The Mann-Whitney test and confidence interval

For the two sample situation (at first for data with no ties), let $X_1, X_2, ..., X_m$ be the data from treatment 1 and let $Y_1, Y_2, ..., Y_n$ be the data from treatment 2. You can picture the data as being arranged in a m x n array where the X_i values are the rows and the Y_j values are the columns. For the m x n pairs of (X_i, Y_j) data in this array, the Mann-Whitney test statistic U is defined as the number of (X_i, Y_j) pairs where $X_i < Y_j$. The null hypothesis for this test is the same as for the Wilcoxon test, that the distributions of X_i and Y_j are the same. Large values of U indicate that treatment 2 has larger observations. Although this test was developed independently of the Wilcoxon test and has its own set of tables, it turns out that this test is exactly equivalent to the Wilcoxon test. To see this, consider ranking together all of the X_i, Y_j data and looking at the rank of a Y_j value:

$$R(Y_j) = (\text{number of } Y's \leq Y_j) + (\text{number of } X's \leq Y_j).$$

If we add these terms over all Y_j values, the term on the left-hand side is just the Wilcoxon rank-sum statistic based on sample 2 (W_2), the first term on the right-hand side is just the sum of the ranks from 1 up to n (equal to $n(n+1)/2$), and the second term on the right-hand side is the Mann-Whitney test statistic, U. Thus, we have that $W_2 = n(n+1)/2 + U$, so that the two tests are within a constant term of each other and are therefore equivalent.

The most useful aspect of the Mann-Whitney test is that we can use the idea of the (X_i, Y_j) pairs to develop a confidence interval for the shift parameter that separates two distributions. That is, if $F_1(x) = F_2(x - \Delta)$, then the parameter Δ is a shift, like the difference in medians, between the two distributions. We consider the set of mn pairwise differences $(X_i - Y_j)$ and find order statistics of these differences to define our confidence interval for Δ, just as we did for the confidence interval for the median $\theta_{.5}$. Using the abbreviation pwd to stand for pairwise differences, the text shows how the distribution of the pwd terms is related to the Mann-Whitney U distribution:

$$P(\text{pwd}(k_a) < \Delta \leq \text{pwd}(k_b)) = P(k_a \leq U \leq k_b - 1).$$

The text then illustrates the use of a table of U percentiles to construct a confidence interval for Δ based on the pwd values. Since we will use the large-sample normal approximation at the end of the chapter, we will wait until that section to construct these confidence intervals.

The Hodges-Lehmann estimate of Δ

From the Mann-Whitney approach based on pairwise differences we also get a estimate of the shift parameter Δ. The Hodges-Lehmann estimate of Δ is the median of the pairwise differences $X_i - Y_j$.

Note: for tied values of X_i and Y_j, we count .5 toward the U statistic value. This rule leads to a Mann-Whitney test that is equivalent to the Wilcoxon rank-sum test adjusted for ties.