Statistical Methods and Computing, STAT:2010

Instructor: Cowles

Lab 10

Apr. 26, 2019

1 Nonparametric alternatives to the paired t-test

The sign test and Wilcoxon's signed rank test are two nonparametric tests that can be used for paired data when the assumptions of the paired t-test are not met. Results of all three tests are reported in the "Tests of Location" section of SAS `proc univariate` output.

The sign test can be used even when quantitative measurements are not available. It uses only the signs (positive, negative, or zero) of values. It has the weakest assumptions but also the least power of the three tests.

The null hypothesis of the sign test and Wilcoxon's signed rank test is that the median of the population of differences is 0.

$$H_0 : M = 0$$

The alternative hypothesis may be either one-sided or two-sided .

The dataset called `arthritis.txt` shows the number of hours of relief provided by two analgesic drugs in twelve patients in a crossover trial of arthritis pain. Each patient took one of the drugs for two weeks, then had a two-week washout period, and then took the other drug for two weeks. The data are paired, with a measurement for each drug on each patient. The columns are

```
drugA -- hours of relief provided by drug A
drugB -- hours of relief provided by drug B
```

We wish to use these data to determine whether one drug or the other provides longer pain relief. We would like to use a paired t-test if its assumptions are met.

```
options linesize = 75 ;

data arthritis ;
input drugA drugB ;
diff = drugB - drugA;
datalines ;
2.0  3.5
3.6  5.7
2.6  2.9
2.6  2.4
7.3  9.9
3.4  3.3
```



```
14.9 16.7
6.6 6.0
2.3 3.8
2.0 4.0
6.8 9.1
8.5 20.9
;
run ;

proc print ;
run ;

proc univariate plot ;
var diff ;
run ;

proc means mean median min max ;
        var drugA drugB ;
run ;
```

Can the paired t-test be used safely for these data?

The results of the three tests are:

Tests for Location: Mu0=0

Test	-Statistic-		-----p Value------	
Student's t	t	2.167709	Pr > |t|	0.0530
Sign	M	3	Pr >= |M|	0.1460
Signed Rank	S	32	Pr >= |S|	0.0088

We will ignore the t-test results. The Wilcoxon signed rank test is our best choice here. We have quantitative data, which the sign test would waste. The signed rank test gives strong evidence (p = 0.0088 for the two-sided test) that the two population medians are not equal. The sign test has much lower power, and, as expected, it gives a larger p-value.

If we look at summary statistics for each drug separately, we see that drug B provided longer pain relief than drug A in this sample.

Variable	Mean	Median	Minimum	Maximum

```
drugA          5.2166667       3.5000000       2.0000000       14.9000000
drugB          7.3500000       4.8500000       2.4000000       20.9000000
-------------------------------------------------------------------------
```

2 Nonparametric alternative to the two-independent-sample t-test

The Wilcoxon rank sum test may be used when two independent simple random samples of a quantitative variable have been collected from two populations, but the assumptions of the two-independent-sample t-test are not met. The Wilcoxon rank sum test converts numeric values to ranks. Thus, it is not sensitive outliers as the t-test is. It may be used to test the null hypothesis that the medians of two populations are the same.

$$H_0 : M_1 = M_2$$
$$H_A : M_1 \neq M_2$$
$$or$$
$$H_A : M_1 > M_2$$
$$or$$
$$H_A : M_1 < M_2$$

However, this usage of the Wilcoxon rank sum test requires the assumption that the two population distributions are the same shape. However, without making this assumption, the Wilcoxon rank sum test may be used to test more general hypotheses as follows:

$$H_0 : \text{the two distributions are the same}$$
$$H_A : \text{one has values that are systematically larger}$$

We will work example 28.1 from the online chapter of the textbook. "Does the presence of small numbers of weeds reduce the yield of corn? Lambs-quarter is a common weed in corn fields. A researcher planted corn at the same rate in eight small plots of ground, then weeded the corn rows by hand to allow no weeds in four randomly selected plots and exactly three lambs-quarter plants per meter of row in the other four plots."

```
data corn ;
input weeds yield ;
datalines ;
0 166.7
0 172.2
0 165.0
0 176.9
3 158.6
3 176.4
3 153.1
```



```
3 156.0
;
run ;

proc univariate plot ;
var yield ;
by weeds ;
run ;

proc npar1way wilcoxon ;
class weeds ;
var yield ;
run ;
```

The sample sizes are very small (4 in each sample), and there is an outlier in the 3-weeds sample. We should not use the two-sample t-test. We use `proc npar1way` to perform the Wilcoxon rank sum test.

The results are:

```
            Wilcoxon Scores (Rank Sums) for Variable yield
                      Classified by Variable weeds

                                Sum of      Expected       Std Dev          Mean
   weeds          N             Scores     Under H0      Under H0         Score

   0              4               23.0         18.0      3.464102         5.750
   3              4               13.0         18.0      3.464102         3.250


                        Wilcoxon Two-Sample Test

                      Statistic             23.0000

                      Normal Approximation
                      Z                      1.2990
                      One-Sided Pr >  Z      0.0970
                      Two-Sided Pr > |Z|     0.1939

                      t Approximation
                      One-Sided Pr >  Z      0.1175
                      Two-Sided Pr > |Z|     0.2351
```

As the problem is stated (whether weeds *reduce* the yield), we are doing a one-sided test. The normal approximation gives a p-value of 0.0970, while the t approximation gives 0.1175. If we are doing our test at significance level 0.05, we cannot reject in either case. This is not surprising, since such small samples will result in low power for our test.

3 A nonparametric alternative to ANOVA

The Kruskal-Wallis test may be used to test hypotheses about the distributions of a quantitative variable in three or more populations when the assumptions required for ANOVA are not met. If the shapes of the distributions of the variable can be assumed to be the same in all populations, then the null hypothesis of the Kruskal-Wallis test may be expressed as equality of medians in all populations. If that assumption is not met, then the null hypothesis is better expressed as the variable having the same distribution in all of the populations.

We will revisit an example from lab 8.

Research by Singh et al. (1999) as reported in the journal *Clinical Immunology and Immunopathology* is concerned with immune abnormalities in autistic chilcren. As part of their research, they took measurements on the serum concentration of an antigen in three samples of children, autistic children, normal children, and mentally-handicapp0ed children (non-Down's-syndrome). All children were 10 years old or younger.

This dataset contains two variables:

```
concentration of the antigen (in units per milliliter of serum)
group,    coded A for autistic
               N for normal
               M for mentally handicapped
```

Here is SAS code for reading and plotting the data.

```
data autistic ;
input conc group $ ;
datalines ;
<data>
;
run ;

proc sort data = autistic ;
by group ;
run ;

proc univariate plot data = autistic ;
var conc ;
by group ;
run ;

proc means data = autistic ;
var conc ;
by group ;
run ;
```


The boxplots reveal outliers, particularly in the autistic sample, and the sample standard deviations are very different. ANOVA should not be used. Also, the shapes of the sample data in the three groups suggest that the shapes of the population distributions are not the same. Thus, we will use the Kruskal-Wallis test and will phrase our hypotheses as:

H_0 : antigen concentrations have the same distribution in all three populations
H_A : antigen concentrations are systematically higher in some populations than others

```
proc npar1way wilcoxon ;
       class group ;
       var conc ;
run ;
```

The NPAR1WAY Procedure

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

group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	23	1094.00	828.0	81.355202	47.565217
M	15	491.50	540.0	70.964383	32.766667
N	33	970.50	1188.0	86.706081	29.409091

Average scores were used for ties.

Kruskal-Wallis Test

Chi-Square	10.9635
DF	2
Pr > Chi-Square	0.0042

The data provide very strong evidence (p = 0.0042) against H_0. We conclude that the distributions of antigen concentration are not the same in the three populations.