# TMA 4255 Applied Statistics
# Exercise 10

## Problem 1

One wants to examine the machines' capacities in a factory by recording how many units is produced by each machine in a specified time. The 4 workers $A_1, ..., A_4$ are randomly assigned to operate the 4 machines $M_1, ..., M_4$ and then the following observations are recorded:

| Worker/Machine | $M1$ | $M2$ | $M3$ | $M4$ |
|---|---|---|---|---|
| $A_1$ | 76 | 77 | 81 | 78 |
| $A_2$ | 69 | 71 | 72 | 68 |
| $A_3$ | 72 | 78 | 80 | 74 |
| $A_4$ | 71 | 74 | 75 | 68 |

a) Assume first that the workers capabilities do not influence the number of units produced. We want to test if the capacities of the machines differ. Which model can we use? Perform the test. What is your conclusion?

b) If the workers' influence on the number of units produced were taken into account, what model should then be used and what test would you use?

c) (Calulate by hand!)
What estimator should be used for the number of produced units by machine $M_2$? Also find a 90% confidence interval for the expectation when you use the estimate of $\sigma^2$ from the model in b).

**MINITAB:** If the data are put in column C1 in Minitab, A's levels (1,2,3,4) in C2 and M's levels in C3 one may use the following commands:

One-way
Stat → ANOVA → One-way
Response: C1
Factor: C3

Two-way without interactions
Stat → ANOVA → Two-way
Response: C1
Row factor: C2
Column factor: C3

**R:**

```
WMds <- data.frame("units"=c(76,77,81,78,69,71,72,68,72,78,80,74,71,74,75,68),
"workers"=as.factor(rep(1:4,each=4)),"machines"=as.factor(rep(1:4,4)))
# one way anova
anova(lm(units~machines,data=WMds))
# two way anova
obj2w <- lm(units~machines+workers,data=WMds)
anova(obj2w)
```

## Problem 2

Poor quality of a thread may cause it to break when weaving material. The quality of a thread is here seen as the ability not to break.

The kind of thread that is used for weaving a particular cloth can be spun of one of three kinds of cotton $A_1, A_2, A_3$ and of one of four kinds of silk $B_1, ..., B_4$. A cloth was weaved using a certain length of thread for each of the 12 possible combinations of thread $A_iB_j$ and the experiment was repeated once. The number of times $Y_{ijk}$ the thread broke during one length was recorded.

The results were:

Let $Y_{ijk}$ be the number of times the thread broke during weaving in the $k$th trial, $k = 1, 2$ for the combination $A_iB_j$. $Y_{ijk}$ are assumed to be independent and normally distributed with the same unknown variance $\sigma^2$ and expectation

$$E(Y_{ijk}) = \mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \text{ where } \sum_i \alpha_i = \sum_i \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$$

| Cotton/Silk | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|---|---|---|---|---|
| $A_1$ | 65 | 76 | 63 | 62 |
| | 68 | 69 | 59 | 69 |
| $A_2$ | 61 | 69 | 61 | 72 |
| | 63 | 62 | 66 | 71 |
| $A_3$ | 51 | 57 | 61 | 61 |
| | 53 | 54 | 52 | 67 |

a) Explain what the parameters $\mu$, $\alpha_i$, $\beta_j$ and $\gamma_{ij}$ represent and find estimators for the parameters.

b) Set up the ANOVA. How would you test if there is a significant interaction between the cotton and slik?

c) Then test if $A$ or $B$ has significant effect. Comment on the results of the analysis.

### Minitab

The response values are put in C1 and the level values for A in C2, and the level values for B in C3. Then use the command: Stat→ ANOVA → Two-way, and select response, row and column factors. Do not tick of "fit additive model" unless you do not want to model the interation between A and B.

Alternatively, data are available in a file `DATA9_1.MTW` from the www-page.

### R

Read in the data in three separate vectors, one for the reponse, one for the coding av factor A and one for the coding of factor B. Then make a data frame with these three columns and use `lm` and `anova` to perform a two-way ANOVA with interactions.

```
dsy <- c(65,68,61,63,51,53,76,69,69,62,57,54,63,59,61,66,61,52,62,69,72,71,61,67)
dsA <-c(rep(c(1,1,2,2,3,3),4))
dsB <-rep(1:4,each=6)
ds <- data.frame(dsy,as.factor(dsA),as.factor(dsB))
colnames(ds)<- c("y","A","B")

res <- anova(lm(y~A*B,data=ds))
```