Homework questions, week 4

Econ 103

1 Daily Homework questions

The questions in bold font are due on **Wednesday 15th June**. You do not need to hand in the questions that are not in bold, though these will be useful to complete for your own understanding.

Lecture 12: Sampling Distributions and Estimation I

Textbook questions:
Chapter 6: **1, 5, 7**

Additional questions: none

Lecture 13: Sampling Distributions and Estimation II

Textbook questions:
Chapter 7: 1, **9**, 13, **17**, **18**, 19

Additional questions: none

The answer in the back of the book for 7-19 is wrong. *I will provide full solutions to 7-13 since it's hard, 7-18 since it's even-numbered, and 7-19 since the book is wrong.*

Lecture 14: Confidence Intervals I

Textbook questions:
Chapter 8: **1**, 3, 5, **7**

Additional questions:

1. **For this question assume that we have a random sample from a normal distribution with unknown mean but *known* variance.**

(a) Suppose that we have 36 observations, the sample mean is 5, and the population variance is 9. Construct a 95% confidence interval for the population mean.

(b) Repeat the preceding with a population variance of 25 rather than 9.

(c) Repeat the preceding with a sample size of 25 rather than 36.

(d) Repeat the preceding but construct a 50% rather than 95% confidence interval.

(e) Repeat the preceding but construct a 99% rather than a 50% confidence interval.

Lecture 15: Confidence Intervals II

Textbook questions: none

Additional questions:

1. **All other things equal, how would the following change the width of a confidence interval for the mean of a normal population? Explain.**

   (a) The sample mean is smaller.

   (b) The population mean is smaller.

   (c) The sample standard deviation is smaller.

   (d) The sample size is smaller.

2. **In this question you will carry out a simulation exercise similar to the one I used to make the plot of twenty confidence intervals from lecture 14. R Tutorial 4 will be useful for this question - I suggest you complete this before you do this question. To hand in this question, please email me a file with your R code in it (send your R script file).**

   (a) Write a function called `my.CI` that calculates a confidence interval for the mean of a normal population when the population standard deviation is known. It should take three arguments: `data` is a vector containing the observed data from which we will calculate the sample mean, `pop.sd` is the population standard deviation, and `alpha` controls the confidence level (e.g. `alpha = 0.1` for a 90% confidence interval). Your function should return a vector whose first element is the lower confidence limit and whose second element is the upper confidence limit. Test out your function on a simple example to make sure it's working properly.

   (b) Write a function called `CI.sim` that takes a single argument `sample.size`. Your function should carry out the following steps. First generate `sample.size` draws from a standard normal distribution. Second, pass your sample of standard normals to `my.CI` with `alpha` set to 0.05 and `pop.sd` set to 1. Third, return the resulting confidence interval. Test your function on a sample of size 10. (What we're doing

here is constructing a 95% confidence interval for the mean of a normal population using simulated data. The population mean is in fact zero, but we want to see how our confidence interval procedure works. To do this we "pretend" that we don't know the population mean and only know the population variance. Think about this carefully and make sure you understand the intuition.)

(c) Use `replicate` to construct 10000 confidence intervals based on simulated data using the function `CI.sim` with `sample.size` equal to 10. (Note that `replicate` will, in this case, return a matrix with *2* rows and *10000* columns. Each column correponds to one of the simulated confidence intervals. The first row contains the lower confidence limit while the second row contains the upper confidence limit.) Calculate the proportion of the resulting confidence intervals contain the true population mean. Did you get the answer you were expecting?

(d) Repeat the preceding but rather than using `CI.sim` write a new function called `CI.sim2`. This new function should be identical to `CI.sim` except that, when calling `my.CI`, it sets `pop.sd = 1/2` rather than 1. How do your results change? Try to provide some intuition for any differences you find.

2 R Tutorials

You should complete R Tutorial #4 by **Thursday 16th June**.

R tutorials will be posted in Piazza, with solution code.