**Stat 401 A: Lab 8 self-assessment**

Part 1: The data in pvalues.csv are the unadjusted p-values from 4,289 t-tests.

Note: JMP users will not be able to do this part (without a lot of work) because JMP requires the raw data. Sorry.

1. If you consider a test with a p-value < 5% as “significant”, how many tests are significant?

2. If you use a false discovery rate of 5%, how many tests are significant (i.e., are “discoveries”)?

Note: If using SAS and create a data set with the “significant” results (by either criterion), the notes in the log file indicate the number of observations in the subset. If using R, the dim() function returns the dimensions (# rows, # columns) of a data frame and the length() function returns the length of a vector. A statement like sum(pvalue < 0.05) returns the number of times the logical is true (because a logical TRUE is a 1 in a computation and a logical FALSE is a 0).

Part 2: The data in GFCLOCKS.txt are data on the auction price of antique clocks. Consider predicting the price (in British pounds) from the age (in years). You can ignore the other two variables in the data set. Fit the linear regression line.

3. What is the estimated slope?

4. What the units for the slope?

**Answers:**

1. 1406 tests have p values less than 0.05

2.767 tests are discoveries, using a 5% fdr.

3 and 4. 10.5 pounds / year