Quantitative Methods II
Assignment 1
August 28, 2011

This is the first homework assignment for the course. Its purpose is to begin using a statistical program to get at the answers we need. All statistics are an attempt to gather information about a process (or population) from a sample of data generated by the process. As such, we will never be able to answer a statistical question with absolute certainty, only with statements of confidence ranges and of expected Type I Error rates.

This assignment covers tests for measures of central tendency, boxplots, and validating those tests, using Monte Carlo techniques.

If you need assistance using `R`, *do not* hesitate to ask for it. To get the most out of such assistance, you will need to explicitly explain your issue, attach the code you have already written, and ask earlier than Sunday.

⋆ ⋆ ⋆

When you hand in this assignment, you will email to me two separate files, your typed answers to the questions asked in the homework and a separate script file. The script file allows me to check that you did the correct analysis. The answer file allows me to see that you can answer the questions in complete and coherent sentences.

The email must include, as its subject line:

`POLS6123: Assignment 1`

Note. *Make sure you include no code in the write-up. All code needs to be attached to the email in the separate script file.*

Problem: High Students [[10]]

The data file `studentHeight` contains height measurements for 10 male students and 10 female students who took a Statistics course in the summer of 2010. Let us assume that the heights of male students and of female students are Normally distributed and that they have the same variance. In symbols, if X_m is the height of male students and X_f is the height of female students, then we are assuming

$$X_m \sim \mathcal{N}(\mu_m, \sigma^2)$$
$$X_f \sim \mathcal{N}(\mu_f, \sigma^2)$$

Part a: First, draw a box plot of the two samples. Comment on whether or not it seems reasonable that the two populations are Normally distributed and have the same variance.

Part b: Now, test the null hypothesis that the average heights of male students are greater than or equal to the heights of female students; that is, test the hypothesis

$$H_0 : \mu_m \geq \mu_f$$

Note. *This problem will be graded based upon how well you determine the correct test to perform, how well you perform that test, and how well you explain your choice.*

Problem: Higher Students? [[15]]

The previous problem made several assumptions about the populations. These assumptions were done to make it easier to determine the appropriate statistical test. This problem removes all of the assumptions in the previous problem.

Let us define μ_m as the average height of male students at Oklahoma State University. Let us also define μ_f as the average height of female students at Oklahoma State University. The data is contained in the `studentHeight` data file.

Test the null hypothesis that the average male student is no shorter than the average female student; that is, test

$$H_0 : \mu_m \geq \mu_f$$

Note. *This problem will be graded based upon how well you determine the correct test to perform, how well you perform that test, and how well you explain your choice.*