Three Nonparametric Tests

We have been using z- and t-statistics for testing hypotheses about a population mean or for comparing two population means. The z-statistic is used for large random samples selected from populations with few limitations on the probability distribution of the underlying population. The t-statistic is used for small samples selected at random from normally distributed populations. What do we do when working with a small sample from a nonnormal population? In Section 6.6 we considered a distribution-free method called the sign test that required few assumptions about the underlying population. The sign test is used to test hypotheses about a population median rather than a population mean. We will now consider three nonparametric tests for comparing two populations. In each case we will assume the probability distributions from which the samples are collected are continuous.

Two Nonparametric Tests for Comparing Two Populations: Independent Sampling

Example 1: We wish to compare reaction times for adult males under the influence of Drug A with reaction times for those under the influence of Drug B. Populations of reaction-time measurements are known to be frequently skewed to the right. So a t-test should not be used to compare the reaction times of the two drugs.

Suppose subjects were randomly assigned to each of two groups, one group to receive Drug A and the other Drug B. The reaction time for each subject was measured at the completion of the experiment. The reaction times in seconds are shown in the table below.

↓	C1	C2
	Drug A	Drug B
1	1.96	2.11
2	2.24	2.43
3	1.71	2.07
4	2.41	2.71
5	1.62	2.50
6	1.93	2.84
7		2.88
8		

The population of reaction times for Drug A is that which could conceptually be obtained by giving Drug A to all adult males. The population of reaction times for Drug B is defined in like manner. The tests we consider here are designed to compare the two probability distributions. We denote the two probability distributions by D_1 and D_2 respectively.

We will compare the two populations using the Wilcoxon Rank Sum Test for Independent Samples and again using an equivalent test called the Mann-Whitney Test.

↓	C1	C2	C3	C4
	Drug A		Drug B	
1	1.96	4	2.11	6
2	2.24	7	2.43	9
3	1.71	2	2.07	5
4	2.41	8	2.71	11
5	1.62	1	2.50	10
6	1.93	3	2.84	12
7			2.88	13
8				

We rank the sample observations as though they were drawn from the same population. We denote the rank sums for Drug A and Drug B by T_1 and T_2 respectively. $T_1 = 25$ and $T_2 = 66$. In this example $n_1 = 6$ and $n_2 = 7$.

Wilcoxon Rank Sum Test (Example 7.8 in text.)

Ho: $D_1 = D_2$ (The distributions are identical.)
Ha: $D_1 \neq D_2$ (D_1 is shifted to the left or right of D_2.)
Test statistic: $T_1 = 25$ because $n_1 < n_2$.

Rejection region: $T \leq T_L = 28$ or $T \geq T_U = 56$. (Table V for $\alpha = 0.05$)
There is sufficient evidence to reject *Ho* because T_1 is in the rejection region. So we can conclude, at the $\alpha = 0.05$ level that the probability distributions for drugs A and B are not identical.

Mann-Whitney Test (Equivalent to Wilcoxon Rank Sum Test. See pp. 391-392 in text.) MINITAB output is shown below.

Mann-Whitney Test and CI: Drug A, Drug B

```
        N  Median
Drug A  6  1.9450
Drug B  7  2.5000


Point estimate for ETA1-ETA2 is -0.4950
96.2 Percent CI for ETA1-ETA2 is (-0.9497,-0.1099)
W = 25.0
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0184
```

A Nonparametric Test for Comparing Two Populations: Paired Differences

Example 2: In this example we consider a situation where consumer preferences for two competing products are compared by having a sample of consumers rate both products. The ratings are paired for each consumer. Suppose each of 10 judges is given a sample of two paper products for comparison on the basis of softness. Each judge rates the softness of each product on a scale from 1 to 10, with higher ratings implying a softer product.

Data for this experiment is shown below. We will employ the ***Wilcoxon Signed Rank Test for Paired Differences***. (See pp. 395-397 in text.)

	A	B	C	D	E	F
1	Judge	Product A	Product B	A - B	lA - Bl	Rank lA - Bl
2	1	6	4	2	2	5
3	2	8	5	3	3	7.5
4	3	4	5	-1	1	2
5	4	9	8	1	1	2
6	5	4	1	3	3	7.5
7	6	7	9	-2	2	5
8	7	6	2	4	4	9
9	8	5	3	2	2	5
10	9	6	7	-1	1	2
11	10	8	2	6	6	10
12						

T_+ = sum of positive ranks = 46

T_- = sum of negative ranks = 9

Ho: The probability distributions of the ratings for products A and B are the same.

Ha: The probability distributions of the ratings differ in location for the two products.

Test statistic: T = smaller of T_+ and T_- = 9

Rejection region: $T \leq 8$ for $\alpha = 0.05$. (Table VI)

There is insufficient evidence to reject *Ho*. That is, we cannot conclude that the two products differ with respect to their softness ratings at the $\alpha = 0.05$ level.

Exercise 1. Exercise 7.67 Use both *Wilcoxon Rank Sum Test* and *Mann-Whitney Test.*

Exercise 2. Exercise 7.81