Lunds University
Mathematical Sciences/Mathematical Statistics 2016

Computer lab 3

1 Introduction to the lab

The computer lab consists of the following parts

- One-way ANOVA.
- Randomized block design.
- Latin Squares.

Try to answer all questions. Do not pass a part until you understand it.

2 One-way ANOVA

In the lectures you have gone through the concept of one-way ANOVA. Randomization is used in experimental design in order to create homogeneous treatment groups and reduce bias. This part of the computer lab will illustrate the completely randomized design approach. We start out by studying the rate of nervous impulse measurements among specimens from 5 species of molluscs. It has been hypothesised that the rate of impulse is independent of the type of mollusc. We want to investigate this by running one-way ANOVA on the samples.

Load the data using the command,

```
Data <- read.table("mollusc.dat")
```

	A	B	C	D
1	48.14	54.75	69.48	71.53
2	35.51	92.24	84.61	64.37
3	46.43	65.5	65.22	81.72
4	56	94.29	85.6	77.6
5	39.68	90.53	80.99	72.38
6	51.46	83.41	61.6	67.43
7	54.47	103.27	77.18	51.34
8	30.15	78.24	93.75	33.71

To keep track of the meaning of the different columns in the data file we name them

```
names(Data)<-c("response","type")
attach(Data)
```

Next we move on to performing the ANOVA, with the response as the dependent variable and the type as the independent variable.

```
result = aov(response~type, Data)
summary(results)
detach(Data)
```

Can we reject the null-hypothesis, and if so, on what significance level? Check the model adequacy by checking the residuals.

3 Randomized block design

Above we studied the one-way fully randomized approach. We now proceed to study the concept of randomized blocks. Randomized blocks are typically introduced when there is some known variability among different test groups that we know will effect the result. By introducing blocking we are able to better control sources of variability that we are not interested in (that have been blocked).

In this section we want to study how the yield of penicillin is affected by the manufacturing process. We are studying 4 different manufacturing ways, hence called treatments, A, B, C and D. Each treatment was used once per blend and a total of five blends were used (G. Box, J, Hunter, W. Hunter p. 146).

To control for the variability between the blends we use a blocked design. Once again we will use ANOVA as analysis tool.

Start with loading and naming the data

```
Data1 <- read.table("penicilin.dat")
names(Data1) <- list("Treatment","Blend","Yield")
attach(Data1)
```

Similarly, create a vector of blocking factors for each element in the response data

```
Blend <- factor(Blend)           # blocking factor
```

We are now ready to run ANOVA on the randomized blocks

```
results <- aov(Yield ~ Treatment + Blend)
summary(results)
detach(Data1)
```

Is there evidence for differences between treatments? For blocks? Check the model adequacy by checking the residuals.

4 Latin Squares

In this section we set out to study the effect of two variables on the response. This method is called Latin Squares and is useful when the experimenter wishes to control the variance in two different directions. Essentially the Latin Square design is an example of an incomplete block design where there is a single treatment and two blocking variables. To illustrate the concept we will look at data air pollution from different gas blends. The data was collected using four cars, four drivers and four types of gas (G. Box, J, Hunter, W. Hunter p. 157).

Start by creating proper datarrame:

```
drivers <- c(rep("Driver1",1), rep("Driver2",1), rep("Driver3",1),
rep("Driver4",1))
cars <- c(rep("Car1",4), rep("Car2",4), rep("Car3",4), rep("Car4",4))
chemicals <- c("A","D","B","C", "B","C","D","A", "D","A","C",
"B", "C","B","A","D")
pollution <- c(19,23,15,19, 24,24,14,18, 23,19,15,19, 26,30,16,16)
pollutdata <- data.frame(cars, drivers, chemicals, pollution)
```

Next, we move on to perform the ANOVA analysis.

```
modelfit <- aov(pollution ~ drivers+cars+chemicals, pollutdata)
summary(modelfit)
```

What results did you get? Can we reject the null-hypothesis? Check the model adequacy by checking the residuals.

End of Lab