Battery capacity

Exam exercise: Battery capacity

In this exercise you will study a data set on battery capacity collected by researchers from AAU Energy. The batteries were kept at three different temperatures. The capacity loss Q (in percentage of initial capacity) is the response variable. We are interested in the effect on the response of the temperature and the efficient number of charge cycles FEC that the battery had been exposed to.

In the analysis, we will work with the log transformed variables `logQ` and `logFEC`. Moreover, we will consider `Temperature` as a categorical variable with the three levels 35°, 40°, and 45° Celcius.

We begin by loading the `mosaic` package and reading in the data. Since we consider `Temperature` a categorical variable, we make this variable a factor.

```
library(mosaic)
capacity<-read.delim("https://asta.math.aau.dk/datasets?file=capacity.txt",sep="")
capacity$Temperature<-factor(capacity$Temperature)
```

1. Make a scatter plot of the log transformed variables `logQ` and `logFEC` where points are colored by temperature level.

2. The code below finds the correlation between `logQ` and `logFEC` for each level of temperature. Explain the concept of correlation and interpret the results in relation to the plot from Question 1.

```
tempGrupper=split(capacity,capacity$Temperature)
lapply(tempGrupper,function(x) cor(x$logQ,x$logFEC))
```

```
## $`35`
```

```
## [1] 0.9813515
##
## $`40`
## [1] 0.9967649
##
## $`45`
## [1] 0.9943055
```

3. Consider a multiple regression model without interaction with `logQ` as the response variable and `logFEC` and `Temperature` as predictors. Write out the model equation using dummy variables. What is the interpretation of the parameters?

4. Edit the code below to fit the model from Question 3. Explain the output from the code. Your explanation should as a minimum include:

- What is the prediction equation?
- Explain the calculation of `t value` and determination and interpretation of the p-value.
- What is the interpretation of `Multiple R-squared`?
- Make an overall F-test for the null-hypothesis that there is no effect of any of the predictors.

```
# model<-lm(...~...,data=...)
# summary(model)
```

5. Investigate whether or not there is an interaction between the effect of `Temperature` and the effect of `logFEC` as predictors of `logQ` (Hint: With more than two groups you need to use the `anova` function to make the test, see p. 31 in the slides from the second lecture of the module).

```
# model1<-lm(...~...,data=...)
```

6. Plot the model with and without interaction and explain the difference.

7. The code below computes the residuals $y_i - \hat{y}_i$ in the model without interaction. Draw a boxplot of the residuals and comment on the result.

```
# residuals<-model$residuals
```