From Theorem A, we have

COROLLARY A

Under the null hypothesis H_0: $F = G$,

$$E(U_Y) = \frac{mn}{2}$$

$$\mathrm{Var}(U_Y) = \frac{mn(m+n+1)}{12} \qquad \blacksquare$$

For m and n both greater than 10, the null distribution of U_Y is quite well approximated by a normal distribution,

$$\frac{U_Y - E(U_Y)}{\sqrt{\mathrm{Var}(U_Y)}} \sim N(0, 1)$$

(Note that this does not follow immediately from the ordinary central limit theorem; although U_Y is a sum of random variables, they are not independent.) Similarly, the distribution of the rank sum of the X's or Y's may be approximated by a normal distribution, since these rank sums differ from U_Y only by constants.

EXAMPLE B Referring to Example A, let us use a normal approximation to the distribution of the rank sum from method B. For $n = 13$ and $m = 8$, we have from Corollary A that under the null hypothesis,

$$E(T) = \frac{8(8 + 13 + 1)}{2} = 88$$

$$\sigma_T = \sqrt{\frac{8 \times 13(8 + 13 + 1)}{12}} = 13.8$$

T is the sum of the ranks from method B, or 51, and the normalized test statistic is

$$\frac{T - E(T)}{\sigma_T} = -2.68$$

From the tables of the normal distribution, this corresponds to a p-value of .007 for a two-sided test, so the null hypothesis is rejected at level $\alpha = .01$, just as it was when we used the exact distribution. For this set of data, we have seen that the t test with the assumption of equal variances, the t test without that assumption, the exact Mann-Whitney test, and the approximate Mann-Whitney test all reject at level $\alpha = .01$. $\blacksquare$
