AM 166 Problem Set 4 2/28/03

15.17 The hypothesis to be tested is

H_0: the population distributions for plastics 1 and 2 are the same

H_a: the population distributions differ in location

We rank the 12 observations in order of magnitude across groups. The data, with corresponding ranks, are shown in the table at the right.

Plastic 1	Plastic 2
15.3 (2)	21.2 (9)
18.7 (6)	22.4 (11)
22.3 (10)	18.3 (5)
17.6 (4)	19.3 (8)
19.1 (7)	17.1 (3)
14.8 (1)	27.7 (12)

The two possible values for U are

$$U_A = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_A = 36 + \frac{6(7)}{2} - W_A = 57 - 30 = 27$$

$$U_B = n_1 n_2 + \frac{n_2(n_2+1)}{2} - W_B = 36 + \frac{6(7)}{2} - 48 = 9$$

Since we have agreed to use the smaller value of U as a test statistic, a lower-tailed rejection region must be determined so that α is close to .10. Notice that the hypothesis to be tested is actually two-tailed. That is, both large and small values of U will tend to contradict the null hypothesis. Hence, although we will only consider the area below some critical value of U (denoted by U_0) in determining α for the test, there is a similar area in the upper tail of the distribution. Thus the area below the critical U is actually $\frac{\alpha}{2}$ and must be doubled to obtain the value for α.

Referring to Table 8 and indexing $n_1 = n_2 = 6$, a rejection region is determined such that

$$P(U \leq U_0) = \frac{\alpha}{2} = .05$$

Hence we will reject if $U \leq 7$, with $\frac{\alpha}{2} = .0465$ or $\alpha = .093$. The minimum of U_1 and U_2 is $U = 9$, and H_0 is not rejected. We cannot conclude that the populations differ in location.

15.18a. The data with their corresponding ranks are given in the table at the right. Notice that the observations corresponding to ranks 9 and 10 and 13 and 14 were tied. Hence an average rank was assigned to both (9.5 and 13.5, respectively).

	A	*B*
	6.1 (1)	9.1 (16)
	9.2 (17)	8.2 (8)
	8.7 (12)	8.6 (11)
	8.9 (13.5)	6.9 (2)
	7.6 (5)	7.5 (4)
	7.1 (3)	7.9 (7)
	9.5 (18)	8.3 (9.5)
	8.3 (9.5)	7.8 (6)
	9.0 (15)	8.9 (13.5)
	14.8 (1)	27.7 (12)
Rank Sum	94	77

Then the two possible values for U are

$U_A = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_A = 126 - 94 = 32$

$U_B = n_1 n_2 + \frac{n_2(n_2+1)}{2} - W_B = 126 - 77 = 49$

Thus, $\mu = 32$. Referring to Table 8 and indexing $n_1 = n_2 = 9$, the p-value is $2P(U \le 32) = 2(.2447) = .4894$.

b. Since the data are presented in an unpaired manner (that is, the two samples are independent and random), the procedure is identical to that used in Section 10.6 of the text. The analysis is as follows:

(1) H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$

(2) $$s^2 = \frac{\sum_j y_{1j}^2 - \left[\frac{\left(\sum_j y_{1j}\right)^2}{n_1}\right] + \sum_j y_{2j}^2 - \left[\frac{\left(\sum_j y_{2j}\right)^2}{n_2}\right]}{n_1+n_2-2}$$

$$= \frac{625.06 - \left[\frac{(74.4)^2}{9}\right] + 599.22 - \left[\frac{(73.2)^2}{9}\right]}{16} = .8675$$

(3) Test statistic: $t = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.1333}{\sqrt{.8675\left(\frac{1}{9}+\frac{1}{9}\right)}} = .30$

(4) Since $|.3| < 1.337$, the p-value $> 2(.1) = .2$.

15.19 If the alternative hypothesis is true, we would expect the batteries from plant A to fail later than the batteries from plant B. Hence the observations from plant A will be ranked near the end of the sequence and the U statistic (i.e., the number of observations from plant A that precede each observation from plant B) will be small. Hence, small values of U will tend to contradict the null hypothesis, and a lower-tailed rejection region is desired. Remember that the test statistic U has expected value $E(U) = \frac{n_1 n_2}{2}$ and variance $V(U) = \frac{n_1 n_2(n_2+n_2+1)}{12}$ when the null hypothesis is true. Also, for large values of n_1 and n_2, the quantity

$$Z = \frac{U - E(U)}{\sigma_\mu}$$

will be approximately normal with mean 0 and variance 1. Once this z value has been calculated, the null hypothesis will be rejected if $z < -1.645$ (cf. Chapter 10). The following data are available: $n_1 = n_2 = 15$, $W_A = 276$, and $W_B = 189$. Notice that observations are presented in order of failure, so that

$W_A = 1 + 5 + 7 + 8 + 13 + 15 + 20 + 21 + 23 + 24 + 25 + 27 + 28 + 29 + 30 = 276$

The value of U is

$$U_A = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_A = 345 - 276 = 69$$

Also,

$$E(U) = \frac{n_1 n_2}{2} = \frac{225}{2} = 112.5$$

and

$$V(U) = \frac{n_1 n_2(n_1+n_2+1)}{12} = \frac{1}{12}(15)(15)(31) = 581.25$$

Thus,

$$z = \frac{U - E(U)}{\sigma_\mu} = \frac{69 - 112.5}{\sqrt{581.25}} = \frac{-43.5}{24.1} = -1.80$$

The null hypothesis is rejected since $z = -1.80$ falls in the rejection region.

15.20 Although $n_2 = 3$, we use a large-sample approximation (Table 8 stops at 10, and $n_1 = 12$). The data with corresponding ranks are shown in the table at the right. Then

DDT	Diazinon
16 (13)	7.8 (9)
5 (6.5)	1.6 (2)
21 (15)	1.3 (1)
19 (14)	
10 (12)	
5 (6.5)	
8 (10)	
2 (3.5)	
7 (8)	
2 (3.5)	
4 (5)	
9 (11)	

$$U_A = 12(3) + \frac{12(13)}{2} - 108 = 6$$

$$U_B = 12(3) + \frac{3(4)}{2} - 12 = 30$$

$$E(U) = \frac{12(3)}{2} = 18$$

and

$$V(U) = \frac{12(3)(16)}{12} = 48$$

The test statistic is

$$z = \frac{U - E(U)}{\sqrt{V(U)}} = \frac{6-18}{\sqrt{48}} = -1.732$$

The p-value is $P(|z| > 1.732) = 2(.0418) = .0836$. Since $.0836 < .10$, the null hypothesis is rejected. There is a difference in the locations of the two populations.

15.51 **a.** Since we are interested in a difference in recovery rates, let

$p = P(\text{recovery rate for } A \text{ exceeds } B \text{ at a given hospital})$

M = number of times A exceeds B

The hypothesis to be tested is

$$H_0\colon p = \tfrac{1}{2} \quad \text{vs.} \quad H_a\colon p \neq \tfrac{1}{2},$$

and the data are shown in the following table.

Hospital	A	B	Sign of $(A - B)$
1	75.0	85.4	−
2	69.8	83.1	−
3	85.7	80.2	+
4	74.0	74.5	−
5	69.0	70.0	−
6	83.3	81.5	+
7	68.9	75.4	−
8	77.8	79.2	−
9	72.2	85.4	−
10	77.4	80.4	−

Various rejection regions are tried in order to find $\alpha = .10$. (Use Table 1, Appendix III).

Rejection Region	α
$M = 0, M = 10$	.002
$M \leq 1, M \geq 9$	.022
$M \leq 2, M \geq 8$	.110

Using the rejection region $M \leq 2$ or $M \geq 8$, the null hypothesis is rejected since the observed value of M is $m = 2$. We conclude that a difference does exist in the recovery rates for the two drugs.

b. In the above analysis we made no assumptions concerning the underlying distributions of the data. To use the t test, we must be able to assume normality of the distributions and equal variances for the two populations. Since the observations given above are percentages, their distributions may be almost mound-shaped, but the variances will not be equal.

15.52 Let $p = P$(gourmet A's rating exceeds gourmet B's for a given meal) and $M =$ a number of meals for which gourmet A exceeds B. The hypothesis to be tested is

$$H_0\colon\ p = \tfrac{1}{2} \qquad \text{vs.} \qquad H_a\colon\ p \neq \tfrac{1}{2}.$$

The sign test will be used with M as the test statistic. Notice that for this exercise, $n = 17$, since a tied rating was given to meals 7, 14, and 20. The value of the test statistic is $m = 8$. Various two-tailed rejection regions are tried in order to find a region with $\alpha = .05$. The calculations are shown below.

Rejection Region	$\alpha = P\left(\text{reject } H_0 \mid p = \tfrac{1}{2}\right)$
$M \le 2, M \ge 15$	$2\left[\binom{17}{0} + \binom{17}{1} + \binom{17}{2}\right]\left(\tfrac{1}{2}\right)^{17} = .00235$
$M \le 4, M \ge 13$	$2\left[\binom{17}{0} + \binom{17}{1} + \binom{17}{2} + \binom{17}{3} + \binom{17}{4}\right]\left(\tfrac{1}{2}\right)^{17} = .04904$

Since the second region gives $\alpha = .05$, we choose to reject H_0 if $M \le 4$ or $M \ge 13$. Since the observed value of M is $m = 8$, the null hypothesis is not rejected. There is insufficient evidence to indicate a difference between the two gourmets.

15.53 For the Wilcoxon signed-rank test, the differences and the ranks of their absolute values are given below for $n = 17$ differences.

d_i	Rank $\lvert d_i \rvert$	d_i	Rank $\lvert d_i \rvert$
−2	10.5	−3	15
−1	4.5	3	15
3	15	3	10.5
1	4.5	−2	10.5
−1	4.5	−1	4.5
3	15	1	4.5
−1	4.5	−2	10.5
−3	15	1	4.5
1	4.5		

Then $T^+ = 73.5$ and $T^- = 79.5$. With $\alpha = .05$ and $n = 17$, the lower portion of the rejection region is $T \le 35$ (see Table 9). Since the observed value of T is $T = 73.5$, the null hypothesis is not rejected, as in Exercise 15.52.

15.54 Having made no assumptions concerning the underlying distribution of the populations, we cannot use a parametric test of means; rather we use the nonparametric Mann-Whitney U test to test the equivalence of the population distributions. The rank sums are $W_A = 126$ and $W_B = 45$. With $n_1 = n_2 = 9$,

$$U_A = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_A = 9(9) + \frac{9(10)}{2} - 126 = 0$$

$$U_B = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_B = 9(9) + \frac{9(10)}{2} - 45 = 81$$

For $n_1 = n_2 = 9$ in Table 8, the lower tail of the two-tailed rejection region is $U \le 18$ with $\alpha = 2(.0252) = .0504$. The observed value of U is $U = 0$, and the null hypothesis is rejected. We conclude that the deaf children do differ from the hearing children in eye movement rate.

15.56 a. The measurements are ordered according to magnitude and ranked from the "outside in," as described in this exercise. The resulting ranks are shown in the table at the right. The hypothesis to be tested is

$$H_0: \sigma_A^2 = \sigma_B^2 \quad \text{vs.} \quad H_a: \sigma_A^2 > \sigma_B^2$$

Instrument	Response	Rank
A	1060.21	1
B	1060.24	3
A	1060.27	5
B	1060.28	7
B	1060.30	9
B	1060.32	8
A	1060.34	6
A	1060.36	4
A	1060.40	2

and the test statistic is the Mann-Whitney U. If the alternative hypothesis is true (that is, the variance for instrument A is greater than the variance for instrument B), then the measurements for instrument A should be very low and very high in the sequence of measurements. Hence they will be assigned the lower ranks, and the sum of ranks for the A observations will be small. A one-tailed test of hypothesis is required, with α near .05. Calculating the U values, we obtain

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_A = 5(4) + \frac{5(6)}{2} - 18 = 17$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - W_B = 20 + 10 - 27 = 3$$

The rejection region will be $U \leq 3$ with $\alpha = .056$. The test statistic falls in the rejection region, and hence the null hypothesis is rejected.

b. Calculate

$$s_1^2 = \frac{\sum_j y_{1j}^2 - \left[\frac{\left(\sum_j y_{1j}\right)^2}{n_1}\right]}{n_1 - 1} = \frac{.0230}{4} = .00575$$

$$s_2^2 = \frac{\sum_j y_{2j}^2 - \left[\frac{\left(\sum_j y_{2j}\right)^2}{n_2}\right]}{n_2 - 1} = \frac{.0035}{3} = .00117$$

Then the test statistic is

$$F = \frac{s_1^2}{s_2^2} = \frac{.00575}{.00117} = 4.914$$

The critical value of F (with 4 and 3 degrees of freedom) is $F_{.05} = 9.12$. The null hypothesis is not rejected.