Chapter 2

Exercise 2.1 For each student, the ranks of the observations are:

Student	Month 1	Month 2	Month 3	Month 4
1	2	4	3	4
Rank	1	3.5	2	3.5
2	0	1	3	4
Rank	1	2	3	4
3	4	5	4	7
Rank	1.5	3	1.5	4
4	3	3	4	3
Rank	2	2	4	2
5	0	0	1	3
Rank	1.5	1.5	3	4
6	4	3	5	5
Rank	2	1	3.5	3.5
7	5	5	4	2
Rank	3.5	3.5	2	1
Rank Total	12.5	16.5	19.0	22.0

Some observations are tied, therefore, formula (2.2) will be used to calculate the Friedman rank test statistics. We have $n = 7$, $k = 4$, $R_1 = 12.5$, $R_2 = 16.5$, $R_3 = 19.0$, $R_4 = 22.0$, and the sum of squares of all ranks is $\sum_{i=1}^{7}\sum_{j=1}^{4} r_{ij}^2 = 205.5$. The test statistic is given as

$$Q = \frac{7(4-1)\Big[(1/7)\big(12.5^2+16.5^2+19.0^2+22.0^2\big)-(7)(4)(4+1)^2/4\Big]}{205.5-(7)(4)(4+1)^2/4} = 4.7705.$$

Using Table A.5 with $k = 4$, $n = 7$, and $\alpha = 0.05$, we obtain the critical value of 7.8. Since the observed statistic doesn't exceed the critical value, we fail to reject the null hypothesis $H_0 : \theta_{month\,1} = \theta_{month\,2} = \theta_{month\,3} = \theta_{month\,4}$. The conclusion is that the location parameters for the distribution of the number of books read don't differ by month.

In SAS, we type the code:

```
data books;
   input student month response @@;
 datalines;
1 1 2 1 2 4 1 3 3 1 4 4 2 1 0 2 2 1 2 3 3 2 4 4
3 1 4 3 2 5 3 3 4 3 4 7 4 1 3 4 2 3 4 3 4 4 4 3
5 1 0 5 2 0 5 3 1 5 4 3 6 1 4 6 2 3 6 3 5 6 4 5
7 1 5 7 2 5 7 3 4 7 4 2
;

proc sort data=books;
   by student;
run;

proc rank data=books out=ranked;
  var response;
    by student;
     ranks rank;
run;
```

```
proc freq data=ranked;
  table student*month*rank/noprint cmh;
run;
```

The test statistic and P-value produced by SAS are

```
Alternative Hypothesis   Value     Prob
Row Mean Scores Differ  4.7705   0.1894
```

The large P-value indicates that H_0 should not be rejected. This is in sync with our previous conclusion.

EXERCISE 2.2 Within-customer ranks of the data are:

Customer	Letter	Phone	Text
1	0.4	0.3	0.1
Rank	3	2	1
2	0.8	0.4	0.3
Rank	3	2	1
3	0.5	0.4	0.1
Rank	3	2	1
4	0.7	0.6	0.2
Rank	3	2	1
5	0.6	0.3	0.2
Rank	3	2	1
6	0.6	0.5	0.4
Rank	3	2	1
7	0.6	0.4	0.3
Rank	3	2	1
8	0.7	0.6	0.2
Rank	3	2	1
Rank Total	24	16	8

Because no ties are observed, the Friedman rank test statistic is computed according to (2.1). For $k = 3$, $n = 8$, $R_1 = 24$, $R_2 = 16$, and $R_3 = 8$,

$$Q = \frac{12(24^2 + 16^2 + 8^2)}{(8)(3)(3+1)} - (3)(8)(3+1) = 16\,.$$

From Table A.5, the critical value corresponding to $\alpha = 0.01$ is 9, thus, we reject the null hypothesis. The conclusion is that the three methods of customer contact differ. The SAS below outputs the test statistic and the P-value.

```
data TVcustomers;
    input customer $ contact $ probability @@;
  datalines;
1 letter 0.4 1 phone 0.3 1 text 0.1
2 letter 0.8 2 phone 0.4 2 text 0.3
3 letter 0.5 3 phone 0.4 3 text 0.1
4 letter 0.7 4 phone 0.6 4 text 0.2
5 letter 0.6 5 phone 0.3 5 text 0.2
6 letter 0.6 6 phone 0.5 6 text 0.4
7 letter 0.6 7 phone 0.4 7 text 0.3
8 letter 0.7 8 phone 0.6 8 text 0.2
```

```
;

proc sort data=TVcustomers;
  by customer;
run;

proc rank data=TVcustomers out=ranked;
  var probability;
    by customer;
      ranks rank;
run;

proc freq data=ranked;
   table customer*contact*rank/noprint cmh;
run;
```

The output is

```
Alternative Hypothesis    Value    Prob
Row Mean Scores Differ  16.0000  0.0003
```

The small P-value supports our conclusion that the methods differ significantly. To investigate which methods differ, we conduct a pairwise comparison using the Wilcoxon signed-rank test. The pairwise differences in observations are shown in the table below.

Customer	Letter-Phone	Letter-Text	Phone-Text
1	0.1	0.3	0.2
2	0.4	0.5	0.1
3	0.1	0.4	0.3
4	0.1	0.5	0.4
5	0.3	0.4	0.1
6	0.1	0.2	0.1
7	0.2	0.3	0.1
8	0.1	0.5	0.4

To test whether differences between pairs of methods exist, for each column of ranks compute the test statistic $T = \min(T^+, T^-)$. For Letter vs. Phone, $T = T^- = 9$, for the other two tests, $T = T^- = 0$. The critical value for $n = 8$, the two-sided alternative, and significance level $\alpha = 0.05$ is 3. Hence, we conclude that Letter and Phone do not differ significantly. The critical value for $n = 8$, two-sided alternative, and $\alpha = 0.01$ is 0. This means that Letter and Text as well as Phone and Text are statistically different even at a 0.01 level of significance.

Running SAS produces the same result. The code is

```
data contact;
   input letter phone text;
       diff_lp=letter-phone;
          diff_lt=letter-text;
             diff_pt=phone-text;
cards;
0.4 0.3 0.1
0.8 0.4 0.3
0.5 0.4 0.1
0.6 0.7 0.2
0.6 0.3 0.2
```

```
0.6 0.5 0.4
0.4 0.6 0.3
0.7 0.6 0.2
;

proc univariate data=contact;
   var diff_lp diff_lt diff_pt;
run;
```

The output is as follows:

• To compare Letter to Phone, the test statistic is $S = n(n+1)/4 - T^- = 8(8+1)/4 - 9 = 18 - 9 = 9$.

```
Test            -Statistic-     ----p Value----
Signed Rank     S          9    Pr >= |S|   0.2656
```

• To compare Letter to Text (or Phone to Text), the test statistic is $S = n(n+1)/4 - T^- = 8(8+1)/4 - 0 = 18 - 0 = 18$.

```
Test            -Statistic-     ----p Value----
Signed Rank     S         18    Pr >= |S|   0.0078
```

The conclusion coincides with the one drawn by hand. Contacts by Letter and by Phone are not significantly different, whereas contact by Text differs from both by Letter and by Phone.

EXERCISE 2.3 The ranks are assigned as follows:

Fish Pond	Lead Contents (in ppb)						Total
A	3	4	4	5	7	8	
Rank	1	3	3	5.5	9	10	31.5
B	10	11	11	12	15	18	
Rank	12.5	14.5	14.5	16	17	18	92.5
C	4	5	6	6	9	10	
Rank	3	5.5	7.5	7.5	11	12.5	47

Since tied ranks are assigned, we will be using the definition given by (2.4) to compute the Kruskal-Wallis test statistic. We know that $n_1 = n_2 = n_3 = 6$, $N = n_1 + n_2 + n_3 = 18$, $R_1 = 31.5$, $R_2 = 92.5$, and $R_3 = 47$. The number of tied 4's is $T_1 = 3$, 5's is $T_2 = 2$, 6's is $T_3 = 2$, 10's is $T_4 = 2$, and 11's is $T_5 = 2$.

The denominator in (2.4) is equal to

$$1 - \frac{(3^3 - 3) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2)}{18^3 - 18} = 0.99174,$$

and, hence, the H-statistic is derived as

$$H = \Big[\frac{12\big(31.5^2/6 + 92.5^2/6 + 47^2/6\big)}{(18)(18+1)} - 3(18+1)\Big]/0.00826 = 11.8552.$$

From Table A.6, the critical value corresponding to $n_1 = n_2 = n_3 = 6$ and $\alpha = 0.05$ is 5.719. Thus, we reject the null hypothesis and conclude that the fish ponds differ in lead contents.

We run the following code in SAS:

```
data lead_contents;
 input pond $ lead @@;
 datalines;
 A 3 A 4 A 4 A 5 A 7 A 8
 B 10 B 11 B 11 B 12 B 15 B 18
 C 4 C 5 C 6 C 6 C 9 C 10
 ;

proc npar1way data=lead_contents wilcoxon;
   class pond;
      var lead;
   exact;
run;
```

The output is

```
          Kruskal-Wallis Test
Chi-Square                     11.8552
Exact      Pr >= Chi-Square  2.379E-04
```

Since the P-value is very small, H_0 is rejected. Next, to see which ponds differ in lead contents, we conduct pairwise two-tailed Wilcoxon rank-sum tests.

- To compare ponds A and C, we write

Fish Pond	Lead Contents (in ppb)						Total
A	3	4	4	5	7	8	
Rank	1	3	3	5.5	9	10	31.5
C	4	5	6	6	9	10	
Rank	3	5.5	7.5	7.5	11	12	46.5

The test statistic is $W = 31.5$, the sum of the ranks in the first sample since samples are of equal sizes. The lower-tailed critical value from Table A.2 for $n_1 = n_2 = 6$ and $\alpha = 0.05$ is $W_L = 26$ and the upper-tailed one is $W_U = 52$. Since $W_L < W < W_U$, the null hypothesis should not be rejected. We conclude at the 5% significance level that there is no difference in lead content between ponds A and C.

- To compare ponds B and C, we write

Fish Pond	Lead Contents (in ppb)						Total
B	10	11	11	12	15	18	
Rank	6.5	8.5	8.5	10	11	12	56.5
C	4	5	6	6	9	10	
Rank	1	2	3.5	3.5	5	6.5	21.5

The test statistic is $W = 56.5$, the sum of the ranks in the first sample since the sizes of the two samples are the same. The upper-tailed critical value from Table A.2 for $n_1 = n_2 = 6$ and $\alpha = 0.01$ is $W_U = 55$. Thus, the null hypothesis should be rejected even at the 1% level of significance. We conclude that the lead content in pond B differs from that in pond C.

The code in SAS that does the pairwise testing is given as

```
proc npar1way data=lead_contents wilcoxon;
   class pond;
     var lead;
```

```
         exact;
   where (pond ne 'B');
run;

proc npar1way data=lead_contents wilcoxon;
   class pond;
     var lead;
         exact;
   where (pond ne 'A');
run;
```

The relevant output is

- for testing A vs. C

```
Wilcoxon Two-Sample Test Statistic (S) 31.5000
Exact Test Two-Sided
Pr >= |S - Mean| 0.2489
```

- for testing B vs. C

```
Wilcoxon Two-Sample Test
Statistic (S) 56.5000
Exact Test
Two-Sided Pr >= |S - Mean| 0.0043
```

The P-value is larger than 0.05 when comparing A and C, and is less than 0.01 when comparing B and C. These results yield the same conclusion as above.

EXERCISE 2.4 After assigning ranks, we get

24°C	Rank	28°C	Rank	32°C	Rank	36°C	Rank
88	15	67	4	93	16	86	13
54	1	72	5	82	11	87	14
65	3	76	7	84	12	81	10
55	2	80	9	78	8	73	6
Total	21		25		47		43

There are no ties among the ranks, therefore, (2.3) will be used for computation of the Kruskal-Wallis test statistic. We have that $n_1 = n_2 = n_3 = n_4 = 4$, $N = 16$, $R_1 = 21$, $R_2 = 25$, $R_3 = 47$, and $R_4 = 43$. The H-statistics is

$$H = \frac{12(21^2/4 + 25^2/4 + 47^2/4 + 43^2/4)}{16(16+1)} - 3(16+1) = 5.5147.$$

Next, we look up the critical value in Table A.6. For $\alpha = 0.05$ and the sample sizes 4, 4, 4, and 4, the critical value is 7.235. The observed test statistic is smaller than the critical value, indicating that the null hypothesis should not be rejected. The conclusion is that the germination rates don't differ. No post-hoc pairwise comparison is necessary.

The code in SAS is:

```
data germination;
   input temperature $ rate @@;
 datalines;
 24 88 24 54 24 65 24 55
 28 67 28 72 28 76 28 80
 32 93 32 82 32 84 32 78
 36 86 36 87 36 81 36 73
 ;

proc npar1way data=germination wilcoxon;
   class temperature;
     var rate;
   exact;
run;
```

The output is

```
        Kruskal-Wallis Test
Chi-Square                   5.5147
Exact      Pr >= Chi-Square  0.1349
```

The P-values is larger than 0.05, hence the null is not rejected.