## STAT 621: Simulation Exercise

**<u>Instructions:</u>** Write a simulation program to investigate the performance of one (or more) of the hypothesis tests we discussed in class (Sign, Wilcoxon, Mann-Whitney, Spearman's and Kendall's correltion). You may examine any question you want, but be sure to make it as specific as possible. Some suggestions are below.

- How does the type of distribution that generated the data affect power of the Wilcoxon test? How about the sign test? Do nonsymmetric distributions have a large impact on the performance of the Wilcoxon test?
- How does power behave as the true $\theta$ gets farther away from $\theta_0$, for any of the procedures we've discussed?
- Compare the power of the Wilcoxon signed rank test, the sign test and the usual Normal hypothesis test (t-test or z-test). Are there certain conditions (e.g. sample sizes) when one is more powerful than the others? Do they both achieve their stated significance level?
- Compare the exact version of the Wilcoxon test to the normal approximation. Which is more powerful for small sample sizes?
- Compare power for the Pearson, Spearman and Kendall tests of association. How do these compare when data are normal, or non-normal? How does sample size affect power? Compare power for linear vs. monotonic functions.

**Note:** The goal of your simulation is to measure the performance of a hypothesis test or tests under different conditions. Performance might be measured in terms of power or achieved significance level. Monte Carlo simulation can be used to estimate either one.

1. Remember that Power=$P$(reject $H_0$). This can be estimated as follows using Monte Carlo simulation. First specify a set of conditions including:

    - sample size
    - true value of $\theta$
    - distribution of the data
    - other stuff....

    Simulate a large number $N$ of different data sets satisfying these conditions. Each time test your hypothesis and record the p-value or just whether or not the null is rejected. Then the power of the test under that set of conditions is estimated as

    $$\text{Power} = \frac{\text{no. times reject } H_0}{\text{no. simulated data sets}}$$

    By changing one condition at at time (e.g., repeat the process with a new sample size) you can learn how power depends on various factors.

2. Another important characteristic of a test is the achieved significance level $\alpha$. Remember $\alpha = P(\text{reject } H_0 | H_0 \text{ true})$. Certain violations of the assumptions (e.g. non-independent data or non-symmetric data) might lead the true significance level to differ quite a bit from what is advertized. Estimating the achieved $\alpha$ level is done basically the same way as described above. But here data are simulated under the null that $\theta = 0$ (but maybe they are from nonsymmetric distributions for example). Then $\alpha$ is estimated as above.

$$\alpha = \frac{\text{no. times reject } H_0}{\text{no. simulated data sets}}$$

Refer to the R code in the class examples to help get you started.