STAT 672, Midterm Exam Part II, Spring 2011

Due 4:00PM, Thursday March 24

This assignment is a take-home midterm exam. You **may not** collaborate with *any* other person (whether in the class or not). You **may** use any reading material (class notes, books, etc.) you wish. Professor Bailey will answer questions. Please include commands and output you used to answer the questions.

3 Problems. 50 points total.

(16pts) 1. Let X have a Cauchy distribution and let Y have a standard normal distribution. I have simulated a random sample from each of these distributions to use as two datasets. I have used the R function `rcauchy(100)` and `rnorm(100)`, in case you would like more information. You can find these two data sets in a file off the course webpage and you can use the read.table command:

```
testdat <- read.table("http://www.rohan.sdsu.edu/~babailey/stat672/testdat.txt",
header=T)
```

Note: The names of the variables in the file are x and y.

(a) First consider the random sample from the Cauchy distribution ($X's$). Use the Wilcoxon Sign Rank test to test if the population median is equal to zero against the two-sided alternative. What do you conclude at the $\alpha = 0.10$ significance level? Are the assumptions of the Wilcoxon Sign Rank test satisfied? Explain.

(b) You should have two random samples of size $n = m = 100$ from these distributions. Use the Wilcoxon Rank Sum procedure to test $H_0 : \Delta = 0$ against the two-sided alternative. What do you conclude at the $\alpha = 0.10$ significance level? Are the assumptions of the Wilcoxon Rank Sum test satisfied? Explain.

Turn over for Problems 2 and 3:

(16pts) 2. Consider the two-sample location problem. This problem is general, but you may find it useful to use some data for illustration. For illustration, you can use the data in Table 4.3 that is described in Problem 4.1, for an example dataset.

The data is available off the class web page at:

```
http://www.rohan.sdsu.edu/~babailey/stat672/t4-3.txt
```

(a) If the smallest observation for Y is made arbitrarily small, will the Wilcoxon Rank Sum test still reject? Explain.

(b) If the largest observation for Y is made arbitrarily large, will the Wilcoxon Rank Sum test still reject? Explain.

(c) Answer (a) and (b) for the t-test.

(18pts) 3. Let $X_1, X_2, \ldots, X_6$ be random sample of size $n = 6$ from a continuous distribution. Suppose we question whether they are observations of a random sample, in fact we suspect a parabolic trend. Let $R_i = \text{rank}(X_i)$ and take $a_1 = a_6 = 9, a_2 = a_5 = 4, a_3 = a_4 = 1$. A statistic that could be used to test the alternative (parabolic trend) hypothesis is

$$L = \sum_{i=1}^{6} a_i R_i.$$

(a) Under the assumption (H_0) that the n random variables are actually observations from a random sample from a continuous distribution, find $\mu_L = E(L)$ and $\sigma_L^2 = Var(L)$.

(b) Consider the following sequence of observations,

```
> x
 [1] 1.8 1.0 1.5 1.4 1.9 1.6
```

The rejection region is of the form $L \geq c$. Use the normal approximation to calculate the approximate p-value for the test. What is your conclusion at the $\alpha = 0.05$ significance level?