TMA4255 Applied Statistics
Exercise 9

Problem 1

One wants to examine the machine's 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 σ^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: Factorial experiments

Use statistical software to solve this problem, see end of problem for hints for commands in MINITAB and R.

One wants to examine if small changes on 4 critical dimensions in a carburettor will affect the horse powers produced with a standard engine with six cylinders. The data from a 2^4 factorial experiment are given below

Dimensions				Response
A	B	C	D	y
-	-	-	-	14.6
+	-	-	-	24.8
-	+	-	-	12.3
+	+	-	-	20.1
-	-	+	-	13.8
+	-	+	-	22.3
-	+	+	-	12.0
+	+	+	-	20.0
-	-	-	+	16.3
+	-	-	+	23.7
-	+	-	+	13.5
+	+	-	+	19.4
-	-	+	+	11.3
+	-	+	+	23.6
-	+	+	+	11.2
+	+	+	+	21.8

a) Estimate the main effects and the interactions using statistical software. Plot main effects and two factor interactions.

b) Write down the regression model that corresponds to this analysis

c) Why is there no result for s^2 in the software output? Assume that $\sigma^2 = 4$ is known from experience. Which effects are now significantly different from 0? Find a 95% confidence interval for the most important effects.

d) If you assume that all three-way and four-way interactions are 0, how can you then estimate σ^2 and σ^2_{effect}? How can you now find the significant effects? Show the theory and do the analysis using statistical software.

MINITAB

Construction of design:

The following commands will produce the same design and in the same order as in the exercise:

Stat → DOE → Factorial → Create Factorial Design:
Number of factors: 4
Type of design: 2-level factorial (default generators)
Designs: Full factorial
Options: [] Randomize runs
x Store design in worksheet

Put the values of the response in C9 afterwards (yes, manually enter the 16 numbers into the column, and remember MINITAB use , and not . as decimal separator). If we would perform real experiments we would randomize them, but to get a clearer overview we skip it here.

Analysis: When interactions up to order 4 is included, all possible effects are included and

Stat → DOE → Analyze Factorial Design:
Terms: 2 or (4)
Graphs: x Normal
x Pareto
x Normal plot
x Residual vs fits
x Residuals vs variables: A B C D

the model will be fitted perfectly. But there will be no degrees of freedom left to estimate the variance of the error. When interactions up to order 2 is included, the variance of the error can be estimated from the interactions of order 3 and 4. Since the fitting is no longer perfect, we can also make residual plots. The table of variance appears automatically.

R

You first need to install the library `FrF2` from the packages tab, or writing `install.packages(''FrF2'')`, answer where to put the library (if you are not superman on you machine) and choose Norway=50 for download. To load the library either write `library(FrF2)` or tick off the package at the packages tab in the lower right window of Rstudio.

Construction of design:

```
plan <- FrF2(nruns=16,nfactors=4,randomize=FALSE)
y <- c(14.6,24.8,12.3,20.1,13.8,22.3,12.0,20.0,16.3,23.7,13.5,19.4,11.3,23.6,11.2,21.8)
plan <- add.response(plan,y)
```

If we would perform real experiments we would randomize them, but to get a clearer overview (when adding the response) we skip it here. Now we have an ordinary data set up to be used with lm, as we know from the regression part of the course.

For parts d and e, we add the argument `blocks` to the `FrF2` function. For 4 blocks we need to allow for aliasing with two-factor interactions (to the block generator).

```
design1<-FrF2(16,4,blocks="ABCD",randomize=FALSE)
summary(design1)
design2 <-FrF2(16,4,blocks=4,alias.block.2fis=TRUE))
summary(design2)
design3 <-FrF2(16,4,blocks=c("ABC","AD"),alias.block.2fis=TRUE))
summary(design3)
```

Analysis: Now we fit linear models, but the notation to include 4th order terms is $(.)^4$ and to include up to 2nd order terms is $(.)^2$. And, remember that effects are 2∗coefficients in the regression.

```
lm4 <- lm(y~(.)^4,data=plan)
summary(lm4)
anova(lm4) # to see the seqSS mentioned in the solutions to d)
effects <- lm4$coeff*2
```

For Normal plots showing the effects this called `DanielPlot`, plot of main effects is called `MEPlot` and plots of interactions are called `IAPlot`. I have not yet found a nice Pareto-plot, let me know if you see one (the sorted barplot with effects).

```
DanielPlot(lm4,half=FALSE,alpha=0.05)
MEPlot(lm4)
IAPlot(lm4)
```