Exercises, lecture 3

19. februar 2010

To find the p-values asked for in the questions you probably need to use Matlab (see Matlab commands below). Approximate answers can sometimes be found by using tables in the book.

Exercise 1 Problem 22.4.1 in the book. Find the p-value.

Exercise 2 Problem 22.4.3 in the book. Find the p-value.

Exercise 3 Problem 22.4.9 in the book. Find the p-value.

Exercise 4 Problem 22.4.12 in the book. Find the p-value.

Exercise 5 Find the p-value in problem 22.3.5 in the book (solved last time).

Exercise 6 Find the p-value in problem 22.3.10 in the book (solved last time).

Exercise 7 Matlab exercise:

1. Download the two data files data1.txt og data2.txt from:
   `http://www.math.aau.dk/~rubak/teaching/2010/nano4/data/`
   (right click one the files and save them to your computer)
2. Load data1 and data2 in Matlab using the commands `x=tblread('data1.txt')` and `y=tblread('data2.txt')`. (Remember to have the files in the working directory.)
3. Make QQ-plots of the two datasets. Do they look normal?
4. The first dataset is the BMI of 25 students measured the day before the Christmas break. The second dataset has the BMIs from the day after the break.
5. Is there a significant change in BMI? What is the p-value?
6. Assume instead the measurements are respectively the girls' and the boys' BMI in a class of 50 students.
7. Is there a significant difference in BMI? What is the p-value?

Exercise 8 Unfinished exercises from previous lectures.

Extra Matlab commands:

In the following are a couple of extra Matlab commands. You still need to refer to the list of commands on the exercise sheet from last lecture.

- `q = chi2inv(p, df)` returns the p quantile of the χ^2 distribution with `df` degrees of freedom.
- `p = chi2cdf(x, df)` calculates $P(X \leq x)$ for $X \sim \chi^2(\texttt{df})$. This is used to find p-values.
- `p = tcdf(x, df)` calculates $P(X \leq x)$ for $X \sim t(\texttt{df})$. This is used to find p-values.
- `p = normcdf(x, m, sigma)` calculates $P(X \leq x)$ for $X \sim N(\texttt{m}, \texttt{sigma}^2)$. This is used to find p-values.
- `[h,p,ci,stats] = ttest2(x, y, alpha)` tests the null hypothesis of `x` and `y` having same mean at significance level `alpha`. The returned results are:
  1. `h` is 0 if the null hypothesis is accepted and 1 if it is rejected (the alternative is accepted).
  2. `p` is the 'p'-value.
  3. `ci` is the confidence interval for the mean.
  4. `stats` is the value of the test statistic t.
- `qqplot(x)` draws a QQ-plot of the values in `x` compared to the normal distribution.