# Stat 587: Lab 1

This lab has 3 different activities: two are hands-on, one is computer-based. You can do them in any order

**1: Rocks:**

At the front of the room is a collection of 99 rocks. Your goal is to estimate the total weight of the 99 rocks. This can be done by estimating the average weight of a rock and multiplying by 99.

The rocks are numbered and must stay in order.

a) look at the rocks and choose the one that you believe best represents the average weight of the population. Record that rock number on your data sheet

b) look at the rocks and choose the five whose average you believe best represents the average weight of the population. Record those five rock numbers on your data sheet.

I have the weights for each rock and will summarize the results for class on Friday.

**2: Randomization tests:**

This is a hands-on introduction to material that will be covered in lecture Friday and/or Monday.

This exercise deals with made-up data, so I know the parameters. The data are imagined to arise from a designed experiment, with treatments randomly assigned to experimental units. In this experiment 14 subjects were randomly assigned to either a control (Group A) or an active (Group B) treatment. Here are the data:

Control (Group A): 0.44, 0.33, 1.02, 0.28, 5.12, 1.84, 2.36

Active (Group B): 0.39, 1.83, 5.66, 8.06, 24.93, 2.97, 13.27

The ‘investigators’ are interested in whether the treatment has any effect. They customarily measure treatment effect in two different ways:

1. the difference between the means
2. the ratio of the two means

The sample averages are: Control: 1.63 and Active: 8.16

For the actual experiment, the observed difference is 6.53 (= 8.16 - 1.63) and the observed ratio = 5.01 (= 8.16 / 1.63)

We want to know the probability that this difference and this ratio (or more extreme values) happened by chance.

Concept: If there is no effect of the treatment, then the labels (A or B) do not matter. Hence, the observed sample of 7 observations from Group A are a simple random sample of 7 observations from the universe of 14 possible observations (union of Group A and Group B). The sampling distribution of any desired test statistic under H0 can be estimated by randomization:

What you do:

Tear out each observation.

Combine all 14 observations into one pile.

Draw a simple random sample of 7 observations from the combined set of 14 observations. This sample is taken **without replacement**. Assign these observations to Group A. The remaining 7 observations are Group B.

Record the values of the 7 observations in group A.

I will compute the differences and ratios and share the results.

**3: Introduction to computing**

Goals:

1) Learn more about each of the three supported computer programs (JMP, R, and SAS)

2) Install (or get access to) your preferred program

3) Learn how to read data into the program you want to use

4) Learn how to compute means and standard deviations.

A summary of the three programs is in the “Quick comparison of JMP, R, and SAS” document on the class main page.

The lab materials are organized by program. The introduction to XX documents (for JMP, R, or SAS) describe how to access each program and their basic principles.

For R and SAS, each week I will provide code and a pdf document explaining that code.

For JMP, I will provide a Word document explaining how to obtain the necessary analyses.

Use of each language is demonstrated using the creativity data (a Chapter 1 case study). The data is in creativity.csv.

Please take time to work through the lab material. The lab material usually solves a lot of problems that arise when you start to do the HW.

Most weeks, I will provide a self assessment exercise; answers are included with the exercise. These are for your benefit. No need to turn in these exercises. Some folks like to check their understanding; others just want to start on the HW. Either approach is fine. If doing the self assessments fits your learning style, please do them. If not, then don’t.

Ask if you have any questions.