Statistics 428: Homework 3, Spring 2017

(Due February 26, 2017)

Problem 1: (Two-Sample T-test vs Wilcoxon Ranked Sum Test)

Suppose we have two independent samples of equal size $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_n$, with cumulative distribution functions F_x and F_y, respectively. We wish to test the null hypothesis $H_0 : F_x = F_y$ versus an alternative hypothesis of **stochastic dominance** $H_a : (1 - F_y(t)) < (1 - F_x(t))$ for all t for which F_x and F_y are nonzero.

Of course, if $(1 - F_x) > (1 - F_y)$ then the corresponding means have the relationship $\mu_x > \mu_y$. With this in mind, one reasonable test statistic is the familiar two-sample t-statistic

$$t = \frac{\bar{y} - \bar{x}}{\sqrt{s_p^2/n + s_p^2/n}}$$

where s_p^2 is the usual pooled variance estimator under the assumption of a common variance $\sigma^2 = \sigma_x^2 = \sigma_y^2$. When $F_x = F_y$ and are the cdfs of a normal distribution, t has a t-distribution with $2n - 2$ degrees of freedom.

An alternative is the Wilcoxon ranked sum test. It is constructed as follows in the special case of equal sample sizes n in the x and y samples:

1. In the pooled sample $\{x_1, ..., x_n, y_1, ..., y_n\}$ replace each observation with its rank, ranking from smallest to largest, $R = \{r_1, r_2, ..., r_n, r_{n+1}, r_{n+2}, ..., r_{2n}\}$.
2. Let $W = \sum_{i=1}^{n} r_i$, the sum of the ranks corresponding to the x-sample.
3. Find the p-value for the one-sided alternative $H_a : (1 - F_y) < (1 - F_x)$, which is $p = P[W \geq w]$, where w is the observed value. Note that when $(1 - F_y) < (1 - F_x)$ we expect larger observations to come from X than from Y, meaning that W will be larger than expected under the null hypothesis. It is possible to find an exact p-value under the null hypothesis, without knowing the exact form of $F = F_x = F_y$. This is due to replacing the observations with their ranks. When $F_x = F_y$ any permutation of $\{1, ..., n, n + 1, ..., 2n\}$ is an equally likely realization of the rank vector R.

However, when n is large it can be fast and accurate to use the asymptotic distribution of the statistic. Specifically, for large n, when the null hypothesis is true

$$z = \frac{W - \mu_w}{\sigma_w} \sim N(0, 1)$$

where $\mu_w = n^2 + n/2$ and $\sigma_w = \sqrt{\frac{2n^3+n^2}{12}}$. We can then refer to the normal distribution to obtain the p-value.

The problem is to do a full power and type 1 error rate analysis of the two methods, setting significance level $\alpha = 0.05$. Try to carefully describe your simulation conditions for each case below, and provide a bound for the standard error of your type 1 error and power estimates.

a. Do a Monte Carlo study of power and type 1 error rate when $X \sim N(0,1)$ and $Y \sim N(-\Delta, 1)$ where $\Delta \geq 0$. Note that $\Delta = 0$ satisfies the null hypothesis. The t reference distribution for the t-statistic will be exact, but you should conduct some independent simulations to see how large n should be before replacing the exact p-value for W with the p-value from the normal approximation. Show appropriate tables and figures to study how power depends on n and Δ.

b. Repeat the study, but this time let $f_x(x) = e^{-x}$ for $x > 0$, and let $f_y(y) = \lambda e^{-\lambda y}$ for $y > 0$ and $\lambda \geq 1$. The null condition in this case is $\lambda = 1$. Again perform a Monte Carlo analysis of how power depends on n and λ, and how type 1 error rate depends on n. Note that the t-distribution for t will not be exactly correct in this case.

c. This time let $X \sim \chi^2(1)$ and $Y \sim \chi^2(\nu)$ for $0 < \nu \leq 1$. Perform a Monte Carlo analysis of power and type 1 error rate for t and W.

Provide a summary discussion of your conclusions, comparing t and W.

For problem 1 you may use the functions t.test() and wilcox.test(). Read the help files for these functions carefully. For instance, you should look for how to specify the equal variance assumption in t.test(), and in wilcox.test() you can see how to specify whether you want an exact p-value or the asymptotic approximation.

Problem 2: (Comparison of Confidence Intervals)

Let $x_1, x_2, ..., x_n$ be a random sample from F_x and consider the problem of obtaining a confidence interval of level $1-\alpha$ for the median $\theta = F_x^{-1}(0.5)$. Let $\hat{\theta}$ denote the sample median. We discussed two ways of constructing confidence intervals using the bootstrap. One method for statistics $\hat{\theta}$ that are asymptotically normal is to form the interval

$$(\hat{\theta} - z_{\alpha/2} se(\hat{\theta}), \hat{\theta} + z_{\alpha/2} se(\hat{\theta})) \tag{1}$$

where $se(\hat{\theta})$ is the bootstrap estimate of the standard error.

The second method used the ordered bootstrap replicates,

$$\hat{\theta}^*_{(1)} \leq \hat{\theta}^*_{(2)} \leq ... \leq \hat{\theta}^*_{(B)},$$

to form the interval

$$(\hat{\theta}^*_{(B\alpha/2)}, \hat{\theta}^*_{(B(1-\alpha/2))}) \tag{2}$$

Conduct a Monte Carlo study comparing the coverage probability and mean lengths of intervals constructed by the two methods. Let $\alpha = 0.10$ and study sample sizes of $n = 20, 50, 100$ with data coming from distributions $N(0, 1)$, exponential($\lambda = 1$), Cauchy, and Uniform(0,1). Show that your standard error for the coverage probability estimates is under 0.005. Give a summary of your conclusions and which method you would recommend.