Statistical Methods Fall 2016

Assignment 3: Estimation and testing

Deadline: December 8, 23.59h

Topics of this assignment

The exercises below concern topics that were covered in Lectures 5, 6, 7 and 8: estimation of population proportion and mean, tests for population proportion or mean based on one sample, tests for differences of two means or proportions based on two samples (see the respective sections in Chapters 6, 7, and 8 of the book and the slides of Lectures 5–8). Before making the assignment, study these topics. Numbers of exercises in the book refer to the twelfth edition (Pearson New International Edition).

How to make the exercises? See Assignment 1.

If you are asked to perform a test, do not only give the conclusion of your test, but report:

– the hypotheses in terms of the population parameter of interest;
– the significance level;
– the test statistic and its distribution under the null hypothesis;
– the observed value of the test statistic (the observed score);
– the P-value or the critical values;
– whether or not the null hypothesis is rejected and why.
If applicable, also phrase your conclusion in terms of the context of the problem.

Theoretical exercises

For the three theoretical exercises below use Tables 2 and 3 from the Appendix in the book to find critical values. Do not use R. If you need to use a t-distribution with the number of degrees of freedom not included in Table 3, report the number of degrees of freedom, and use the critical value based on a t-distribution with the next lower number of degrees of freedom found in the table.

Exercise 3.1 Exercise 32 from Section 6.2 (Fortune Tellers), use a confidence level of 95% (instead of 98%).

Exercise 3.2 Exercise 16 from Section 8.4 (Self-Reported and Measured Male Heights). Instead of constructing a confidence interval, use a significance level of $\alpha = 0.1$ to test the claim that males aged 12–16 correctly report their heights. Some characteristics of the data that you may or may not use are: $\bar{x}_1 = 65.583$, $\bar{x}_2 = 66.583$, $s_1 = 5.282$, $s_2 = 4.969$, $s_d = 3.520$.
(Follow the detailed instructions about testing presented above).

R-exercises

Hints concerning R:

- Recall: A sample from a certain distribution can be obtained in R with the function `rdist(n,par)` where `dist` stands for the name of the distribution, `n` for the sample size, and `par` for the relevant parameters: `x=rnorm(50,5,1)`, `x=rexp(25,1)`, `x=runif(30,-1,1)`, `x=rt(10,df=5)`, `x=rchisq(25,df=8)`. For example, the function `rnorm(n,mean,sd)` generates a sample of size `n` from the normal distribution with expectation `mean` and standard deviation `sd`. The parameters of the other distributions are documented in the help file.

- The R-functions `pnorm` and `qnorm` can be used for computing probabilities and quantiles of normally distributed random variables. Similarly, `pt(...,df=k)` and `qt(...,df=k)` can be used to compute probabilities and quantiles of a t-distributed random variable with k degrees of freedom.

- For computing a confidence interval based on a t-distribution and for performing a t-test the R-function `t.test` can also be used. For example, to perform a t-test with significance level 5% for testing the null hypothesis $\mu = 10$ against the alternative hypothesis $\mu > 10$, using the data set `example` use the command `t.test(example,mu=10,alt="greater")`. The arguments of `t.test` can be adjusted to test with other significance levels and other null and alternative hypotheses as well: study the help file of `t.test`.

- Note that the function `t.test` reports also the P-value. Use `t.test(...)$p.value` to access it without printing the whole output of `t.test()`.

- In order to get a proportion of elements of the vector `x` that are, for instance, less than some value `M` you can use `mean(x<M)`.

- The R-function `t.test` can also be used for *two*-sample problems: put the values of the two samples into two vectors, `x` and `y`, say; the command `t.test(x,y, ...)` performs a two-sample test (based on the t-distribution) and computes a confidence interval as well, with default significance level $\alpha = 5\%$. Values of other arguments, like `"paired = TRUE"` to perform a paired t-test, can be specified by inserting them at the position of the dots.

Exercise 3.3 Alice and Bob work evening shifts at a supermarket. Alice has complained to the manager that she works, on average, much more than Bob. The manager claims that on average they both work the same amount of time. After a short discussion between the manager and Alice, the manager randomly selected 50 evenings when Alice worked, and 50 evenings when Bob worked (not necessarily the same evenings as Alice). The datasets `Alice` and `Bob` in the file `Ass3.RData` contain the number of hours they have worked per evening.

a) Give a point estimate and a 95% confidence interval for the difference of mean working time per evening of Alice and Bob.

b) Investigate the manager's claim with a suitable test. Take significance level 5%. Motivate your choice. (See the first page of the assignment for detailed instructions about testing).

c) Now investigate Alice's claim with a suitable test. Take significance level 5%. Motivate your choice. (See the first page of the assignment for detailed instructions about testing).

d) Find the ratio of the P-values found in part b) and c). How can you explain this value?

Exercise 3.4 Alice from the previous exercise has another concern. By contract the employees are supposed to work 4 hours per evening. Alice claims that the proportion of evenings on which she worked more than 3.5 hours is larger than the proportion of evenings on which Bob worked more than 3.5 hours.

a) Based on the data in `Ass3.RData` give a point estimate for the difference in proportion of evenings that Alice and Bob have worked more than 3.5 hours.

b) Investigate Alice's claim with a suitable test. Take significance level 5%. Motivate your choice. (See the first page of the assignment for detailed instructions about testing).

Exercise 3.5 In this exercise we illustrate with simulated samples the concept 'probability of type II error' when testing the null hypothesis $\mu = 0$ versus the alternative hypothesis $\mu \neq 0$. For this we consider samples from $N(\mu, 1)$ for some $\mu \neq 0$.

For parts a)–c): report only a table with proportions of not *rejected hypotheses, and your R-code in the appendix.*

a) We first investigate the case $\mu = -0.5$. Generate a sample x of size 10 from $N(-0.5, 1)$. Test, as if you know that the underlying distribution is normal but do not know its location and standard deviation, with a t-test the null hypothesis that the expectation of the underlying distribution of x, i.e. the population mean, is equal to 0 against the alternative that it is not equal to 0. Take significance level 5%. Do this 100 times and count how many of the 100 times the null hypothesis was *not* rejected.

b) Repeat part a) with samples of size 25, 50, 100.

c) Now repeat part a) and part b) with $\mu = -0.2, 0.3, 1$.

d) Explain how what you did in parts a)–c) is related to the 'probability of type II error'.

e) Do you see a global pattern in the results of the parts a)–c)? If so, which pattern?