6.1 Wilcoxon Rank-Sum Test

Suppose observations $x_1, \dots, x_{n_1}$ and $y_1, \dots, y_{n_2}$ are available. It is assumed that the two samples are drawn from independent populations, and that the underlying distribution of the measurements is unknown. To test a hypothesis regarding the relative position of the location parameters (typically thought of a median rather than a mean) of the underlying distributions, the **Wilcoxon rank-sum test** may be conducted. The location parameters are denoted θ_X and θ_Y, respectively.

The following are the steps of the testing procedure. First, the samples are pooled and the observations are ranked so that the smallest observation receives the rank of 1. If two or more observations are tied, all of them are assigned the same rank which equals the average of the ranks that the observations should have gotten if they were not tied. Then the test statistic W is computed as the sum of the ranks assigned to the observations in the smaller sample. For definiteness, we will assume that the first sample, the sample of size n_1, is smaller. This can always be achieved by renaming samples. If the samples have equal sizes, we take the first sample.

The next step depends on the type of the alternative hypothesis under investigation.

• $H_1 : \theta_X < \theta_Y$ is one-sided asserting that the distribution of Y is shifted to the right with respect to the distribution of X. Under this alternative hypothesis, the test statistic is expected to be small, thus the null hypothesis should be rejected for small values of W, that is, if $W \leq W_L$ where W_L denotes the lower critical value introduced below.

• $H_1 : \theta_X > \theta_Y$ is one-sided stating that the distribution of Y is shifted to the left with respect to the distribution of X. If this alternative is true, then the test statistic tends to be large. Hence, the null hypothesis should be rejected for large values of W, that is, if $W \geq W_U$ where W_U is the upper critical value defined below.

• $H_1 : \theta_X \neq \theta_Y$ is two-sided with the direction of the location shift not specified. Under this alternative, the test statistic should be either small or large, therefore, the null hypothesis is rejected if W falls outside of the critical interval (W_L, W_U).

The lower and upper critical values W_L and W_U depend on the sample sizes n_1 and n_2, significance level α (typically chosen as 0.01 or 0.05), and whether the alternative is one- or two-tailed. The critical values are tabulated (see

Table 2 on page 139 in "Nonparametric Methods in Statistics with SAS Applications" By O.Korosteleva, CRC Press, 2013).

Example. Suppose we would like to compare the location parameters for the distributions of BMI in Communities A and B. We assume that the true distribution of BMI is unknown, and therefore, we resort to a nonparametric Wilcoxon rank-sum test. We pool the data and assign ranks. The result is presented in the following table.

Community A	29.3	31.4	38.7	33.2	30.3
Rank A	1	3	5	4	2
Community B	42.0	39.9	44.5	40.7	38.9
Rank B	9	7	10	8	6

The test statistic is the sum of ranks for Community A. We compute $W = 1+3+5+4+2 = 15$. Suppose we want to test at the 1% significance level $H_0 : \theta_A = \theta_B$ against a one-sided alternative $H_1 : \theta_A < \theta_B$. We look up the lower critical value for $n_1 = n_2 = 5, \alpha = 0.01$, for one-sided alternative. We get $W_L = 16$. Since our test statistic $W = 15$ is smaller than W_L, we reject the null hypothesis in favor of the alternative and conclude that the location parameter of the BMI distribution in Community A is lower than that for Community B.

In SAS:

```
data cohortsAB;
 input community $ BMI @@;
cards;
A 29.3 A 31.4 A 38.7 A 33.2 A 30.3
B 42.0 B 39.9 B 44.5 B 40.7 B 38.9
;

proc npar1way data=cohortsAB Wilcoxon;
   class community;
     var BMI;
   exact;
run;
```

Wilcoxon Two-Sample Test							
				t Approximation		Exact	
Statistic (S)	Z	Pr < Z	Pr > |Z|	Pr < Z	Pr > |Z|	Pr <= S	Pr >= |S-Mean|
15.0000	-2.5067	0.0061	0.0122	0.0167	0.0335	0.0040	0.0079
Z includes a continuity correction of 0.5.							

In R:

```
library(exactRankTests)
BMI.A<- c(29.3, 31.4, 38.7, 33.2, 30.3)
BMI.B<- c(42.0, 39.9, 44.5, 40.7, 38.9)
wilcox.exact(BMI.A, BMI.B, paired=FALSE, alternative="less")
#The test statistic is W-n1(n1+1)/2

Exact Wilcoxon rank sum test

data:  BMI_A and BMI_B
W = 0, p-value = 0.003968
alternative hypothesis: true mu is less than 0
```

Note that the test statistic that R outputs is equal to $15 - 5(5 + 1)/2 = 15 - 15 = 0$. □