Seminar für Finanzökonometrie
Institut für Statistik
Akademiestr. 1/I
http://www.finmetrics.statistik.uni-muenchen.de/studium_lehre
WiSe 2015/16
Port, Henry
Based on exercises from Steffen Unkel

Statistical Geophysics
Exercise Sheet 4

Exercise 1 In this exercise we will investigate again some rain data collected at the volcano Merapi (Indonesia). The file *rain.dat* contains data on the amount of rain (in litres per square metre), measured at two stations (Jrakah and Kaliurang). Suppose the days on which the measurements were taken are different for both stations. The random variables

$$Y_A = \text{rain at station Jrakah}$$

and

$$Y_B = \text{rain at station Kaliurang}$$

are independent and follow a normal distribution with equal variances:

$$Y_A \sim \mathcal{N}(\mu_A, \sigma^2), \quad Y_B \sim \mathcal{N}(\mu_B, \sigma^2)$$

a) We would like to test the research hypothesis that the amount of rain at station Jrakah is different from the amount of rain at Kaliurang ($\alpha = 0.05$). Formulate the null and alternative hypotheses for this test problem.

b) Calculate the test statistic and make a decision whether or not the null hypothesis can be rejected. Do the following:

   i) Read the data into R and determine means and empirical variances.

   ii) Calculate the pooled empirical variance and calculate the test statistic.

   iii) To which value do you have to compare the test statistic in order to get a decision?

   iv) Is there a faster way in R to carry out a Student's t-Test?

c) Carry out a test for the hypothesis that there is less rain at station Jrakah than at station Kaliurang ($\alpha = 0.05$).

d) Suppose the 20 days on which the measurements were taken are the same for both stations, that is, the samples are dependent. Carry out an appropriate test to check

the assumption that there is a difference in the amount of rain at the two stations ($\alpha = 0.05$).

Exercise 2 The carrier of a public transport network used to assume that the proportion of passengers with a season ticket is 35%. In a survey 112 of 350 respondents turned out to be owners of a season ticket. Carry out an appropriate statistical test at the 5%-level to check the research hypothesis that the proportion of the owners of season tickets is different from 35%.

Exercise 3 A child protection agency investigates the situation of foster children. The question of interest is if foster children are better integrated in a family when foster parents also have children of their own. Eight foster parents with own children and six foster parents without own children took part in the study. An integration score was developed by means of a questionnaire. The higher the score, the better the child is integrated in the family. The following scores were identified:

Foster Parents	Score							
With own children x_i	8	13	16	20	24	17	18	25
Without own children y_i	12	9	13	11	19	15		

Check the question above with the Wilcoxon rank-sum test ($\alpha = 0.1$). You may use the fact that $w_\alpha(n, m) = w_{0.1}(8, 6) = 50$ and $w_{1-\alpha}(n, m) = n(n + m + 1) - w_\alpha(n, m)$. Note that in R the command $qwilcox(0.9, 8, 6)$ will return a critical value that is not appropriate for comparison with the observed test statistic of the Wilcoxon rank-sum test introduced in the lecture.

Exercise 4 In an empirical study examining the smoking habits ten smokers were asked how many cigarettes they smoke on average per day. The results of this study are as follows:

26 34 5 20 50 44 18 39 29 19

a) Use the Wilcoxon signed-rank test to examine the research hypothesis that the median of the number of smoked cigarettes per day is larger than 25 ($\alpha = 0.1$). You may use the critical value $w^+_{1-\alpha}(n) = w^+_{0.9}(10) = 39$. Note that in R the command $qsignrank(0.9, 10)$ returns a value of 40. In the tutorial we will also discuss the one-sample sign test as a an alternative to the Wilcoxon signed-rank test.

b) Suppose that the number of smoked cigarettes per day is normally distributed. Carry out an appropriate parametric test and compare your result with the results you obtained in (a).