MIOT H5014

Statistical Analysis for Engineers

Assignment 2 on Hypothesis Testing

Semester 2, 2016/7

Summary

This assignment involves writing a computer program. It will generate a number of data values from a normal distribution with known parameters. The program will then calculate a confidence interval for the mean and variance. This step will be repeated a large number of times and the probability of the parameters being in their Intervals will be empirically calculated.

Purpose of the program

Your program should do the following:

1. Read in a sample size number N_1.
2. Ask the user for an appropriate level of significance α.
3. Generate N_1 values of a normally distributed variable X and calculate the values of the sample mean and standard deviation.
4. Calculate the $100(1 - \alpha)\%$ confidence interval for this sample of values. Your programme should be able to deal with the possibility that N_1 is a large number or small.
5. Repeat the above steps a very large number of times N_2 and calculate the proportion of times the mean is inside the confidence interval. Comment on this result.

Submission and Assessment

Your answer should consist of:

1. A printout of the program code with a signed cover sheet stating that the code is your own work.
2. An email from your ITB email address with the code as text in the body of the email.
3. Your code should include brief comments within the body of the algorithm explaining the function of each loop, function or other elements of the program.
4. You may be asked to provide an explanation and a demonstration of the program working.

Note

In the C programming language, the function random() yields a pseudo-random number with certain properties. A good textbook on C will set this out this for you and show how to convert it to a uniform variable. Recall that if U is a uniform random variable then it can be transformed to a variable Z with the standard normal distribution by the transformation

$$X = \Phi^{-1}(U), \text{ where } \Phi \text{ is the CDF for N(0,1).}$$

Alternatively a spreadsheet may be used to generate a text file containing a very large number of values of a normally distributed random variable and the programme could read in this data. If you take the second option, you should show that your data is indeed normally distributed.