TMA 4255 Applied Statistics

Exercise 9

MINITAB/R commands are at the end of each problem.

Problem 1

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)?

c) Are there different model assumptions if you perform a two-sample t-test or a one-way ANOVA with one factor with two levels?

Investigate if the model assumptions are met for analysis of the copper data using a one-way ANOVA.

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

Problem 2

The content of copper in 4 different bronze alloys is to be compared. The following observations have been collected.

Data

Row	CU_1	CU_2	CU_3	CU_4
1	83,09	83,01	83,02	83,01
2	83,04	82,96	83,10	82,90
3	83,06	82,99	83,05	83,04
4	83,04	83,03	83,04	83,01
5	83,05	83,00	83,08	82,96
6	83,03	82,97		
7	83,05	82,99		
8	82,98			

a) Make a Boxplot for the 4 bronze-alloys. Do you find any reason to believe there might be differences. Are there potential outliers.

b) Perform a one way analysis of variance. What is the conclusion from the F-test, using a significance level of 5%. Do you find any serious violence of the assumptions.

c) Find out how the alloys differ from each other.

You can load the data into MINITAB by the command:
https://www.math.ntnu.no/emner/TMA4255/2023v/files/Bronzealloys.txt and press OK
Stat → ANOVA → One-way (data in separate columns)
Responses C1-C4 and choose from the menu.

You can read the data into R using the following commands:
copper=c(83.09,83.04, ... ,82.96)
Denoting the four types of alloys as A, B, C and D you can make a column
alloys=c("A", ... ,"A","B",...,"B", "C", ...,"C","D", ... ,"D")
Then the following commands may be useful

```
> boxplot(copper~alloys)
> mod1=aov(copper~alloys)
> summary(mod1)
> qqnorm(mod1$residuals)
> plot(mod1$fit,mod1$residuals)
> TukeyHSD(aov(copper~alloys))
```

You can load the data into MINITAB by the command:
https://www.math.ntnu.no/emner/TMA4255/2023v/files/Bronzealloys.txt and press OK
Stat → ANOVA → One-way (data in separate columns)
Responses C1-C4 and choose from the menu.

You can read the data into R using the following commands:
copper=c(83.09,83.04, ... ,82.96)
Denoting the four types of alloys as A, B, C and D you can make a column
alloys=c("A", ... ,"A","B",...,"B", "C", ...,"C","D", ... ,"D")
Then the following commands may be useful

```
> boxplot(copper~alloys)
> mod1=aov(copper~alloys)
> summary(mod1)
> qqnorm(mod1$residuals)
> plot(mod1$fit,mod1$residuals)
> TukeyHSD(aov(copper~alloys))
```