Chapter 15: Nonparametric Statistics

15.1 Let Y have a binomial distribution with $n = 25$ and $p = .5$. For the two–tailed sign test, the test rejects for extreme values (either too large or too small) of the test statistic whose null distribution is the same as Y. So, Table 1 in Appendix III can be used to define rejection regions that correspond to various significant levels. Thus:

Rejection region	α
$Y \le 6$ or $Y \ge 19$	$P(Y \le 6) + P(Y \ge 19) = .014$
$Y \le 7$ or $Y \ge 18$	$P(Y \le 7) + P(Y \ge 18) = .044$
$Y \le 8$ or $Y \ge 17$	$P(Y \le 8) + P(Y \ge 17) = .108$

15.2 Let p = P(blood levels are elevated after training). We will test H_0: $p = .5$ vs H_a: $p > .5$.

a. Since $m = 15$, so p–value = $P(M \ge 15) = \binom{17}{15}.5^{17} + \binom{17}{16}.5^{17} + \binom{17}{16}.5^{17} = 0.0012$.

b. Reject H_0.

c. $P(M \ge 15) = P(M > 14.5) \approx P(Z > 2.91) = .0018$, which is very close to part **a**.

15.3 Let $p = P$(recovery rate for A exceeds B). We will test H_0: $p = .5$ vs H_a: $p \ne .5$. The data are:

Hospital	A	B	Sign(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	–

a. From the above, $m = 2$ so the p–value is given by $2P(M \le 2) = .110$. Thus, in order to reject H_0, it would have been necessary that the significance level $\alpha \ge .110$. Since this is fairly large, H_0 would probably not be rejected.

b. The t–test has a normality assumption that may not be appropriate for these data. Also, since the sample size is relatively small, a large–sample test couldn't be used either.

15.4 **a.** Let $p = P$(school A exceeds school B in test score). For H_0: $p = .5$ vs H_a: $p \ne .5$, the test statistic is M = # of times school A exceeds school B in test score. From the table, we find $m = 7$. So, the p–value = $2P(M \ge 7) = 2P(M \le 3) = 2(.172) = .344$. With $\alpha = .05$, we fail to reject H_0.

b. For the one–tailed test, H_0: $p = .5$ vs H_a: $p > .5$. Here, the p–value = $P(M \ge 7) = .173$ so we would still fail to reject H_0.

15.5 Let $p = P$(judge favors mixture B). For H_0: $p = .5$ vs H_a: $p \neq .5$, the test statistic is M = # of judges favoring mixture B. Since the observed value is $m = 2$, p–value = $2P(M \leq 2)$ = 2(.055) = .11. Thus, H_0 is not rejected at the $\alpha = .05$ level.

15.6 **a.** Let $p = P$(high elevation exceeds low elevation). For H_0: $p = .5$ vs H_a: $p > .5$, the test statistic is M = # of nights where high elevation exceeds low elevation. Since the observed value is $m = 9$, p–value = $P(M \geq 9) = .011$. Thus, the data favors H_a.

b. Extreme temperatures, such as the minimum temperatures in this example, often have skewed distributions, making the assumptions of the t–test invalid.

15.7 **a.** Let $p = P$(response for stimulus 1 is greater that for stimulus 2). The hypotheses are H_0: $p = .5$ vs H_a: $p > .5$, and the test statistic is M = # of times response for stimulus 1 exceeds stimulus 2. If it is required that $\alpha \leq .05$, note that

$$P(M \leq 1) + P(M \geq 8) = .04,$$

where M is binomial($n = 9$, $p = .5$) under H_0. Our rejection region is the set {0, 1, 8, 9}. From the table, $m = 2$ so we fail to reject H_0.

b. The proper test is the paired t–test. So, with H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, the summary statistics are $\overline{d} = -1.022$ and $s_D^2 = 3.467$, the computed test statistic is $|t| = \frac{|-1.022|}{\sqrt{\frac{3.467}{9}}} = 1.65$ with 8 degrees of freedom. Since $t_{.025} = 2.306$, we fail to reject H_0.

15.8 Let $p = P(B$ exceeds $A)$. For H_0: $p = .5$ vs H_a: $p \neq .5$, the test statistic is M = # of technicians for which B exceeds A with $n = 7$ (since one tied pair is deleted). The observed value of M is 1, so the p–value = $2P(M \leq 1) = .125$, so H_0 is not rejected.

15.9 **a.** Since two pairs are tied, $n = 10$. Let $p = P$(before exceeds after) so that H_0: $p = .5$ vs H_a: $p > .5$. From the table, $m = 9$ so the p–value is $P(M \geq 9) = .011$. Thus, H_0 is not rejected with $\alpha = .01$.

b. Since the observations are counts (and thus integers), the paired t–test would be inappropriate due to its normal assumption.

15.10 There are n ranks to be assigned. Thus, $T^+ + T^-$ = sum of all ranks = $\sum_{i=1}^{n} i = n(n+1)/2$ (see Appendix I).

15.11 From Ex. 15.10, $T^- = n(n+1)/2 - T^+$. If $T^+ > n(n+1)/4$, it must be so that $T^- < n(n+1)/4$. Therefore, since $T = \min(T^+, T^-)$, $T = T^-$.

15.12 **a.** Define d_i to be the difference between the math score and the art score for the i^{th} student, $i = 1, 2, \ldots, 15$. Then, $T^+ = 14$ and $T^- = 106$. So, $T = 14$ and from Table 9, since $14 < 16$, p–value < .01. Thus H_0 is rejected.
b. H_0: identical population distributions for math and art scores vs. H_a: population distributions differ by location.

15.13 Define d_i to be the difference between school A and school B. The differences, along with the ranks of $|d_i|$ are given below.

	1	2	3	4	5	6	7	8	9	10
d_i	28	5	–4	15	12	–2	7	9	–3	13
rank $\|d_i\|$	13	4	3	9	7	1	5	6	2	8

Then, $T^+ = 49$ and $T^- = 6$ so $T = 6$. Indexing $n = 10$ in Table 9, $.02 < T < .05$ so H_0 would be rejected if $\alpha = .05$. This is a different decision from Ex. 15.4

15.14 Using the data from Ex. 15.6, $T^- = 1$ and $T^+ = 54$, so $T = 1$. From Table 9, p–value < .005 for this one–tailed test and thus H_0 is rejected.

15.15 Here, R is used:

```
> x <- c(126,117,115,118,118,128,125,120)
> y <- c(130,118,125,120,121,125,130,120)
> wilcox.test(x,y,paired=T,alt="less",correct=F)

        Wilcoxon signed rank test

data:  x and y
V = 3.5, p-value = 0.0377
alternative hypothesis: true mu is less than 0
```

The test statistic is $T = 3.5$ so H_0 is rejected with $\alpha = .05$.

15.16 **a.** The sign test statistic is $m = 8$. Thus, p–value = $2P(M \geq 8) = .226$ (computed using a binomial with $n = 11$ and $p = .5$). H_0 should not be rejected.

b. For the Wilcoxon signed–rank test, $T^+ = 51.5$ and $T^- = 14.5$ with $n = 11$. With $\alpha = .05$, the rejection region is $\{T \leq 11\}$ so H_0 is not rejected.

15.17 From the sample, $T^+ = 44$ and $T^- = 11$ with $n = 10$ (two ties). With $T = 11$, we reject H_0 with $\alpha = .05$ using Table 9.

15.18 Using the data from Ex. 12.16:

d_i	3	6.1	2	4	2.5	8.9	.8	4.2	9.8	3.3	2.3	3.7	2.5	–1.8	7.5
$\|d_i\|$	3	6.1	2	4	2.5	8.9	.8	4.2	9.8	3.3	2.3	3.7	2.5	1.8	7.5
rank	7	12	3	10	5.5	14	1	11	15	8	4	9	5.5	2	13

Thus, $T^+ = 118$ and $T^- = 2$ with $n = 15$. From Table 9, since $T^- < 16$, p–value < .005 (a one–tailed test) so H_0 is rejected.

15.19 Recall for a continuous random variable Y, the median ξ is a value such that $P(Y > \xi) = P(Y < \xi) = .5$. It is desired to test $H_0: \xi = \xi_0$ vs. $H_a: \xi \neq \xi_0$.

a. Define $D_i = Y_i - \xi_0$ and let M = # of negative differences. Very large or very small values of M (compared against a binomial distribution with $p = .5$) lead to a rejection.

b. As in part a, define $D_i = Y_i - \xi_0$ and rank the D_i according to their absolute values according to the Wilcoxon signed–rank test.

15.20 Using the results in Ex. 15.19, we have H_0: $\xi = 15{,}000$ vs. H_a: $\xi > 15{,}000$ The differences $d_i = y_i - 15000$ are:

d_i	–200	1900	3000	4100	–1800	3500	5000	4200	100	1500
$\|d_i\|$	200	1900	3000	4100	1800	3500	5000	4200	100	1500
rank	2	5	6	8	4	7	10	9	1	3

a. With the sign test, $m = 2$, p–value $= P(M \le 2) = .055$ ($n = 10$) so H_0 is rejected.

b. $T^+ = 49$ and $T^- = 6$ so $T = 6$. From Table 9, $.01 < p$–value $< .025$ so H_0 is rejected.

15.21 **a.** $U = 4(7) + \frac{1}{2}(4)(5) - 34 = 4$. Thus, the p–value $= P(U \le 4) = .0364$

b. $U = 5(9) + \frac{1}{2}(5)(6) - 25 = 35$. Thus, the p–value $= P(U \ge 35) = P(U \le 10) = .0559$.

c. $U = 3(6) + \frac{1}{2}(3)(4) - 23 = 1$. Thus, p–value $= 2P(U \le 1) = 2(.0238) = .0476$

15.22 To test: H_0: the distributions of ampakine CX–516 are equal for the two groups
H_a: the distributions of ampakine CX–516 differ by a shift in location

The samples of ranks are:

Age group											
20s	20	11	7.5	14	7.5	16.5	2	18.5	3.5	7.5	$W_A = 108$
65–70	1	16.5	7.5	14	11	14	5	11	18.5	3.5	$W_B = 102$

Thus, $U = 100 + 10(11)/2 - 108 = 47$. By Table 8,

$$p\text{–value} = 2P(U \le 47) > 2P(U \le 39) = 2(.2179) = .4358.$$

Thus, there is not enough evidence to conclude that the population distributions of ampakine CX–516 are different for the two age groups.

15.23 The hypotheses to be tested are:
H_0: the population distributions for plastics 1 and 2 are equal
H_a: the populations distributions differ by location

The data (with ranks in parentheses) are:

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

By Table 8 with $n_1 = n_2 = 6$, $P(U \le 7) = .0465$ so $\alpha = 2(.0465) = .093$. The two possible values for U are $U_A = 36 + \frac{6(7)}{2} - W_A = 27$ and $U_B = 36 + \frac{6(7)}{2} - W_B = 9$. So, $U = 9$ and thus H_0 is not rejected.

15.24 **a.** Here, $U_A = 81 + \frac{9(10)}{2} - W_A = 126 - 94 = 32$ and $U_B = 81 + \frac{9(10)}{2} - W_B = 126 - 77 = 49$. Thus, $U = 32$ and by Table 8, p–value = $2P(U \le 32) = 2(.2447) = .4894$.

b. By conducting the two sample t–test, we have H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 = \mu_2 \ne 0$. The summary statistics are $\bar{y}_1 = 8.267$, $\bar{y}_2 = 8.133$, and $s_p^2 = .8675$. The computed test stat. is $|t| = \frac{|.1334|}{\sqrt{.8675\left(\frac{2}{9}\right)}} = .30$ with 16 degrees of freedom. By Table 5, p–value > 2(.1) = .20 so H_0 is not rejected.

c. In part **a**, we are testing for a shift in distribution. In part **b**, we are testing for unequal means. However, since in the t–test it is assumed that both samples were drawn from normal populations with common variance, under H_0 the two distributions are also equal.

15.25 With $n_1 = n_2 = 15$, it is found that $W_A = 276$ and $W_B = 189$. Note that although the actual failure times are not given, they are not necessary:

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

Thus, $U = 354 - 276 = 69$ and since $E(U) = \frac{n_1 n_2}{2} = 112.5$ and $V(U) = 581.25$,

$$z = \frac{69 - 112.5}{\sqrt{581.25}} = -1.80.$$

Since $-1.80 < -z_{.05} = -1.645$, we can conclude that the experimental batteries have a longer life.

15.26 R:

```
> DDT <- c(16,5,21,19,10,5,8,2,7,2,4,9)
> Diaz <- c(7.8,1.6,1.3)
> wilcox.test(Diaz,DDT,correct=F)

        Wilcoxon rank sum test

data:  Diaz and DDT
W = 6, p-value = 0.08271
alternative hypothesis: true mu is not equal to 0
```

With $\alpha = .10$, we can reject H_0 and conclude a difference between the populations.

15.27 Calculate $U_A = 4(6) + \frac{4(5)}{2} - W_A = 34 - 34 = 0$ and $U_B = 4(6) + \frac{6(7)}{2} - W_B = 45 - 21 = 24$. Thus, we use $U = 0$ and from Table 8, p–value = $2P(U \le 0) = 2(.0048) = .0096$. So, we would reject H_0 for $\alpha \approx .10$.

15.28 Similar to previous exercises. With $n_1 = n_2 = 12$, the two possible values for U are

$$U_A = 144 + \frac{12(13)}{2} - 89.5 = 132.5 \text{ and } U_B = 144 + \frac{12(13)}{2} - 210.5 = 11.5,$$

but since it is required to detect a shift of the "B" observations to the right of the "A" observations, we let $U = U_A = 132.5$. Here, we can use the large–sample approximation. The test statistic is $z = \frac{132.5 - 72}{\sqrt{300}} = 3.49$, and since $3.49 > z_{.05} = 1.645$, we can reject H_0 and conclude that rats in population "B" tend to survive longer than population A.

15.29 H_0: the 4 distributions of mean leaf length are identical, vs. H_a: at least two are different.
R:
```
> len <-
c(5.7,6.3,6.1,6.0,5.8,6.2,6.2,5.3,5.7,6.0,5.2,5.5,5.4,5.0,6,5.6,4,5.2,
3.7,3.2,3.9,4,3.5,3.6)
> site <- factor(c(rep(1,6),rep(2,6),rep(3,6),rep(4,6)))
> kruskal.test(len~site)

        Kruskal-Wallis rank sum test

data:  len by site
Kruskal-Wallis chi-squared = 16.974, df = 3, p-value = 0.0007155
```

We reject H_0 and conclude that there is a difference in at least two of the four sites.

15.30 **a.** This is a completely randomized design.

b. R:
```
> prop<-c(.33,.29,.21,.32,.23,.28,.41,.34,.39,.27,.21,.30,.26,.33,.31)
> campaign <- factor(c(rep(1,5),rep(2,5),rep(3,5)))
> kruskal.test(prop,campaign)

        Kruskal-Wallis rank sum test

data:  prop and campaign
Kruskal-Wallis chi-squared = 2.5491, df = 2, p-value = 0.2796
```

From the above, we cannot reject H_0.

c. R:
```
> wilcox.test(prop[6:10],prop[11:15], alt="greater")

        Wilcoxon rank sum test

data:  prop[6:10] and prop[11:15]
W = 19, p-value = 0.1111
alternative hypothesis: true mu is greater than 0
```

From the above, we fail to reject H_0: we cannot conclude that campaign 2 is more successful than campaign 3.

15.31 **a.** The summary statistics are: TSS = 14,288.933, SST = 2586.1333, SSE = 11,702.8. To test H_0: $\mu_A = \mu_B = \mu_C$, the test statistic is $F = \frac{2586.1333/2}{11,702.8/12} = 1.33$ with 2 numerator and 12 denominator degrees of freedom. Since $F_{.05} = 3.89$, we fail to reject H_0. We assumed that the three random samples were independently drawn from separate normal populations with common variance. Life–length data is typically right skewed.

b. To test H_0: the population distributions are identical for the three brands, the test statistic is $H = \frac{122}{15(16)}\left(\frac{36^2}{5} + \frac{35^2}{5} + \frac{49^2}{5}\right) - 3(16) = 1.22$ with 2 degrees of freedom. Since $\chi^2_{.05}$ = 5.99, we fail to reject H_0.

15.32 **a.** Using R:

```
> time<–c(20,6.5,21,16.5,12,18.5,9,14.5,16.5,4.5,2.5,14.5,12,18.5,9,
1,9,4.5, 6.5,2.5,12)
> strain<-factor(c(rep("Victoria",7),rep("Texas",7),rep("Russian",7)))
>
> kruskal.test(time~strain)

        Kruskal-Wallis rank sum test

data:  time by strain
Kruskal-Wallis chi-squared = 6.7197, df = 2, p-value = 0.03474
```

By the above, p–value = .03474 so there is evidence that the distributions of recovery times are not equal.

b. R: comparing the Victoria A and Russian strains:

```
> wilcox.test(time[1:7],time[15:21],correct=F)

        Wilcoxon rank sum test

data:  time[1:7] and time[15:21]
W = 43, p-value = 0.01733
alternative hypothesis: true mu is not equal to 0
```

With p–value = .01733, there is sufficient evidence that the distribution of recovery times with the two strains are different.

15.33 R:

```
> weight <- c(22,24,16,18,19,15,21,26,16,25,17,14,28,21,19,24,23,17,
18,13,20,21)
> temp <- factor(c(rep(38,5),rep(42,6),rep(46,6),rep(50,5)))
>
> kruskal.test(weight~temp)

        Kruskal-Wallis rank sum test

data:  weight by temp
Kruskal-Wallis chi-squared = 2.0404, df = 3, p-value = 0.5641
```

With a p–value = .5641, we fail to reject the hypothesis that the distributions of weights are equal for the four temperatures.

15.34 The rank sums are: $R_A = 141$, $R_B = 248$, and $R_C = 76$. To test H_0: the distributions of percentages of plants with weevil damage are identical for the three chemicals, the test statistic is $H = \frac{12}{30(31)}\left(\frac{141^2}{10} + \frac{248^2}{10} + \frac{76^2}{10}\right) - 3(31) = 19.47$. Since $\chi^2_{.005} = 10.5966$, the p–value is less than .005 and thus we conclude that the population distributions are not equal.

15.35 By expanding H,

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{k} n_i\left(\overline{R}_i^2 - 2\overline{R}_i\frac{n+1}{2} + \frac{(n+1)^2}{4}\right)$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k} n_i\left(\frac{R_i^2}{n_i^2} - (n+1)\frac{R_i}{n_i} + \frac{(n+1)^2}{4}\right)$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} + \frac{12}{n}\sum_{i=1}^{k} R_i + \frac{3(n+1)}{n}\sum_{i=1}^{k} n_i$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} + \frac{12}{n}\left(\frac{n(n+1)}{2}\right) + \frac{3(n+1)}{n}\cdot n$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(n+1)\,.$$

15.36 There are 15 possible pairings of ranks: The statistic H is

$$H = \frac{12}{6(7)}\sum R_i^2\,/\,2 - 3(7) = \frac{1}{7}\left(\sum R_i^2 - 147\right).$$

The possible pairings are below, along with the value of H for each.

pairings			H
(1, 2)	(3, 4)	(5, 6)	32/7
(1, 2)	(3, 5)	(4, 6)	26/7
(1, 2)	(3, 6)	(5, 6)	24/7
(1, 3)	(2, 4)	(5, 6)	26/7
(1, 3)	(2, 5)	(4, 6)	18/7
(1, 3)	(2, 6)	(4, 5)	14/7
(1, 4)	(2, 3)	(5, 6)	24/7
(1, 4)	(2, 5)	(3, 6)	8/7
(1, 4)	(2, 6)	(3, 5)	6/7
(1, 5)	(2, 3)	(4, 6)	14/7
(1, 5)	(2, 4)	(3, 6)	6/7
(1, 5)	(2, 6)	(3, 4)	2/7
(1, 6)	(2, 3)	(4, 5)	8/7
(1, 6)	(2, 4)	(3, 5)	2/7
(1, 6)	(2, 5)	(3, 4)	0

Thus, the null distribution of H is (each of the above values are equally likely):

h	0	2/7	6/7	8/7	2	18/7	24/7	26/7	32/7
$p(h)$	1/15	2/15	2/15	2/15	2/15	1/15	2/15	2/15	1/15

15.37 R:

```
> score <- c(4.8,8.1,5.0,7.9,3.9,2.2,9.2,2.6,9.4,7.4,6.8,6.6,3.6,5.3,
2.1,6.2,9.6,6.5,8.5,2.0)
> anti <- factor(c(rep("I",5),rep("II",5),rep("III",5),rep("IV",5)))
> child <- factor(c(1:5, 1:5, 1:5, 1:5))
> friedman.test(score ~ anti | child)

        Friedman rank sum test

data:  score and anti and child
Friedman chi-squared = 1.56, df = 3, p-value = 0.6685
```

a. From the above, we do not have sufficient evidence to conclude the existence of a difference in the tastes of the antibiotics.

b. Fail to reject H_0.

c. Two reasons: more children would be required and the potential for significant child to child variability in the responses regarding the tastes.

15.38 R:

```
> cadmium <- c(162.1,199.8,220,194.4,204.3,218.9,153.7,199.6,210.7,
179,203.7,236.1,200.4,278.2,294.8,341.1,330.2,344.2)
> harvest <- c(rep(1,6),rep(2,6),rep(3,6))
> rate <- c(1:6,1:6,1:6)
> friedman.test(cadmium ~ rate | harvest)

        Friedman rank sum test

data:  cadmium and rate and harvest
Friedman chi-squared = 11.5714, df = 5, p-value = 0.04116
```

With $\alpha = .01$ we fail to reject H_0: we cannot conclude that the cadmium concentrations are different for the six rates of sludge application.

15.39 R:

```
> corrosion <- c(4.6,7.2,3.4,6.2,8.4,5.6,3.7,6.1,4.9,5.2,4.2,6.4,3.5,
5.3,6.8,4.8,3.7,6.2,4.1,5.0,4.9,7.0,3.4,5.9,7.8,5.7,4.1,6.4,4.2,5.1)
> sealant <- factor(c(rep("I",10),rep("II",10),rep("III",10)))
> ingot <- factor(c(1:10,1:10,1:10))
> friedman.test(corrosion~sealant|ingot)

        Friedman rank sum test

data:  corrosion and sealant and ingot
Friedman chi-squared = 6.6842, df = 2, p-value = 0.03536
```

With $\alpha = .05$, we can conclude that there is a difference in the abilities of the sealers to prevent corrosion.

15.40 A summary of the ranked data is

Ear	A	B	C
1	2	3	1
2	2	3	1
3	1	3	2
4	3	2	1
5	2	1	3
6	1	3	2
7	2.5	2.5	1
8	2	3	1
9	2	3	1
10	2	3	1

Thus, $R_A = 19.5$, $R_B = 26.5$, and $R_C = 14$.
To test: H_0: distributions of aflatoxin levels are equal
H_a: at least two distributions differ in location

$F_r = \frac{12}{10(3)(4)}[(19.5)^2 + (26.5)^2 + (14)^2] - 3(10)(4) = 7.85$ with 2 degrees of freedom. From Table 6, $.01 < p$–value $< .025$ so we can reject H_0.

15.41 **a.** To carry out the Friedman test, we need the rank sums, R_i, for each model. These can be found by adding the ranks given for each model. For model A, $R_1 = 8(15) = 120$. For model B, $R_2 = 4 + 2(6) + 7 + 8 + 9 + 2(14) = 68$, etc. The R_i values are:

120, 68, 37, 61, 31, 87, 100, 34, 32, 62, 85, 75, 30, 71, 67

Thus, $\sum R_i^2 = 71{,}948$ and then $F_r = \frac{12}{8(15)(16)}[71{,}948 - 3(8)(16)] = 65.675$ with 14 degrees of freedom. From Table 6, we find that p–value $< .005$ so we soundly reject the hypothesis that the 15 distributions are equal.

b. The highest (best) rank given to model H is lower than the lowest (worst) rank given to model M. Thus, the value of the test statistic is $m = 0$. Thus, using a binomial distribution with $n = 8$ and $p = .5$, p–value $= 2P(M = 0) = 1/128$.

c. For the sign test, we must know whether each judge (exclusively) preferred model H or model M. This is not given in the problem.

15.42 H_0: the probability distributions of skin irritation scores are the same for the 3 chemicals vs. H_a: at least two of the distributions differ in location.
From the table of ranks, $R_1 = 15$, $R_2 = 19$, and $R_3 = 14$. The test statistic is

$$F_r = \frac{12}{8(3)(4)}[(15)^2 + (19)^2 + (14)^2] - 3(8)(4) = 1.75$$

with 2 degrees of freedom. Since $\chi^2_{.01} = 9.21034$, we fail to reject H_0: there is not enough evidence to conclude that the chemicals cause different degrees of irritation.

15.43 If $k = 2$ and $b = n$, then $F_r = \frac{2}{n}\left(R_1^2 + R_2^2\right) - 9n$. For $R_1 = 2n - M$ and $R_2 = n + M$, then

$$\begin{aligned}F_r &= \frac{2}{n}\left[(2n-M)^2 + (n+M)^2\right] - 9n \\ &= \frac{2}{n}\left[(4n^2 - 4nM + M^2) + (n^2 + 2nM + M^2) - 4.5n^2\right] \\ &= \frac{2}{n}(-.5n^2 - 2nM + 2M^2) \\ &= \frac{4}{n}(M^2 - nM - \tfrac{1}{4}n^2) \\ &= \frac{4}{n}(M - \tfrac{1}{2}n)^2\end{aligned}$$

The Z statistic from Section 15.3 is $Z = \dfrac{M - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} = \dfrac{2}{\sqrt{n}}(M - \tfrac{1}{2}n)$. So, $Z^2 = F_r$.

15.44 Using the hints given in the problem,

$$F_r = \tfrac{12b}{k(k+1)}\sum\left(\overline{R}_i^2 - 2\overline{R}_i\overline{R} + \overline{R}^2\right) = \tfrac{12b}{k(k+1)}\sum\left(R_i^2/b^2 - (k+1)R_i/b + (k+1)^2/4\right)$$

$$= \tfrac{12b}{k(k+1)}\sum R_i^2/b^2 - \tfrac{12}{k}\tfrac{bk(k+1)}{2} + \tfrac{12b(k+1)k}{4k} = \tfrac{12}{bk(k+1)}\sum R_i^2 - 3b(k+1).$$

15.45 This is similar to Ex. 15.36. We need only work about the 3! = 6 possible rank pairing. They are listed below, with the R_i values and F_r. When $b = 2$ and $k = 3$, $F_r = \frac{1}{2}\Sigma R_i^2 - 24$.

Block 1	2	R_i
1	1	2
2	2	4
3	3	6

$F_r = 4$

Block 1	2	R_i
1	1	2
2	3	5
3	2	5

$F_r = 3$

Block 1	2	R_i
1	2	3
2	1	3
3	3	6

$F_r = 3$

Block 1	2	R_i
1	2	3
2	3	5
3	1	4

$F_r = 1$

Block 1	2	R_i
1	3	4
2	1	3
3	2	5

$F_r = 1$

Block 1	2	R_i
1	3	4
2	2	4
3	1	4

$F_r = 0$

Thus, with each value being equally likely, the null distribution is given by

$$P(F_r = 0) = P(F_r = 4) = 1/6 \text{ and } P(F_r = 1) = P(F_r = 3) = 1/3.$$

15.46 Using Table 10, indexing row (5, 5):
a. $P(R = 2) = P(R \le 2) = .008$ (minimum value is 2).
b. $P(R \le 3) = .040$.
c. $P(R \le 4) = .167$.

15.47 Here, $n_1 = 5$ (blacks hired), $n_2 = 8$ (whites hired), and $R = 6$. From Table 10,

$$p\text{–value} = 2P(R \le 6) = 2(.347) = .694.$$

So, there is no evidence of nonrandom racial selection.

15.48 The hypotheses are H_0: no contagion (randomly diseased)
H_a: contagion (not randomly diseased)
Since contagion would be indicated by a grouping of diseased trees, a small numer of runs tends to support the alternative hypothesis. The computed test statistic is $R = 5$, so with $n_1 = n_2 = 5$, p–value = .357 from Table 10. Thus, we cannot conclude there is evidence of contagion.

15.49 **a.** To find $P(R \le 11)$ with $n_1 = 11$ and $n_2 = 23$, we can rely on the normal approximation. Since $E(R) = \frac{2(11)(23)}{11+23} + 1 = 15.88$ and $V(R) = 6.2607$, we have (in the second step the continuity correction is applied)

$$P(R \le 11) = P(R < 11.5) \approx P(Z < \tfrac{11.5-15.88}{\sqrt{6.2607}}) = P(Z < -1.75) = .0401.$$

b. From the sequence, the observed value of $R = 11$. Since an unusually large or small number of runs would imply a non–randomness of defectives, we employ a two–tailed test. Thus, since the p–value = $2P(R \le 11) \approx 2(.0401) = .0802$, significance evidence for non–randomness does not exist here.

15.50 **a.** The measurements are classified as A if they lie above the mean and B if they fall below. The sequence of runs is given by

$$\underline{A\ A\ A\ A\ A}\ \underline{B\ B\ B\ B\ B\ B}\ \underline{A}\ \underline{B}\ \underline{A}\ \underline{B}\ \underline{A}$$

Thus, $R = 7$ with $n_1 = n_2 = 8$. Now, non–random fluctuation would be implied by a small number of runs, so by Table 10, p–value = $P(R \le 7) = .217$ so non–random fluctuation cannot be concluded.

b. By dividing the data into equal parts, $\bar{y}_1 = 68.05$ (first row) and $\bar{y}_2 = 67.29$ (second row) with $s_p^2 = 7.066$. For the two–sample t–test, $|t| = \frac{|68.05-67.27|}{\sqrt{7.066\left(\frac{2}{8}\right)}} = .57$ with 14 degrees of freedom. Since $t_{.05} = 1.761$, H_0 cannot be rejected.

15.51 From Ex. 15.18, let A represent school A and let B represent school B. The sequence of runs is given by

$$\underline{A}\ \underline{B}\ \underline{A}\ \underline{B}\ \underline{A}\ \underline{B\ B\ B}\ \underline{A}\ \underline{B\ B}\ \underline{A\ A}\ \underline{B}\ \underline{A}\ \underline{B}\ \underline{A\ A}$$

Notice that the 9th and 10th letters and the 13th and 14th letters in the sequence represent the two pairs of tied observations. If the tied observations were reversed in the sequence of runs, the value of R would remain the same: $R = 13$. Hence the order of the tied observations is irrelevant.

The alternative hypothesis asserts that the two distributions are not identical. Therein, a small number of runs would be expected since most of the observations from school A would fall below those from school B. So, a one–tailed test is employed (lower tail) so the p–value = $P(R \le 13) = .956$. Thus, we fail to reject the null hypothesis (similar with Ex. 15.18).

15.52 Refer to Ex. 15.25. In this exercise, $n_1 = 15$ and $n_2 = 16$. If the experimental batteries have a greater mean life, we would expect that most of the observations from plant B to be smaller than those from plant A. Consequently, the number of runs would be small. To use the large sample test, note that $E(R) = 16$ and $V(R) = 7.24137$. Thus, since $R = 15$, the approximate p–value is given by

$$P(R \le 15) = P(R < 15.5) \approx P(Z < -.1858) = .4263.$$

Of course, the hypotheses H_0: the two distributions are equal, would not be rejected.

15.53 R:

```
> grader <- c(9,6,7,7,5,8,2,6,1,10,9,3)
> moisture <- c(.22,.16,.17,.14,.12,.19,.10,.12,.05,.20,.16,.09)
> cor(grader,moisture,method="spearman")
[1] 0.911818
```

Thus, $r_S = .911818$. To test for association with $\alpha = .05$, index .025 in Table 11 so the rejection region is $|r_S| > .591$. Thus, we can safely conclude that the two variables are correlated.

15.54 R:

```
> days <- c(30,47,26,94,67,83,36,77,43,109,56,70)
> rating <- c(4.3,3.6,4.5,2.8,3.3,2.7,4.2,3.9,3.6,2.2,3.1,2.9)
> cor.test(days,rating,method="spearman")

        Spearman's rank correlation rho

data:  days and rating
S = 537.44, p-value = 0.0001651
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.8791607
```

From the above, $r_S = -.8791607$ and the p–value for the test H_0: there is no association is given by p–value = .0001651. Thus, H_0 is rejected.

15.55 R:

```
> rank <- c(8,5,10,3,6,1,4,7,9,2)
> score <- c(74,81,66,83,66,94,96,70,61,86)
> cor.test(rank,score,alt = "less",method="spearman")
```

```
        Spearman's rank correlation rho

data:  rank and score
S = 304.4231, p-value = 0.001043
alternative hypothesis: true rho is less than 0
sample estimates:
       rho
-0.8449887
```

a. From the above, $r_S = -.8449887$.

b. With the p–value = .001043, we can conclude that there exists a negative association between the interview rank and test score. Note that we only showed that the correlation is negative and not that the association has some specified level.

15.56 R:

```
> rating <- c(12,7,5,19,17,12,9,18,3,8,15,4)
> distance <- c(75,165,300,15,180,240,120,60,230,200,130,130)
> cor.test(rating,distance,alt = "less",method="spearman")

        Spearman's rank correlation rho

data:  rating and distance
S = 455.593, p-value = 0.02107
alternative hypothesis: true rho is less than 0
sample estimates:
       rho
-0.5929825
```

a. From the above, $r_S = -.5929825$.

b. With the p–value = .02107, we can conclude that there exists a negative association between rating and distance.

15.57 The ranks for the two variables of interest x_i and y_i corresponding the math and art, respectively) are shown in the table below.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$R(x_i)$	1	3	2	4	5	7.5	7.5	9	10.5	12	13.5	6	13.5	15	10.5
$R(y_i)$	5	11.5	1	2	3.5	8.5	3.5	13	6	15	11.5	7	10	14	8.5

Then, $r_S = \dfrac{15(1148.5) - 120(120)}{\sqrt{[15(1238.5) - 120^2]^2}} = .6768$ (the formula simplifies as shown since the samples of ranks are identical for both math and art). From Table 11 and with $\alpha = .10$, the rejection region is $|r_S| > .441$ and thus we can conclude that there is a correlation between math and art scores.

15.58 R:

```
> bending <- c(419,407,363,360,257,622,424,359,346,556,474,441)
> twisting <- c(227,231,200,211,182,304,384,194,158,225,305,235)
> cor.test(bending,twisting,method="spearman",alt="greater")
```

```
        Spearman's rank correlation rho

data:  bending and twisting
S = 54, p-value = 0.001097
alternative hypothesis: true rho is greater than 0
sample estimates:
      rho
0.8111888
```

a. From the above, $r_S = .8111888$.
b. With a p–value = .001097, we can conclude that there is existence of a population association between bending and twisting stiffness.

15.59 The data are ranked below; since there are no ties in either sample, the alternate formula for r_S will be used.

$R(x_i)$	2	3	1	4	6	8	5	10	7	9
$R(y_i)$	2	3	1	4	6	8	5	10	7	9
d_i	0	0	0	0	0	0	0	0	0	0

Thus, $r_S = 1 - \frac{6[(0)^2+(0)^2+\ldots+(0)^2}{10(99)} = 1 - 0 = 1$.

From Table 11, note that 1 > .794 so the p–value < .005 and we soundly conclude that there is a positive correlation between the two variables.

15.60 It is found that $r_S = .9394$ with $n = 10$. From Table 11, the p–value < 2(.005) = .01 so we can conclude that correlation is present.

15.61 **a.** Since all five judges rated the three products, this is a randomized block design.

b. Since the measurements are ordinal values and thus integers, the normal theory would not apply.

c. Given the response to part b, we can employ the Friedman test. In R, this is (using the numbers 1–5 to denote the judges):

```
> rating <- c(16,16,14,15,13,9,7,8,16,11,7,8,4,9,2)
> brand <- factor(c(rep("HC",5),rep("S",5),rep("EB",5)))
> judge <- c(1:5,1:5,1:5)
> friedman.test(rating ~ brand | judge)

        Friedman rank sum test

data:  rating and brand and judge
Friedman chi-squared = 6.4, df = 2, p-value = 0.04076
```

With the (approximate) p–value = .04076, we can conclude that the distributions for rating the egg substitutes are not the same.

15.62 Let $p = P$(gourmet A's rating exceeds gourmet B's rating for a given meal). The hypothesis of interest is H_0: $p = .5$ vs H_a: $p \neq .5$. With M = # of meals for which A is superior, we find that

$$P(M \leq 4) + P(M \geq 13) = 2P(M \leq 4) = .04904.$$

using a binomial calculation with $n = 17$ (3 were ties) and $p = .5$. From the table, $m = 8$ so we fail to reject H_0.

15.63 Using the Wilcoxon signed–rank test,

```
> A <- c(6,4,7,8,2,7,9,7,2,4,6,8,4,3,6,9,9,4,4,5)
> B <- c(8,5,4,7,3,4,9,8,5,3,9,5,2,3,8,10,8,6,3,5)
> wilcox.test(A,B,paired=T)

        Wilcoxon signed rank test

data:  A and B
V = 73.5, p-value = 0.9043
alternative hypothesis: true mu is not equal to 0
```

With the p–value = .9043, the hypothesis of equal distributions is not rejected (as in Ex. 15.63).

15.64 For the Mann–Whitney U test, $W_A = 126$ and $W_B = 45$. So, with $n_1 = n_2 = 9$, $U_A = 0$ and $U_B = 81$. From Table 8, the lower tail of the two–tailed rejection region is $\{U \leq 18\}$ with $\alpha = 2(.0252) = .0504$. With $U = 0$, we soundly reject the null hypothesis and conclude that the deaf children do differ in eye movement rate.

15.65 With $n_1 = n_2 = 8$, $U_A = 46.5$ and $U_B = 17.5$. From Table 8, the hypothesis of no difference will be rejected if $U \leq 13$ with $\alpha = 2(.0249) = .0498$. Since our $U = 17.5$, we fail to reject H_0 (same as in Ex. 13.1).

15.66 **a.** The measurements are ordered below according to magnitude as mentioned in the exercise (from the "outside in"):

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

To test H_0: $\sigma_A^2 = \sigma_B^2$ vs. H_a: $\sigma_A^2 > \sigma_B^2$, we use the Mann–Whitney U statistic. If H_a is true, then the measurements for A should be assigned lower ranks. For the significance level, we will use $\alpha = P(U \leq 3) = .056$. From the above table, the values are $U_1 = 17$ and $U_2 = 3$. So, we reject H_0.

b. For the two samples, $s_A^2 = .00575$ and $s_B^2 = .00117$. Thus, $F = .00575/.00117 = 4.914$ with 4 numerator and 3 denominator degrees of freedom. From R:

```
> 1 - pf(4.914,4,3)
[1] 0.1108906
```

Since the p–value = .1108906, H_0 would not be rejected.

15.67 First, obviously $P(U \le 2) = P(U = 0) + P(U = 1) + P(U = 2)$. Denoting the five observations from samples 1 and 2 as A and B respectively (and $n_1 = n_2 = 5$), the only sample point associated with $U = 0$ is

$$B\ B\ B\ B\ B\ A\ A\ A\ A\ A$$

because there are no A's preceding any of the B's. The only sample point associated with $U = 1$ is

$$B\ B\ B\ B\ A\ B\ A\ A\ A\ A$$

since only one A observation precedes a B observation. Finally, there are two sample points associated with $U = 2$:

$$B\ B\ B\ A\ B\ B\ A\ A\ A\ A \qquad B\ B\ B\ B\ A\ A\ B\ A\ A\ A$$

Now, under the null hypothesis all of the $\binom{10}{5} = 252$ orderings are equally likely. Thus,

$$P(U \le 2) = 4/252 = 1/63 = .0159.$$

15.68 Let Y = # of positive differences and let T = the rank sum of the positive differences. Then, we must find $P(T \le 2) = P(T = 0) + P(T = 1) + P(T = 2)$. Now, consider the three pairs of observations and the ranked differences according to magnitude. Let d_1, d_2, and d_3 denote the ranked differences. The possible outcomes are:

d_1	d_2	d_3	Y	T
+	+	+	3	6
–	+	+	2	5
+	–	+	2	4
+	+	–	2	3
–	–	+	1	3
–	+	–	1	2
+	–	–	1	1
–	–	–	0	0

Now, under H_0 Y is binomial with $n = 3$ and $p = P(A$ exceeds $B) = .5$. Thus,
$P(T = 0) = P(T = 0, Y = 0) = P(Y = 0)P(T = 0 \mid Y = 0) = .125(1) = .125$.

Similarly, $P(T = 1) = P(T = 1, Y = 1) = P(Y = 1)P(T = 1 \mid Y = 1) = ..375(1/3) = .125$, since conditionally when $Y = 1$, there are three possible values for T (1, 2, or 3).

Finally, $P(T = 2) = P(T = 2, Y = 1) = P(Y = 1)P(T = 2 \mid Y = 1) = ..375(1/3) = .125$, using similar logic as in the above.

Thus, $P(T \le 2) = .125 + .125 + .125 = .375$.

15.69 **a.** A composite ranking of the data is:

Line 1	Line 2	Line 3
19	14	2
16	10	15
12	5	4
20	13	11
3	9	1
18	17	8
21	7	6
$R_1 = 109$	$R_2 = 75$	$R_3 = 47$

Thus,

$$H = \frac{12}{21(22)}\left[\frac{109^2}{7} + \frac{75^2}{7} + \frac{47}{7}\right] = 3(22) = 7.154$$

with 2 degrees of freedom. Since $\chi^2_{.05} = 5.99147$, we can reject the claim that the population distributions are equal.

15.70 **a.** R:

```
> rating <- c(20,19,20,18,17,17,11,13,15,14,16,16,15,13,18,11,8,
12,10,14,9,10)
> supervisor <- factor(c(rep("I",5),rep("II",6),rep("III",5),
rep("IV",6)))
> kruskal.test(rating~supervisor)

        Kruskal-Wallis rank sum test

data:  rating by supervisor
Kruskal-Wallis chi-squared = 14.6847, df = 3, p-value = 0.002107
```

With a p–value = .002107, we can conclude that one or more of the supervisors tend to receive higher ratings

b. To conduct a Mann–Whitney U test for only supervisors I and III,

```
> wilcox.test(rating[12:16],rating[1:5], correct=F)

        Wilcoxon rank sum test

data:  rating[12:16] and rating[1:5]
W = 1.5, p-value = 0.02078
alternative hypothesis: true mu is not equal to 0
```

Thus, with a p–value = .02078, we can conclude that the distributions of ratings for supervisors I and III differ by location.

15.71 Using Friedman's test (people are blocks), $R_1 = 19$, $R_2 = 21.5$, $R_3 = 27.5$ and $R_4 = 32$. To test H_0: the distributions for the items are equal vs.
H_a: at least two of the distributions are different

the test statistic is $F_r = \frac{12}{10(4)(5)}\left[19^2 + (21.5)^2 + (27.5)^2 + 32^2\right] - 3(10)(5) = 6.21$.

With 3 degrees of freedom, $\chi^2_{.05} = 7.81473$ and so H_0 is not rejected.

15.72 In R:

```
> perform <- c(20,25,30,37,24,16,22,25,40,26,20,18,24,27,39,41,21,25)
> group <- factor(c(1:6,1:6,1:6))
> method <- factor(c(rep("lect",6),rep("demonst",6),rep("machine",6)))
> friedman.test(perform ~ method | group)

        Friedman rank sum test

data:  perform and method and group
Friedman chi-squared = 4.2609, df = 2, p-value = 0.1188
```

With a p–value = .1188, it is unwise to reject the claim of equal teach method effectiveness, so fail to reject H_0.

15.73 Following the methods given in Section 15.9, we must obtain the probability of observing exactly Y_1 runs of S and Y_2 runs of F, where $Y_1 + Y_2 = R$. The joint probability mass functions for Y_1 and Y_2 is given by

$$p(y_1, y_2) = \frac{\binom{7}{y_1-1}\binom{7}{y_2-1}}{\binom{16}{8}}.$$

(1) For the event $R = 2$, this will only occur if $Y_1 = 1$ and $Y_2 = 1$, with either the S elements or the F elements beginning the sequence. Thus,
$$P(R = 2) = 2p(1, 1) = \tfrac{2}{12{,}870}.$$
(2) For $R = 3$, this will occur if $Y_1 = 1$ and $Y_2 = 2$ or $Y_1 = 2$ and $Y_2 = 1$. So,
$$P(R = 3) = p(1, 2) + p(2, 1) = \tfrac{14}{12{,}870}.$$
(3) Similarly, $P(R = 4) = 2p(2, 2) = \frac{98}{12{,}870}$.

(4) Likewise, $P(R = 5) = p(3, 2) + p(2, 3) = \frac{294}{12{,}870}$.

(5) In the same manor, $P(R = 6) = 2p(3, 3) = \frac{882}{12{,}870}$.

Thus, $P(R \le 6) = \frac{2+14+98+294+882}{12{,}870} = .100$, agreeing with the entry found in Table 10.

15.74 From Ex. 15.67, it is not difficult to see that the following pairs of events are equivalent:

$$\{W = 15\} \equiv \{U = 0\},\ \{W = 16\} \equiv \{U = 2\}, \text{ and } \{W = 17\} \equiv \{U = 3\}.$$

Therefore, $P(W \le 17) = \text{P}(U \le 3) = .0159$.

15.75 Assume there are n_1 "A" observations and n_2 "B" observations, The Mann–Whitney U statistic is defined as

$$U = \sum_{i=1}^{n_2} U_i \,,$$

where U_i is the number of A observations preceding the i^{th} B. With $B_{(i)}$ to be the i^{th} B observation in the combined sample after it is ranked from smallest to largest, and write $R[B_{(i)}]$ to be the rank of the i^{th} ordered B in the total ranking of the combined sample. Then, U_i is the number of A observations the precede $B_{(i)}$. Now, we know there are $(i-1)$ B's that precede $B_{(i)}$, and that there are $R[B_{(i)}] - 1$ A's and B's preceding $B_{(i)}$. Then,

$$U = \sum_{i=1}^{n_2} U_i = \sum_{i=1}^{n_2} [R(B_{(i)}) - i] = \sum_{i=1}^{n_2} R(B_{(i)}) - \sum_{i=1}^{n_2} i = W_B - n_2(n_2+1)/2$$

Now, let $N = n_1 + n_2$. Since $W_A + W_B = N(N+1)/2$, so $W_B = N(N+1)/2 - W_A$. Plugging this expression in to the one for U yields

$$U = N(N+1)/2 - n_2(n_2+1)/2 - W_A = \tfrac{N^2+N+n_2^2+n_2}{2} - W_A$$
$$= \tfrac{n_1^2+2n_1n_2+n_2^2+n_1+n_2-n_2^2-n_2}{2} - W_A = n_1n_2 + \tfrac{n_1(n_1+1)}{2} - W_A \,.$$

Thus, the two tests are equivalent.

15.76 Using the notation introduced in Ex. 15.65, note that

$$W_A = \sum_{i=1}^{n_1} R(A_i) = \sum_{i=1}^{N} X_i \,,$$

where

$$X_i = \begin{cases} R(z_i) & \text{if } z_i \text{ is from sample } A \\ 0 & \text{if } z_i \text{ is from sample } B \end{cases}$$

If H_0 is true,

$$E(X_i) = R(z_i)P[X_i = R(z_i)] + 0\cdot P(X_i = 0) = R(z_i)\tfrac{n_1}{N}$$
$$E(X_i^2) = [R(z_i)]^2\tfrac{n_1}{N}$$
$$V(X_i) = [R(z_i)]^2\tfrac{n_1}{N} - \left(R(z_i)\tfrac{n_1}{N}\right)^2 = [R(z_i)]^2\left(\tfrac{n_1(N-n_1)}{N^2}\right).$$
$$E(X_i, X_j) = R(z_i)R(z_i)P[X_i = R(z_i), X_j = R(z_j)] = R(z_i)R(z_i)\left(\tfrac{n_1}{N}\right)\left(\tfrac{n_1-1}{N-1}\right).$$

From the above, it can be found that $\text{Cov}(X_i, X_j) = R(z_i)R(z_i)\left[\tfrac{-n_1(N-n_1)}{N^2(N-1)}\right]$.

Therefore,

$$E(W_A) = \sum_{i=1}^{N} E(X_i) = \tfrac{n_1}{N}\sum_{i=1}^{N} R(z_i) = \tfrac{n_1}{N}\left(\tfrac{N(N+1)}{2}\right) = \tfrac{n_1(N+1)}{2}$$

and

$$\begin{aligned}
V(W_A) &= \sum_{i=1}^{N} V(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\
&= \tfrac{n_1(N-n_1)}{N^2} \sum_{i=1}^{N} [R(z_i)]^2 - \tfrac{n_1(N-n_1)}{N^2(N-1)} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} R(z_i)R(z_j) - \sum_{i=1}^{N} [R(z_i)]^2 \right] \\
&= \tfrac{n_1(N-n_1)}{N^2} \left[\tfrac{N(N+1)N2N+1)}{6} \right] - \tfrac{n_1(N-n_1)}{N^2(N-1)} \left\{ \left[\sum_{i=1}^{N} R(z_i) \right]^2 - \sum_{i=1}^{N} [R(z_i)]^2 \right\} \\
&= \tfrac{2n_1(N-n_1)(N+1)(2N+1)}{12N} - \tfrac{n_1(N-n_1)}{N^2(N-1)} \left[\tfrac{N^2(N+1)^2}{4} - \tfrac{N(N+1)(2N+1)}{6} \right] \\
&= \tfrac{n_1 n_2 (n_1+n_2+1)}{12} \left[\tfrac{4N+2}{N} - \tfrac{(3N+2)(N-1)}{n(N-1)} \right] = \tfrac{n_1 n_2 (n_1+n_2+1)}{12} .
\end{aligned}$$

From Ex. 15.75 it was shown that $U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_A$. Thus,

$$E(U) = n_1 n_2 + \tfrac{n_1(n_1+1)}{2} - E(W_A) = \tfrac{n_1 n_2}{2}$$
$$V(U) = V(W_A) = \tfrac{n_1 n_2 (n_1+n_2+1)}{12} .$$

15.77 Recall that in order to obtain T, the Wilcoxon signed–rank statistic, the differences d_i are calculated and ranked according to absolute magnitude. Then, using the same notation as in Ex. 15.76,

$$T^+ = \sum_{i=1}^{N} X_i$$

where

$$X_i = \begin{cases} R(D_i) & \text{if } D_i \text{ is positive} \\ 0 & \text{if } D_i \text{ is negative} \end{cases}$$

When H_0 is true, $p = P(D_i > 0) = \frac{1}{2}$. Thus,

$$\begin{aligned}
E(X_i) &= R(D_i)P[X_i = R(D_i)] = \tfrac{1}{2}R(D_i) \\
E(X_i^2) &= [R(D_i)]^2 P[X_i = R(D_i)] = \tfrac{1}{2}[R(D_i)]^2 \\
V(X_i) &= \tfrac{1}{2}[R(D_i)]^2 = [\tfrac{1}{2}R(D_i)]^2 = \tfrac{1}{4}[R(D_i)]^2 \\
E(X_i, X_j) &= R(D_i)R(D_j)P[X_i = R(D_i), X_j = R(D_j)] = \tfrac{1}{4}R(D_i)R(D_j) .
\end{aligned}$$

Then, Cov(X_i, X_j) = 0 so

$$E(T^+) = \sum_{i=1}^{n} E(X_i) = \tfrac{1}{2}\sum_{i=1}^{n} R(D_i) = \tfrac{1}{2}\left(\tfrac{n(n+1)}{2}\right) = \tfrac{n(n+1)}{4}$$

$$V(T^+) = \sum_{i=1}^{n} V(X_i) = \tfrac{1}{4}\sum_{i=1}^{n} [R(D_i)]^2 = \tfrac{1}{4}\left(\tfrac{n(n+1)(2n+1)}{6}\right) = \tfrac{n(n+1)(2n+1)}{24} .$$

Since $T^- = \frac{n(n+1)}{2} - T^+$ (see Ex. 15.10),

$$E(T^-) = E(T^+) = E(T)$$
$$V(T^-) = V(T^+) = V(T) .$$

15.78 Since we use X_i to denote the rank of the i^{th} "X" sample value and Y_i to denote the rank of the i^{th} "Y" sample value,

$$\sum_{i=1}^n X_i = \sum_{i=1}^n Y_i = \tfrac{n(n+1)}{2} \text{ and } \sum_{i=1}^n X_i^2 = \sum_{i=1}^n Y_i^2 = \tfrac{n(n+1)(2n+1)}{6}.$$

Then, define $d_i = X_i - Y_i$ so that

$$\sum_{i=1}^n d_i^2 = \sum_{i=1}^n \left(X_i^2 - 2X_iY_i + Y_i^2\right) = \tfrac{n(n+1)(2n+1)}{6} - 2\sum_{i=1}^n X_iY_i + \tfrac{n(n+1)(2n+1)}{6}$$

and thus

$$\sum_{i=1}^n X_iY_i = \tfrac{n(n+1)(2n+1)}{6} - \tfrac{1}{2}\sum_{i=1}^n d_i^2 .$$

Now, we have

$$\begin{aligned}
r_S &= \frac{n\sum_{i=1}^n X_iY_i - \left(\sum_{i=1}^n X_i\right)\left(\sum_{i=1}^n Y_i\right)}{\sqrt{\left[n\sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i\right)^2\right]}\sqrt{\left[n\sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i\right)^2\right]}} \\
&= \frac{\frac{n^2(n+1)(2n+1)}{6} - \frac{n}{2}\sum_{i=1}^n d_i^2 - \frac{n^2(n+1)^2}{4}}{\frac{n^2(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4}} \\
&= \frac{\frac{n^2(n+1)(n-1)}{12} - \frac{n}{2}\sum_{i=1}^n d_i^2}{\frac{n^2(n+1)(n-1)}{12}} \\
&= 1 - \frac{\frac{n}{2}\sum_{i=1}^n d_i^2}{\frac{n^2(n^2-1)}{12}} \\
&= 1 - \frac{6\sum_{i=1}^n d_i^2}{n(n^2-1)} .
\end{aligned}$$