3 in 4

$t_Q = c(1,-2,1,0,0,0,-1,2,1,1,-2,1)$: kvadratni trend za sorte 1,3 in 4

Interakcije= to je produkt

$s_{H1_L} = K1 * K3 = c(-1,0,1,0,0,0,1,0,-1,0,0,0)$: razlika v linearnem trendu pri sortah 1 in 3

$s_{H1_Q} = c(1,-2,1,0,0,0,-1,2,1,0,0,0)$: razlika v kvadratnem trendu pri sortah 1 in 3

$s_{H2_L} = c(-1,0,1,0,0,0,-1,0,1,2,0,-2)$: razlika v linearnem trendu med sorto 4 in povprečjem sorte 1 in 3

$s_{H2_Q} = c(1,-2,1,0,0,0,1,-2,1,-2,4,-2)$: razlika v kvadratnem trendu med sorto 4 in povprečjem sorte 1 in 3

```
> contrasts(jagode$obravnavanje)=cbind(
+ c(1,1,1,0,0,0,-1,-1,-1,0,0,0), c(1,1,1,0,0,0,1,1,1,-2,-2,-2),
+ c(-1,0,1,0,0,0,-1,0,1,-1,0,1), c(1,-2,1,0,0,0,1,-2,1,1,-2,1),
+ c(-1,0,1,0,0,0,1,0,-1,0,0,0), c(1,-2,1,0,0,0,-1,2,1,0,0,0),
+ c(-1,0,1,0,0,0,-1,0,1,2,0,-2), c(1,-2,1,0,0,0,1,-2,1,-2,4,-2))

> Kontrasti=contrasts(jagode$obravnavanje)

> ## poimenujem 8 kontrastov smiselno, ostala imena so brez veze

> colnames(Kontrasti)=c('.s_H1', '.s_H2', '.t_L', '.t_Q',
, '.s_H1_L', '.s_H1_Q', '.s_H2_L', '.s_H2_Q', '.bla', '.bla', '.bla')

> Kontrasti
      .s_H1 .s_H2 .t_L .t_Q .s_H1_L .s_H1_Q .s_H2_L .s_H2_Q
s12      1      1     -1      1       -1        1       -1        1
s13      1      1      0     -2        0       -2        0       -2
s14      1      1      1      1        1        1        1        1
s22      0      0      0      0        0        0        0        0
s23      0      0      0      0        0        0        0        0
s24      0      0      0      0        0        0        0        0
s32     -1      1     -1      1        1       -1       -1        1
s33     -1      1      0     -2        0        2        0       -2
s34     -1      1      1      1       -1        1        1        1
s42      0     -2     -1      1        0        0        2       -2
s43      0     -2      0     -2        0        0        0        4
s44      0     -2      1      1        0        0       -2       -2
      .bla      .bla      .bla
s12  0.04743437 0.11773849 0.03783499
s13  0.08301015 0.20604237 0.06621124
s14  0.04743437 0.11773849 0.03783499
s22  0.43689856 -0.51830518 -0.67368879
s23 -0.09486919 -0.72014520 0.62110219
s24 -0.87566607 -0.08610769 -0.37305709
s32  0.07115156 0.17660774 0.05675249
s33  0.03557578 0.08830387 0.02837625
s34  0.07115156 0.17660774 0.05675249
s42  0.05929297 0.14717312 0.04729374
s43  0.05929297 0.14717312 0.04729374
s44  0.05929297 0.14717312 0.04729374
```



```

> contrasts(jagode$obravnavanje)=Kontrasti

> LinearModel.4 <- lm(pridelek ~ blok + obravnavanje, data=jagode)

> summary(LinearModel.4)

Call:
lm(formula = pridelek ~ blok + obravnavanje, data = jagode)

Residuals:
    Min       1Q   Median       3Q      Max
-2.50833 -0.86042  0.07917  0.72500  2.95000

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.66149    0.42085   15.828 < 2e-16 ***
blok2           0.69167    0.59226    1.168 0.251240
blok3           1.05833    0.59226    1.787 0.083138 .
blok4           1.45000    0.59226    2.448 0.019840 *
obravnavanje.s_H1  0.59786    0.30760    1.944 0.060507 .
obravnavanje.s_H2  1.07571    0.17321    6.211 5.20e-07 ***
obravnavanje.t_L  -2.64786    0.30760   -8.608 6.00e-10 ***
obravnavanje.t_Q  -0.66595    0.17321   -3.845 0.000521 ***
obravnavanje.s_H1_L -0.16571    0.38356   -0.432 0.668533
obravnavanje.s_H1_Q  0.18109    0.24962    0.725 0.473287
obravnavanje.s_H2_L -0.32393    0.21349   -1.517 0.138709
obravnavanje.s_H2_Q -0.09131    0.12169   -0.750 0.458352
obravnavanje.bla  -0.72520    0.72537   -1.000 0.324701
obravnavanje.bla  -3.85950    0.72537   -5.321 7.16e-06 ***
obravnavanje.bla  -0.86086    0.72537   -1.187 0.243783
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.451 on 33 degrees of freedom
Multiple R-squared:  0.8418, Adjusted R-squared:  0.7748
F-statistic: 12.55 on 14 and 33 DF, p-value: 2.013e-09

```

- Interakcije niso značilne.
- Za te tri sorte skupaj zaznamo kvadratni trend.
- s1 in s3 mejno statistično različni (p=0.06).
- s4 je statistično značilno slabša od povprečja s1 in s3 (p=0.0000).

Kmetijski nasveti:

- Priporočamo sorto 2, ker ima visok pridelek in je neodvisna od časa prekrivanja.
- Odsvetujemo sorto 4, ker ima nizek pridelek.

7 NAMESTO ZAKLJUČKA

[MALI PRINC SE VRNE, 1997](#)

Str. 38, 39

...

MP: Kaj pa računaš?

S: Cel kup stvari. Na primer, ali **Eskimi pojedó več sladoleda poleti ali pozimi**, ali **če so žabe preteklo pomlad kvakale glasneje kot pred desetimi leti**.

MP: Kako smešno. Zakaj pa?

S: Ker sem statistik, tak je moj poklic.... Ljudje strašno radi štejejo, merijo in primerjajo. Brez konca se sprašujejo, **ali redkvice rastejo pri njih bolje kot v sosednji vasi** in **ali ima prebivalstvo tam daljše ali krajše lase**. Na ta vprašanja lahko odgovori statistika...

