

Exercise 2

1. A plant scientist wants to test the effect of a new herbicide on lentils. He considers the following treatments:

1. control (without weeding, no fertilizer)	TR = 1
2. weeding by hand	TR = 2
3. spraying with herbicide before	TR = 3
4. spraying with herbicide after	TR = 4
5. weeding by hand + fertilizer	TR = 5
6. spraying with herbicide before + fertilizer	TR = 6
7. spraying with herbicide after + fertilizer	TR = 7

The experiment was arranged as a block design experiment with 3 blocks (variable BLOCK) and one treatment factor (variable TR). The harvesting weight (variable Y) was used as response. The data is given in the file `lentil.dat`. (`d.len <- read.table('http://stat.ethz.ch/Teaching/Datasets/WBL/lentil.dat', header=TRUE)`)

- a) Check with an analysis of variance if there are differences between treatments.

R hints:

- Reading data in:

```
t.url <- "http://stat.ethz.ch/Teaching/Datasets/WBL/lentil.dat"
d.len <- read.table(t.url, header=T)
d.len$BLOCK <- factor(d.len$BLOCK)
d.len$TR <- factor(d.len$TR)
```

- Plotting the data:

Use the function

```
> plot(as.numeric(d.len$TR), ...)
```

to plot the data (TR on the x-axis and Y on the y-axis).

For a more informative plot use the following combination of R functions to create an empty plot first and then fill it with text:

```
> plot(..., ..., type='n', xlab='TR', ylab='Y')
> text(..., ..., labels=d.len$BLOCK)
```

- b) In order to detect existing differences between treatments, we consider the following contrasts:

Contrast	Treatment						
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
L1	-6	+1	+1	+1	+1	+1	+1
L2	0	-1	-1	-1	+1	+1	+1
L3	0	+2	-1	-1	+2	-1	-1
L4	0	0	-1	+1	0	-1	+1
L5	0	-2	+1	+1	+2	-1	-1
L6	0	0	+1	-1	0	-1	+1

Are these contrasts orthogonal? What questions can be answered by testing them?

- c) Test the contrasts.

R hints:

- Contrasts: Generate a matrix `len.contr` with the contrasts L1 to L6 in the **columns** (either use `cbind` or `matrix`). Then use the following R code:


```
> contrasts(d.lentil$TR) <- len.contr # assign contrasts
> r.len <- aov(...)
> summary.lm(r.len)
```
- ANOVA tables for contrasts: R code:


```
> summary(r.len,split=list(Factorname=list(L1=1,...)))
```

d) Write down the design matrix.

(Source: R.G. Petersen, *Agricultural Field experiments - Design and Analysis*, 1994, p. 113)

2. The response time (in milleseconds) was determined for three types of electrical circuits. The results were:

Type 1	9	12	10	8	15
Type 2	20	21	23	17	30
Type 3	6	5	8	16	7

a) Test the hypothesis, that all types have the same response time.

R hints:

- Creating the data set:


```
> y <- c(9,12,10,8,15,20,21,23,17,30,6,5,8,16,7)
> Ty <- c(...)
> circ <- data.frame(Type=Ty,Y=y)
> circ$Type <- as.factor(circ$Type)
```
- ANOVA table:


```
use the functions
> circ.fit <- aov(...)
> summary(circ.fit)
```

b) Use Tukey's method to compare pairs of treatment means.

R hints:

- Tukey's method:


```
> TukeyHSD(circ.fit,"Type", conf.level=...)
```

c) Construct a set of orthogonal contrasts, assuming that circuit type 2 was different from the other two.

R hints:

- Contrasts:


```
see exercise 1.
```

d) Test the contrasts.

R hints:

- Analysis of contrasts:


```
use the argument "split" to show the contrasts of the fitted data.
> summary(...,split=list(Type=list(L1=1,L2=2)))
```

3. An experiment was run to compare three different methods to measure air pollution. Two different technicians did the measurements, which are given in ppm (parts per million). To control the influence of the technicians a block design was used.

Technician	Treatment		
	A	B	C
1	7	36	2
2	13	44	18

- a) Calculate the overall mean, treatment and block means. Make a 2×3 contingency table with the residuals.
- b) Write down the complete anova table. How large is $\hat{\sigma}$?

R hints:

- First create a vector for each factor and a vector for the measurements (i.e. 3 vectors of length 6)
- then create a data frame (do not forget `as.factor(...)` for the factors).
- You can now use the function
`> aov(Y~techn+treatment,...)`

4. A scientist is interested in how genotype of a strawberry plant affects fruit yield. There are three levels of genotype (AA, AB, BB) and ten plots of land, three plants per plot. Each of the three genotypes is present in each plot. The data is in file `strawb.dat`.

```
st<-read.table("http://stat.ethz.ch/Teaching/Datasets/strawb.dat",header=TRUE)
```

- a) Plot the data.
- b) Do an analysis of variance on the data.
- c) Do an analysis of variance without taking into account land effects.
- d) Compare the results in b) and c). Why are the degrees of freedom different? Which result would you use?