SAS Program Notes
Biostatistics: A Guide to Design, Analysis, and Discovery
Chapter 9: Nonparametric Tests

In the program notes for chapter 9, we discuss non-parametric or distribution-free methods.

Note 9.1 – The Sign Test and Wilcoxon Signed Rank Test

1. The Sign Test

To illustrate the use of the sign test, suppose that we have a group of patients who suffer from arthritis in both hands. Two ointments are available to relieve pain temporarily and reduce swelling and can be applied to hands of patients when needed. A researcher would like to test the effectiveness of ointments A and B on her patient population to see if patients respond to one ointment better than the other. She administers ointments randomly when patients experience a flair-up so that if ointment A is applied to a patient's right hand then ointment B is applied to left hand and vise-versa. Ten minutes later the patient is ask to rank the relief experienced in each hand on a scale of 1 to 5 where 1 is "no relief" and 5 is "complete relief". The table below shows the data on pain relief from 12 patients.

Table 1.

	Ointment	
Patient ID	**A**	**B**
1	2	4
2	4	5
3	3	3
4	3	2
5	1	5
6	3	4
7	4	5
8	4	3
9	2	4
10	2	3
11	4	2
12	3	3

By looking through rows of data the research can tell how many times ointment A was more effective than ointment B just by counting the number of times the score for ointment A was higher than the score for ointment B. If the scores are the same, then there is no noticeable difference between the two ointments. The actual score differences between ointment A and B are provided below:

(ointment A – ointment B): -2, -1, 0, +1, -4, -1, -1, +1, -2, -1, +2, 0

The null hypothesis is that there is no difference between ointments A and B. Basically we are stating that the population median of the differences in scores between the two ointments, represented by Δ, is equal to zero. We can write the following null and alternative hypotheses:

H_0: $\Delta = 0$ versus
H_a: $\Delta \neq 0$

When using the sign test just consider the signs of the differences and eliminate all differences that are equal to 0 because we cannot assign a positive or negative sign to 0. Therefore we are left with the data below:

(ointment A – ointment B): -, -, eliminated, +, -, -, -, +, -, -, +, eliminated

Out of the 10 signs, if there were no difference, we would expect 5 to be positive and 5 to be negative. We could also say that we expect 50%, 1/2, or 0.5 of the signs to be positive. If π represents the probability of a positive sign, then the null and alternative hypotheses can be written as

H_0: $\pi = p = 0.5$ versus
H_a: $\pi \neq p$.

If the sample size is small, we can use the binomial distribution. The SAS commands below can be used to test the whether the proportion of positive signs differs significantly from 50%. We use **PROC FREQ** along with the EXACT statement to obtain the exact p-value of 0.3438 for the two-sided test.

SAS commands:

```
DATA ARTHRITIS;
INPUT CODE @@;
DATALINES;
1 1 0 1 1 1 0 1 1 0
;
PROC FREQ;
  TABLES CODE/BINOMIAL(P=0.5);
  EXACT BINOMIAL;
RUN;
```

SAS output:

```
                 The SAS System
               The FREQ Procedure

                                     Cumulative    Cumulative
CODE     Frequency      Percent       Frequency      Percent
--------------------------------------------------------------
```

```
          0          3       30.00           3       30.00
          1          7       70.00          10      100.00


             Binomial Proportion for CODE = 0
            -----------------------------------
            Proportion (P)                0.3000
            ASE                           0.1449
            95% Lower Conf Limit          0.0160
            95% Upper Conf Limit          0.5840

            Exact Conf Limits
            95% Lower Conf Limit          0.0667
            95% Upper Conf Limit          0.6525

               Test of H0: Proportion = 0.5

            ASE under H0                  0.1581
            Z                            -1.2649
            One-sided Pr <  Z             0.1030
            Two-sided Pr > |Z|            0.2059

            Exact Test
            One-sided Pr <=  P            0.1719
            Two-sided = 2 * One-sided     0.3438

                      Sample Size = 10
```

As another example, consider the data in Table 9.2, here we find the day 1 and 2 caloric intake data for the 14 boys with the more extreme caloric intakes on day1. We would like to test for "reversion towards the mean". The SAS commands below show the entry of the 14 extreme values (7 smallest and 7 largest) for calories for the first day of recording as well as the corresponding second day's values. The variable CODE is '0' to indicate the observation is one of the smallest values for day 1 and CODE is '1' to indicate that the observation is one of the largest values. The variable DIFF is the day 1 value minus the day 2 value. The variable REVERSION is '0' to indicate reversion towards the mean and '1' to indicate otherwise. For example, if one of the smallest day 1 values (CODE=0) has a day 1 and 2 difference that is negative (DIFF<0) then reversion toward the mean occurred (REVERSION=0). Similarly if one of the largest day 1 values (CODE=1) has a day 1 and 2 difference that is positive (DIFF>0) then reversion toward the mean occurred (REVERSION=0).

SAS commands:

```
DATA EXTREME;
INPUT DAY1 DAY2 CODE;
DIFF=DAY1-DAY2;
IF CODE=0 AND DIFF<0 THEN REVERSION=0;
  ELSE IF CODE=0 AND DIFF>0 THEN REVERSION=1;
IF CODE=1 AND DIFF>0 THEN REVERSION=0;
```

```
  ELSE IF CODE=1 AND DIFF<0 THEN REVERSION=1;
DATALINES;
1053  2484  0
4322  2926  1
1753  1054  0
3532  3289  1
2842  2849  1
1505  1925  0
3076  2431  1
1292   810  0
3049  2573  1
3277  2185  1
1781  1844  0
2773  3236  1
1645  2269  0
1723  3163  0
;
PROC FREQ;
  TABLES REVERSION/BINOMIAL(P=0.5);
  EXACT BINOMIAL;
RUN;
```

SAS output:

The SAS System
The FREQ Procedure

REVERSION	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	10	71.43	10	71.43
1	4	28.57	14	100.00

Binomial Proportion
for REVERSION = 0

Proportion (P)	0.7143
ASE	0.1207
95% Lower Conf Limit	0.4776
95% Upper Conf Limit	0.9509
Exact Conf Limits	
95% Lower Conf Limit	0.4190
95% Upper Conf Limit	0.9161

Test of H0: Proportion = 0.5

ASE under H0	0.1336
Z	1.6036
One-sided Pr > Z	0.0544
Two-sided Pr > |Z|	0.1088

```
Exact Test
One-sided Pr >=  P           0.0898
Two-sided = 2 * One-sided    0.1796

          Sample Size = 14
```

2. The Wilcoxon Signed Rank Test

The SAS commands below once again show the entry of the 14 extreme values (7 smallest and 7 largest) for calories for the first day of recording as well as the corresponding second day's values. Recall that variable CODE indicates whether or not the observation is one of the smallest values (CODE=1) or one of the largest values (CODE=0). However, the variable DIFF is the **day 1 value minus the day 2 value** for the 7 largest day 1 observations, and it is the **day 2 value minus the day 1 value** for the 7 smallest day 1 observations.

In this case, we use **PROC UNIVARIATE**, and specify the variable DIFF after the VAR statement. After running the procedure and obtaining the SAS output shown below, we should notice that the results of three tests are provided. We will skip the Student's t, and find the statistic labeled M next to the word Sign under Test in the **PROC UNIVARIATE** output. The value M is the number of positive sample values minus the expected number of values greater than zero if the null hypothesis is true. In our example, there are 10 difference values greater than zero and we would expect 7 (= 14/2) positive values under the null hypothesis. Therefore M=3 in this case. The p-value associated with this test, Prob > | M |, is the p-value for a two-sided alternative hypothesis. Therefore, for our one-sided alternative, we must divide the indicated probability by 2 to obtain the desired p-value. Notice that we have obtained the same results as those displayed for the exact binomial test.

PROC UNIVARIATE also performs the Wilcoxon Signed Rank test. The test statistic is labeled S and is located next to Signed Rank and the corresponding p-value for a two-sided alternative hypothesis is given by Prob > = | S |. The signed rank statistic, S, that SAS uses is the sum of the ranks of the positive values minus the sum expected under the null hypothesis, (n*(n+1)/4). In this example, S is 29.5 (= 82 - 52.5). Since we are performing a one-sided test, the p-value is 1/2 the value that SAS has reported, that is, it is 0.0338. You can consider using a t distribution approximation to calculate the p-value when the sample size is greater than 20. SAS adjusts its calculation of the t statistic for ties in the data.

SAS commands:

```
DATA EXTREME;
INPUT DAY1 DAY2 CODE;
DIFF=DAY2-DAY1;
IF CODE=1 THEN DIFF=DAY1-DAY2;
DATALINES;
1053  2484  0
4322  2926  1
```

```
1753   1054   0
3532   3289   1
2842   2849   1
1505   1925   0
3076   2431   1
1292    810   0
3049   2573   1
3277   2185   1
1781   1844   0
2773   3236   1
1645   2269   0
1723   3163   0
;
PROC UNIVARIATE;
  VAR DIFF;
RUN;
```

SAS output:

```
                     The SAS System
                 The UNIVARIATE Procedure
                    Variable:  DIFF

                Tests for Location: Mu0=0

Test           -Statistic-    -----p Value------

Student's t    t  2.293105    Pr >  |t|   0.0392
Sign           M         3    Pr >= |M|   0.1796
Signed Rank    S      29.5    Pr >= |S|   0.0676
```

Note 9.2 – The Wilcoxon Rank Sum Test or Mann-Whitney Test

In Example 9.5, we applied the Wilcoxon Rank Sum test to the data in Table 9.5 on the proportion of calories from fat for boys in grades 5-6 and 7-8. **PROC NPAR1WAY** followed by WILCOXON is used to perform the Wilcoxon Rank Sum test. The test statistic is the smallest sum of ranks and the p-value is calculated using a normal and t distribution approximation as well as a chi-square approximation. The test statistic, Statistic, is 224.5 as is shown in the text. The two-sided test p-value from: 1) the normal distribution approximation is 0.6358; 2) the t distribution approximation is 0.6390; and 3) from the chi-square approximation is 0.6229.

SAS commands:

```
PROC NPAR1WAY WILCOXON;
```

```
  CLASS GRADE;
  VAR PERCENTFAT;
RUN;
```

SAS output:

```
                         The SAS System

                     The NPAR1WAY Procedure

       Wilcoxon Scores (Rank Sums) for Variable PERCENTFAT
                   Classified by Variable GRADE

                           Sum of      Expected       Std Dev          Mean
GRADE        N             Scores      Under H0      Under H0         Score
___________________________________________________________________________

0           14             224.50         238.0     27.450696     16.035714
1           19             336.50         323.0     27.450696     17.710526

                  Average scores were used for ties.


                      Wilcoxon Two-Sample Test

                   Statistic              224.5000

                   Normal Approximation
                   Z                       -0.4736
                   One-Sided Pr <  Z        0.3179
                   Two-Sided Pr > |Z|       0.6358

                   t Approximation
                   One-Sided Pr <  Z        0.3195
                   Two-Sided Pr > |Z|       0.6390

             Z includes a continuity correction of 0.5.


                        Kruskal-Wallis Test

                   Chi-Square               0.2419
                   DF                            1
                   Pr > Chi-Square          0.6229
```

Note 9.3 – The Kruskal-Wallis Test

PROC NPAR1WAY is also used to perform the Kruskal-Wallis test. For example, consider the data in Table 9.8 on simulated reductions in diastolic blood pressure for three treatment groups. The same commands used above are also used here since the Kruskal-Wallis test is an extension of the Wilcoxon Rank Sum test. SAS shows the sum of the ranks, the expected rank sums and the mean rank for each of the groups. The test statistic is labeled Chi-Square and its p-value is given by Pr > Chi-Square.

SAS commands:

```
PROC IMPORT FILE='C:\TABLE9-8.XLS' OUT=TABLE98 REPLACE;
RUN;

DATA TABLE9_8;
  SET TABLE98;
PROC NPAR1WAY WILCOXON;
  CLASS GROUP;
  VAR REDUCTION;
RUN;
```

SAS output:

```
                              The SAS System

                          The NPAR1WAY Procedure

               Wilcoxon Scores (Rank Sums) for Variable reduction
                        Classified by Variable group

                          Sum of      Expected       Std Dev          Mean
    group        N        Scores      Under H0      Under H0         Score
    ______________________________________________________________________

    0            8        164.50         160.0     28.696466      20.56250
    1           15        436.50         300.0     34.574335      29.10000
    2           16        179.00         320.0     34.956383      11.18750

                      Average scores were used for ties.


                            Kruskal-Wallis Test

                        Chi-Square         19.2065
                        DF                       2
                        Pr > Chi-Square     <.0001
```

Note 9.4 – The Friedman Test

In Example 9.7, the results of a randomized block design are shown to evaluate the effectiveness of five different insecticides along with a control (check) group. We can use the Friedman test to determine if there is a difference in the number of living adult plum curculios emerging from caged areas after being treated by the different insecticides. The Value of the test statistic is 19.6043. The p-value, Prob, is 0.0015.

SAS commands:

```
data insect;
input block insecticide number;
datalines;
1 1  14
1 2   7
1 3   6
1 4  95
1 5  37
1 6 212
2 1   6
2 2   1
2 3   1
2 4 133
2 5  31
2 6 172
3 1   8
3 2   0
3 3   1
3 4  86
3 5  13
3 6 202
4 1  36
4 2  15
4 3   4
4 4 115
4 5  69
4 6 217
;
proc freq;
  tables block*insecticide*number/cmh2 scores=rank noprint;
run;
```

SAS output:

```
                          The SAS System
                        The FREQ Procedure

              Summary Statistics for insect by number
                       Controlling for block

      Cochran-Mantel-Haenszel Statistics (Based on Rank Scores)

  Statistic    Alternative Hypothesis    DF       Value      Prob
  ---------------------------------------------------------------
      1        Nonzero Correlation        1     10.9034    0.0010
      2        Row Mean Scores Differ     5     19.6043    0.0015
```

Total Sample Size = 24