Empirical Methods for Computer Science (CS 5970) Homework 4

December 2, 2008

This homework assignment is due on Tuesday, December 9nd at 5:00pm. Your work may be handed in electronically (use the **Homework 4** digital dropbox on D2L) or in hardcopy form.

This assignment must be done individually: do not share/discuss your answers with others or look at the answers of others.

Question 1

In this question, we will look at comparing two-sample randomization with bootstrap (with randomization) on the statistic *inter-quartile range* (iqr). Implement a matlab function that shows the power of these two statistical tests as a function of the relative standard deviation of the two samples.

Here are the details:

- $N1 = 20$, $N2 = 8$: number of items in the sample.
- $Q = 100$: number of times that we will take N samples and perform the statistical tests in order to estimate the power of the tests.
- $M = 100$: the number resamples that we will take for the bootstrap tests.

- All tests are right-tailed. IE, we are testing the alternative hypothesis that $iqr(S1) > iqr(S2)$.
- Draw the two samples from randn (at the beginning of your function):
  - $S1 = randn(N1, Q)$
  - $S2_{tmp} = randn(N2, Q)$
- For each power test: $S2 = S2_{tmp}/\sigma$
- $\sigma = [1...12]$
- $\alpha = 0.05$

1. (10pts) Show the power for each of the tests as a function of the assumed standard deviation. Discuss the similarities/differences.

2. (10pts) Setting $N2 = 20$, show the power of the two tests. Discuss the similarities/differences.

3. (10pts) Which test would you prefer to use? Why?