

**CEE 5290/CS 5722/ORIE 5340 Heuristic Methods for Optimization**

**Homework 7: Statistical Comparisons**

**Assigned: Fri, Oct 10th, 2014**

**Due: Fri, Oct 17<sup>th</sup>, 2014 @ noon**

**(free extension to Monday Oct 20<sup>th</sup> noon due to fall break)**

**Readings:** Class handouts on statistical testing. Also refer to the “Probability and Statistics Review” and “Hypothesis Testing MATLAB” documents under the Course Documents folder on Blackboard

---

**Statistical Comparisons:** The table below shows the objective function value for the best solution in each trial for three different algorithms applied to the same problem:

Trial	SA	GA	GS
1	91.94	147.90	47.66
2	77.13	97.88	150.53
3	10.93	39.76	97.04
4	18.6	204.48	82.62
5	28.63	488.83	99.89
6	86.52	113.00	76.52
7	64.58	141.97	87.84
8	22.23	53.76	51.73
9	59.75	408.20	147.51
10	134.11	226.95	115.98
Mean	59.44	192.27	95.73
Std. Dev	39.52	148.35	34.90

(i). Make a boxplot for the data provided above (Use the Matlab command BOXPLOT). Comment on your plot: How do the means and variances compare for each of the algorithms? Are there any outliers? Which algorithm performed the best in your opinion and why?

(ii). Plot empirical CDF's for the data above using the plotting position formula provided in class ( $i/[n+1]$ ). Comment on your plot: Which algorithm appears to perform the best? Is there any evidence of *stochastic dominance*?

(iii). Perform all pairwise comparisons (three in all) of mean objective function value of best solution using a two-sample t-test. State your hypothesis. Report your test statistic and p-values for each comparison. At  $\alpha = 0.05$  what is your conclusion for each test? NOTE: to choose  $H_a$ , test if the means are the same at first, then do hypothesis test to see if one mean is higher than the other.

(iv). If you were told that SA and GS had the same starting solution in each trial would you perform a different test for comparing these two algorithms? Explain why or why not. If you decide to perform another test state your hypothesis and report its p-value and your conclusions at  $\alpha = 0.05$ . Compare your test results to those in (iii).

(v). Perform non-parametric comparisons for all pairwise tests performed in (iii). Report your test statistic and p-values for each comparison. At  $\alpha = 0.05$  what is your conclusion for each test? Compare your test results to those in (iii). NOTE: P-value is the probability of obtaining the test statistic observed or higher. If P-value is lower than your alpha (rejection region), then you reject the null hypothesis.

(vi). Summarize your comparisons by different methods. Are your conclusions different in each case? Based on the tests above which algorithm would you pick as the preferable one? Explain your reasons.