Tests for equality of scale

We will examine two-sample tests for equality of scale for two conditions: i) assuming equal location parameters (the Siegel-Tukey and Ansari-Bradley tests), and ii) without assuming equal location parameters (ratio of mean deviances tests).

i) A test for equal scale, given equal location parameters (the Siegel-Tukey and Ansari-Bradley tests)

Since the two groups have equal location parameters, we can use a clever ranking system to allow us to use the Wilcoxon statistic. The Siegel-Tukey test is conducted by combining data from both groups, and ranking in a different way. In the Siegel-Tukey test, the smallest observation has rank 1, then the largest observation has rank 2, the next largest has rank 3, the second smallest has rank 4, the third smallest has rank 5, etc. Then a Wilcoxon rank-sum test is applied.

If you had started the Siegel-Tukey test with the largest observation getting rank 1, the smallest getting rank 2, the second smallest getting rank 3, etc, you are doing essentially the same test but you will typically get a different value for the Wilcoxon statistic. To address this non-uniqueness, the Ansari-Bradley test performs the ranking both ways and averages the ranks. The tables for the Wilcoxon test are no longer applicable, so a permutation test can be conducted.

ii) A test for equal scale, with possibly unequal location parameters (RMD and $\widehat{RMD}$)

If location parameters differ, then the Siegel-Tukey and Ansari-Bradley tests will not be appropriate. Instead, we base tests on deviances, such as

$$\text{dev}_{ix} = X_i - \mu_1, \quad \text{dev}_{iy} = Y_j - \mu_2, (\text{if } \mu_i\text{'s are known})$$

or otherwise

$$\widehat{\text{dev}}_{ix} = X_i - \text{med}_1, \quad \widehat{\text{dev}}_{iy} = Y_j - \text{med}_2, (\text{if } \mu_i\text{'s are unknown}).$$

In the first case, using dev_{ix} and dev_{iy}, compute the statistic:

$$RMD = \frac{\sum_{i=1}^{m} |\text{dev}_{ix}|/m}{\sum_{j=1}^{n} |\text{dev}_{iy}|/n},$$

and obtain a permutation distribution by permuting the dev_{ix} and dev_{iy} terms. If the location parameters are unknown, then compute the permutation distribution of

$$\widehat{RMD} = \frac{\sum_{i=1}^{m} |\widehat{\text{dev}}_{ix}|/m}{\sum_{j=1}^{n} |\widehat{\text{dev}}_{iy}|/n},$$

by permuting the $\widehat{\text{dev}}_{ix}$ and $\widehat{\text{dev}}_{iy}$ values.

The Type I error rate of the $\widehat{RMD}$ -based procedure

As discussed in the text, the Type I error rate of the $\widehat{RMD}$ procedure does vary with the population distribution of the data, however this effect lessens as sample size increases.