TMA4255 Applied Statistics
Exercise 8

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	$z_1 z_2$

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

Use the data given in Table 1. We assume that $X_1, ..., X_n$ and $Y_1, ..., Y_m$ all are independent and normally distributed:

$$E(X_i) = \mu_X \; Var(X_i) = \sigma_X^2, \; i = 1, ..., n$$

$$E(Y_j) = \mu_Y \; Var(Y_i) = \sigma_Y^2, \; j = 1, ..., n$$

Assume that $\sigma_X^2 = \sigma_Y^2$, but unknown.

From A (X_j)	5179	5203	5207	5195	5207	5202	5203	5208	5216	5193
From B (Y_j)	5190	5159	5153	5206	5168	5186	5194	5200		

Table 1: Tensile strength for copper wires

a) Put the data into your statistical software (MINITAB or R) and perform a two sample t-test.
Write down the expressions for the statistics computed.
What is being tested here?
What is the conclusion when the significance level of the test is 1%?

b) By using one-way analysis of variance one can examine if the tensile strength of the copper wires are not equal.
Perform the test.
Explain the statistics in the output and what is being tested here. Why is the p-value for this test the same as in a)?

MINITAB:

Stat → Basic Statistics → 2-Sample t
Samples in different columns (C1 C2)
Alternative: Not equal
√ Assume equal variances

Stat → ANOVA → One-way (Unstacked)
Responses: C1 C2
Graph: three in one

R:

```
dsA <- c(5179,5203,5207,5195,5207,5202,5203,5208,5216,5193)
dsB <- c(5190,5159,5153,5206,5168,5186,5194,5200)
t.test(dsA,dsB,var.equal=TRUE)
copper <- c(dsA,dsB)
coppergrp <- c(rep("A",length(dsA)),rep("B",length(dsB)))
obj <- lm(copper~as.factor(coppergrp))
anova(obj)
plot(obj)
```