CHAPTER 6. ANALYSIS OF VARIANCE (ONE-WAY)

6.1. What is “Analysis of variance” (ANOVA)? 6.2. One-Way Analysis of Variance

6.3. The Least significant difference intervals

6.4. The p-value

Chapter 6: Assignments

We want to investigate if there is an effect of the type of fertilizer applied to apple trees and the production of apples. We randomly select 15 trees and randomly assign them to one of three groups (5 trees per group). We perform a test in which we apply one type of fertilizer (fertilizer 1, 2 or 3) to each group. The data are shown. At a α=0.05, can it be concluded that there is a significant difference in the production of apples depending on which fertilizer is used? Which fertilizer/fertilizers causes a higher/lower production than the other/others?

Fertilizer 1	Fertilizer 2	Fertilizer 3
10	6	5
12	8	9
9	3	12
15	0	8
13	2	4

- I tilfældet, kan det ses at vi har startet med at opskrive dataene i en variabel.
- Derefter har vi inddelt værdierne inde i 1,2 og 3 sektionskategorier.
- Derefter har vi gjort selve sektionsvariablerne til faktorer, fordi de er ikke bogstaver som såden!
- Til sidst har vi brugt one-way anova test og derved fundet forskellene gennem LSD-testen.

```
"Opgave 1"
maaling <- c(10,12,9,15,13,6,8,3,0,2,5,9,12,8,4)
behandling <- c(rep("1",5),rep("2",5),rep("3",5))
kunst <- data.frame(maaling,behandling)
kunst
str(kunst)
kunst$behandling <- as.factor(kunst$behandling)
res.aov <- aov(maaling~behandling,data=kunst)
summary(res.aov)
print(LSD.test(res.aov,"behandling"))
```

- Nu kommer følgende resultater fra kommandolinjen.

```
> str(kunst)
'data.frame':   15 obs. of  2 variables:
 $ maaling   : num  10 12 9 15 13 6 8 3 0 2 ...
 $ behandling: chr  "1" "1" "1" "1" ...
> kunst$behandling <- as.factor(kunst$behandling)
> res.aov <- aov(maaling~behandling,data=kunst)
> summary(res.aov)
            Df Sum Sq Mean Sq F value  Pr(>F)
behandling   2  160.1   80.07   9.168 0.00383 **
Residuals   12  104.8    8.73
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> print(LSD.test(res.aov,"behandling"))
Error in LSD.test(res.aov, "behandling") :
  could not find function "LSD.test"
```

We want to evaluate three different methods to lower the blood pressure of individuals that have been diagnosed with high blood pressure. Eighteen subjects are randomly assigned to three groups (6 per group): the first group takes medication, the second group exercises, and the third one follows a specific diet. After four weeks, the reduction in each person's blood pressure is recorded. Is there a significant difference among the reduction obtained from each of the three methods? If yes, which method was more effective?

Medication	Exercise	Diet
12	5	6
8	9	10
11	2	5
17	0	9
16	1	8
15	3	6

- I tilfældet, kan det ses at vi har startet med at opskrive dataene i en variabel.
- Derefter har vi inddelt værdierne inde i 1,2 og 3 sektionskategorier.
- Derefter har vi gjort selve sektionsvariablerne til faktorer, fordi de er ikke bogstaver som såden!
- Til sidst har vi brugt one-way anova test og derved fundet forskellene gennem LSD-testen.

```
"Opgave 2"
measure <- c(12,8,11,17,16,15,5,9,2,0,1,3,6,10,5,9,8,6)
treatment <- c(rep("3",6),rep("4",6),rep("5",6))
blodtryk <- data.frame(measure,treatment)
blodtryk
str(blodtryk)
blodtryk$treatment <- as.factor(blodtryk$treatment)
res.aov <- aov(measure~treatment,data=blodtryk)
summary(res.aov)
print(LSD.test(res.aov,"treatment"))
```

- Nu kommer følgende resultater fra kommandolinjen.

In the table below, there are randomly selected scores for eight amateur basketball teams in each of five Danish regions, for a particular weekend. Is there sufficient evidence to support that there is a difference in mean scores by region? If yes, which region/s got the highest scores and which one the lowest?

Region Hovedstaden	Region Sjælland	Region Syddanmark	Region Midtjylland	Region Nordjylland
68	78	89	62	57
75	79	87	74	65
95	65	75	71	78
85	67	65	70	88
84	60	84	72	67
88	79	92	72	77
85	57	84	64	72
75	74	72	75	69

- Vi bruger den samme step som de sidste opgaver. Forskellen bliver bare, at vi laver faktorisering på 5 sektioner.

```
"Opgave 3"
score <- c(68,75,95,85,84,88,85,75,78,79,65,67,60,79,57,74,89,87,75,65,84,92,84,72,62,74,71,70,72,72,64,75,57,65,78,88,67,77,72,69
section <- c(rep("HO",8),rep("SJ",8),rep("SY",8),rep("MI",8),rep("NO",8))
regioner <- data.frame(score,section)
regioner
str(regioner)
regioner$section <- as.factor(regioner$section)
res.aov <- aov(score~section,data=regioner)
print(LSD.test(res.aov,"section"))
```

```
> str(regioner)
'data.frame':   40 obs. of  2 variables:
 $ score  : num  68 75 95 85 84 88 85 75 78 79 ...
 $ section: chr  "HO" "HO" "HO" "HO" ...
> regioner$section <- as.factor(regioner$section)
> res.aov <- aov(score~section,data=regioner)
> print(LSD.test(res.aov,"section"))
Error in LSD.test(res.aov, "section") :
  could not find function "LSD.test"
>
```