Wilcoxon Mann Whitney Test

In the independent samples case, we first learned to apply the independent sample t test when we had normality of the sample mean for each sample. If there is a violation of this assumption, we'll apply a non parametric test for independent samples – the Wilcoxon Mann Whitney test.

This test uses the relative position of the data in a rank ordering, unlike the t test which uses the actual values. To use the Wilcoxon Mann Whitney test

- It must be reasonable to regard the data as a random sample from their respective populations
- Observations within each sample must be independent of one another
- The two samples must be independent of one another

Rank Sum Algorithm

(1) Arrange (rank) the observations in each sample from smallest to largest
(2) Count the number of observations in each group that are smaller than each observation in the other group.
   a. Look at the first observation in the first group and count the number of observations in the second group that are smaller than it.
   b. Record that count.
   c. Repeat for all observations in the sample.
   d. Do a, b, and c for the second column.
(3) Add the rankings you recorded from Step (2) for the first column and call this sum K_1. Add the rankings you recorded from Step (2) for the second column. Call this sum K_2. Check your calculations! $K_1 + K_2 = n_1 n_2$
(4) The test statistic is $U_s = \max\{K_1, K_2\}$
(5) Use the test statistic, U_s to bracket the P-value from Table 6

Example 7.39 Soil respiration is a measure of microbial activity in soil, which affects plant growth. In one study, soil cores were taken from two locations in a forest – under an opening in the forest canopy ("gap" location) and a nearby location under heavy tree growth ("growth" location). The amount of carbon dioxide given off by each core was measured (mol CO_2/g soil/hr).

Normal Q-Q Plot (Growth) and **Normal Q-Q Plot** (Gap)

TABLE 7.17 Soil Respiration Data (mol CO_2/g soil/hr) from Example 7.38

Growth				Gap			
17	20	170	315	22	29	13	16
22	190	64		15	18	14	6
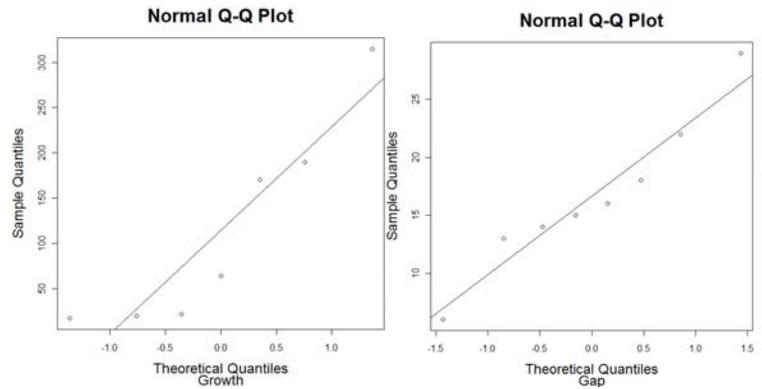
The researcher would like to compare soil respiration under the two conditions (gap and growth), but there is a clear violation of normality, and with small sample sizes the CLT can't help us out. An attempt to transform the growth data (maybe with log(growth)), would make the skewness of the gap distribution even worse.

Employ the Wilcoxon Mann Whitney test at the α = 0.05 significance level to see whether the distributions are shifted in some way from one another.

TABLE 7.18 Wilcoxon-Mann-Whitney Calculations for Example 7.39

Number of Gap Observations That Are Smaller	Y_1 Growth Data	Y_2 Gap Data	Number of Growth Observations That Are Smaller
	17	6	
	20	13	
	22	14	
	64	15	
	170	16	
	190	18	
	315	22	
		29	
$K_1 =$			$K_2 =$