IMBS 1 **ISCID-CO Dunkerque/ULCO**

Mathematics applied to economics and management
Foundations of Descriptive and Inferential Statistics

January 2016 - Final assessment - Session 2 - Semester 1
Time allowed : 1h30 - All documents allowed

Exercice 1 Create the vectors :

1. (1,2,3,...,19,20)
2. (20,19,...,2,1)
3. (1,2,3,...,19,20,19,18,...,2,1)
4. (4, 6, 3) and assign it to the name tmp.

For parts 5., 6. and 7., look at the help for the function rep.

5. (4,6,3, 4,6,3,...,4,6,3) where there are 10 occurrences of 4.
6. (4,6,3, 4,6,3,...,4,6,3,4) where there are 11 occurrences of 4, 10 occurrences of 6 and 10 occurrences of 3.
7. (4,4,...,4, 6,6,...,6, 3,3,...,3) where there are 10 occurrences of 4, 20 occurrences of 6 and 30 occurrences of 3.

Exercice 2 The file "hills.txt" stored in your workspace contains the Scottish hill races data set.

1. Read the file, assigning the result to the object `hills`.
2. Examine the object. Note column and row names
3. Make the columns of the hills object available by name
4. Construct a scatter plot. The function call in this way means the first argument is the horizontal axis.
5. Compute a linear regression of time against distance.
6. Obtain more information about the regression.
7. Add the least squares regression line - note anonymous function call.
8. Obtain some diagnostics plots - note the different arguments to the plot function. Be aware of the prompt in the Console.
9. Create a copy of the hills object.
10. Quit the session.

Exercice 3 Assume that we have registered the height and weight for four people : heights in cm are 180, 165, 160, 193 ; weights in kg are 87, 58, 65, 100. Make two vectors, height and weight, with the data. The bodymass index (BMI) is defined as weight in kg / (height in m)^2. Make a vector with the BMI values for the four people, and a vector with the natural logarithm to the BMI values. Finally make a vector with the weights for those people who have a BMI larger than 25.

Exercice 4

1. In the R package MASS there is a dataset called `cats`. Run the following commands :

```
> library(MASS)
> data(cats)
```

Have a look at the dataset. The variables Bwt and Hwt give the weight of the body (kg) and the heart (g), respectively. There are both male and female cats. Make a dataset with the data from males only (use the function `split`).

2. Make a scatterplot of the data for the male cats (Bwt on x-axis, Hwt on y-axis). Does it look reasonable to use a linear regression model for the data ?
3. Fit a linear regression model (`regModel`) for the male cats, that allows for prediction of the heart weight given the body weight. Add the fitted regression line to the scatterplot from the previous question.
4. Find the coefficients of the fitted line. How large is the expected difference in heart weight for two cats with a difference of 1 kg in bodyweight ? How large is the expected difference in heart weight for two cats with a difference of 100 g in bodyweight ?
5. Use the estimates to find the expected heart weight for a male cat that weights 3 kg.