140207

Examination – Statistics for the course: The research process

Ignacio Rangel. Clinical Medicine. Örebro University

Bengt Löfstrand

Score: pass with 62.5 % of overall score, e.g. 10p

It is allowed to bring in: a book in statistics and a pocket calculator

1.- indicate if the following statements are true or false –no further comments to the answers are required- **total 3p**

a) In a two tailed test (assume a t-test) with a significance level of 1%, the risk for a Type I error is equal to 0.05

(0.5p)

b) The basic first steps in hypothesis testing are the formulation of the null hypothesis and the alternative hypothesis. Failure to reject H_0 after performing the test statistics leads to its acceptance

(0.5p)

c) Small p values indicate large effects

(0.5p)

d) In the analysis of two quantitative variables, a numerical indicator that explains the variation of *y* based on the variation of *x* is ρ

(0.5p)

e) The following combinations of correlation- and regression-coefficients are possible:

1.- r=1 b=0,5

2.- r= -0.5 b= -1

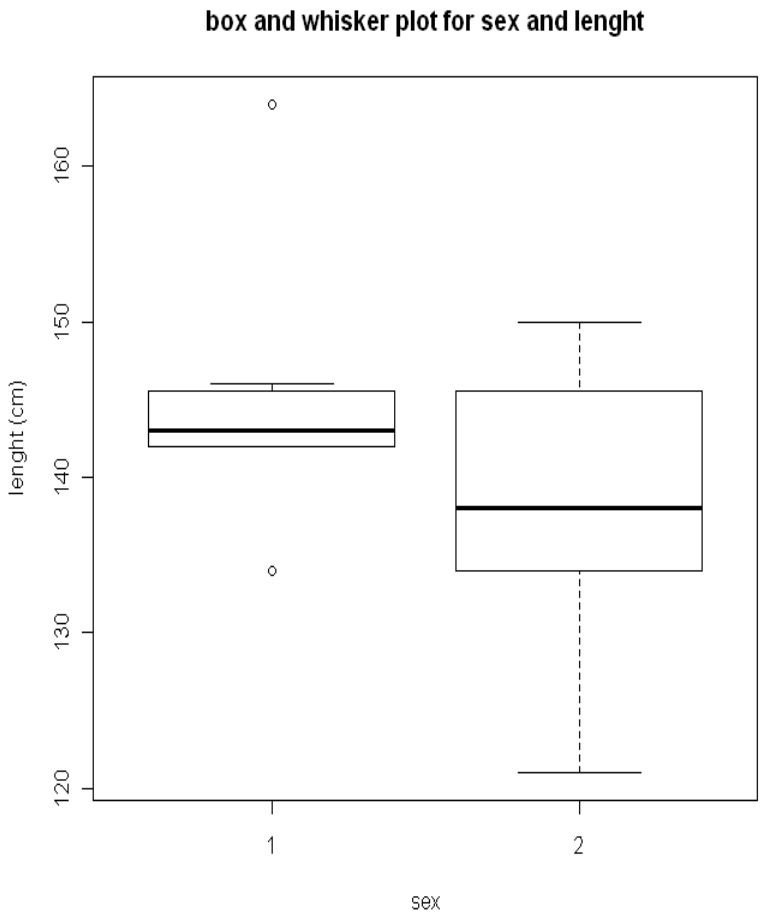
(0.5p)

f) Hypothesis testing: by changing α from 0,05 to 0,01 we decrease the risk to incorrectly claim that there is a difference between, i.e. two treatments when in reality there are no such differences. Thus, we increase strictness of the test analysis and by doing so it is easier to miss differences that in reality exist.

(0.5p)

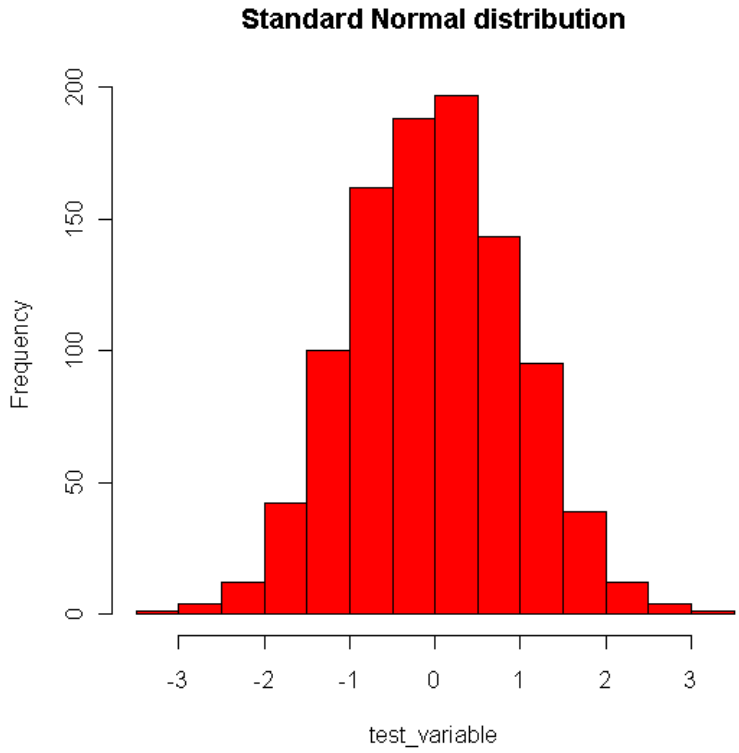
2.- Have a look at the next figure. It is the results of anthropometric measurements carried out on a group of 10-year old children. Sex 1 = boys, sex 2 = girls. Describe the results and explain how to interpret the different parts of the box and whisker plot.

(1.5p)



3.- The next figure represents a histogram of a standard normal distribution. Explain it!

(1.5 p)



4.- In a study where researchers wanted to investigate if there was a correlation between exposure for passive smoking and miscarriage, a total of 250 women were asked if they were exposed for passive smoking. The answer alternatives were “never”, “sometimes” and “often”. Results were as follows: 40 answered “never”, 120 answered “sometimes” and the rest answered “often”. **Total 2p**

a) What kind of data is this? Make a figure to summarise the results, present an “average” point as well as a distribution measurement of the data

(2p)

5.- In relation to blood grouping, it is known that approximately 15% of the Swedish population is Rh- negative. A sample consisting of 40 persons with a particular disease was 20% Rh-negative. Was the Rh-negative proportion of people with this disease different from the rest of the Swedish population? Base your conclusion on the application of hypothesis testing development. **Total 3p**

(3p)

6.- Assume that the weight among 10-year old girls is normally distributed and that the average weight is 33,5 kg. From a free-school in a town in southern Sweden, where children are served only vegetarian food, the average weight among a randomly selected sample of 41 students (same age and sex) was 32,3 kg (standard deviation = 3,75 kg). Is it possible to conclude that the average weight in the sampled free-school was different from the standard average weight of 10-year old girls? Base your conclusion on the application of hypothesis testing development. **Total 3p**

(3p)

Questions from: Bengt Löfstrand

(2p)

1 Glaucoma is an eye desease characterized by increased pressure in the eye. You made a study testing the hypothesis that these patients have higher blood pressure than normal. 200 patients with glaucoma were in the study and the mean of systolic blood pressure was 140 mm Hg and standard deviation 25 mm Hg.

a Give a 95 % confidence interval for the true value of blood pressure for patients with glaucoma.

b The mean of blood pressure for people without glaucoma and in the same ages is 130 mm Hg. Referring to this study, is there a correlation between glaucoma and blood pressure?