TMA4255 Applied Statistics
Exercise 2

Problem 1

A new type of tyres, N tyres, is now available. A bus company wants to determine whether the new type has better quality than the old type, G tyres, which has been used so far.

The company conducts the following experiment: For each of the 9 buses, one of the rear tyres is randomly provided with a N tyre and the other one with a G tyre. After a certain time of driving the wear (i.e. the decrease in the depth of the grooves) is measured on each of the 18 tyres. The wear on the N tyre and G tyre from bus j is denoted by N_j and G_j respectively and we define $D_j = G_j - N_j$, $j = 1, ..., 9$.

$D_1, ..., D_9$ are assumed to be independent and normally distributed with expectation δ and variance $\sigma^2 = 36$. On the basis of the observed values of $D_1, ..., D_9$ one wants to say if a N tyre on the average has better quality than a G tyre.

Bus no. j	1	2	3	4	5	6	7	8	9
G tyre (G_j)	35	45	61	89	58	74	120	34	55
N tyre (N_j)	31	42	55	79	59	79	108	36	46
D_j	4	3	6	10	-1	-5	12	-2	9

Table 1: Results from the bus experiment

a) Why would it be natural to treat this situation as a hypothesis testing problem with $H_0 : \delta = 0$ and $H_1 : \delta > 0$?

What is the reason that one has chosen the experimental plan described above?

b) Suggest a test for the hypothesis testing problem based on $\bar{D} = \frac{1}{9}\sum_{j=1}^{9} D_j$. What is the conclusion if the results of the experiment are as given in Table 1? Use a significance level of 5 %.

c) How would the test be carried out if the assumption in the introduction is changed so that σ^2 is unknown?

d) How would the test in a) be carried out if there were 9 buses with N tyres and 9 buses with G tyres?

What would be the conclusion if the results were as in Table 1 and all N_j and G_j are assumed to be independent and normally distributed with the same unknown variance? Also look at the case when the variances are assumed not to be equal.

Compare the results in b) and c) with the results in d) and comment.

Problem 2

Two manufacturers A and B produce the same kind of copper wires. A customer is interested in knowing whether there is any difference in the tensile strength of the copper wires from the two manufacturers, and has therefore measured the strength in a random sample from A's and B's production.

Let $X_1, ..., X_n$ denote the tensile strength in a random sample from A's production and let $Y_1, ..., Y_m$ denote the tensile strength in a random sample from B's production.

The data are given in Table 2.

It is assumed that $X_1, ..., X_n$ and $Y_1, ..., Y_m$ are all independent and normally distributed.

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

$$\mathrm{E}(Y_j) = \mu_Y,\ \mathrm{Var}(Y_j) = \sigma_Y^2,\ ,j = 1, ..., m$$

a) Assume that σ_X^2 and σ_Y^2 are unknown. We want to test

$$H_0 : \sigma_X^2 = \sigma_Y^2 \text{ vs. } H_1 : \sigma_X^2 \neq \sigma_Y^2$$

What is the conclusion if the significance level of the test should be 5% and the observations are as in Table 2?

b) Find a 90% confidence interval for σ_X^2/σ_Y^2 and compare this interval with the result in a). How can you find a 90% confidence interval for σ_X/σ_Y?

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 2: Tensile strength for copper wires

Minitab hints

Both data sets are found in the file `data2.MTW`.

Problem 1: Z-test, t-test and two sample t-test can be found under

Stat → Basic Statistics

Problem 2: The quantiles for the F-test can be found under

Calc → Probability distributions → F

R hints

Problem 1: Data are found in
tt dataE2P1.txt, and can be read using
`ds <- read.table("dataE2P1.txt",header=TRUE)` after saving the file to the appropirate

place, or directly from the www using

```
ds <- read.table("http://www.math.ntnu.no/~mettela/TMA4255/Data/dataE2P1.txt",header=TRUE)
```

A z-test (known sigma) is not available in R (since this situation practically never arises except with the purpose of teaching...), so you can either program your own function, or download the TeachingDemos library and use the z.test in this library.

To install the TeachingDemos library you write `install.packages("TeachingDemos")` and choose Norway for downloading. Then start the library by typing `library(TeachingDemos)`. Then `?z.test` will give you help about the z-test.

`t.test`: both one and two sample.

Problem 2: Data are found in
tt dataE2P2.txt with the 10+8 numbers as one long vector. Read them all (after you have saved the file) and divide into to vectors by

```
ds <- scan("dataE2P2.txt")
dsA <-ds[1:10]
dsB <- ds[11:18]
```

Or, reading directly from www by

```
ds <- scan("http://www.math.ntnu.no/~mettela/TMA4255/Data/dataE2P2.txt")
```

`var.test` for F-test, and `qf` for critical values in the F-distribution.