TMA4255 Applied Statistics
Exercise 6

Observe: MINITAB and R commands in the end of Problem 2, R-script from the course www-page.

Problem 1: A simple 2^2 experiment

This problem should be solved using pen and paper! No software.

Supose in a two-level experiment with two factors (regressors) z_1 (for factor A) and z_2 (for factor B) the design matrix is given as

Experiment no.	Const.term	A	B	AB
1	1	-1	-1	1
2	1	1	-1	-1
3	1	-1	1	-1
4	1	1	1	1
	const	z_1	z_2	z_1z_2

Assume the model is given by $Y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_{12} z_1 z_2 + \varepsilon$, $\varepsilon \sim N(0, \sigma^2)$. Let us for simplicity assume that all ε are set to 0.

a) Find the main effects of z_1 and z_2 and their two-factor interaction $z_1 \cdot z_2$.
That is, first write down the expression for the expected effect for z_1 as a formula using y_1, y_2, y_3, y_4, that is, assume this effect is $(y_2 + y_4)/2 - (y_1 + y_3)/2$ (often we call this A). Then use the regression equation to replace the ys by βs and z's. Observe that you get the answer $2\beta_1$. Do the same for the other two effects listed.

b) Suppose you just run the two first experiments only. That is, while experimenting with z_1 you keep z_2 at its low level. What would then be the main effect of z_1? What would be the main effect of z_1 if you instead keep z_2 at its high level? Do this in the same manner as you did for a).

c) What does the results in b) tell you about varying one factor at a time when interactions are present?

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

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

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

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. (These last two questions you need to do by hand.)

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.

Warning: Blocking will not be covered before after the supervision of this exercise, so items e and f below should be done at a later date.

e) Assume now that the experiment is to be done in two blocks. Let the blocks be determined by the four-factor interaction ABCD. Which effects can now be estimated unconfounded with the block effect?

f) How would you perform the experiment in four blocks? (Try different options).

MINITAB

Construction of design:

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

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.

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

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)
```