P8130: Biostatistical Methods I

Lecture 11: Non-parametric Tests

Cody Chiuzan, PhD
Department of Biostatistics
Mailman School of Public Health (MSPH)

Lectures 11: Outline

- Non-parametric tests
  - Sign test for one-sample/paired-samples
  - Wilcoxon Signed-Rank test for paired samples <-> paired t-test
  - Wilcoxon Rank-Sum for two-independent samples <-> independent t-test

Non-parametric Tests

- Almost all the tests covered thus far relied on the normality assumption
- We assumed that samples had underlying normal distributions
  - Small samples usually do not have a normal distribution
- If this normality assumption fails, we need to use non-parametric (distribution-free methods)
  - The non-parametric methods do not impose a specific form of the distribution
  - The hypotheses to be tested relate more to the nature of the distribution as whole rather than specific values for the parameters

Sign Test

- One-sample test for location
- Assumption: data come from a continuous distribution
- The null hypothesis states that the distribution of a random variable X is symmetric about zero.
- Testing the symmetry around zero is commonly used with paired data, where X is the difference between the two sets of values.
- Sign test only considers the ‘sign’ but not the magnitude of the differences

Sign Test

- Let $d_1, d_2, \ldots, d_n$ be the differences between *n* pairs with an underlying continuous distribution
- Let Δ be the median of the distribution:
  - If $\Delta = 0$, then the positive and negative differences are equally likely
  - If $\Delta > 0$, then a positive difference is more likely than a negative difference
  - If $\Delta < 0$, then a positive difference is less likely than a negative difference
- For two-sided test, the hypotheses would be:

$$H_0: \Delta = 0 \; vs \; H_1: \Delta \neq 0 \quad (1)$$

i.e., the positive and negative differences are equally likely.

Sign Test

- Each difference can either be positive and negative
  - Zero differences are not counted with either positive and negatives
- If we ignore the zero differences, and let C be the number of positive differences, under the null hypothesis:

$C \sim Bin(n^*, 0.5)$, n^* is the total number of ‘+’ and ‘-’ differences.

- Hypotheses stated in (1) resume to one-sample binomial test:

$$H_0:\ \Delta = 0 \quad vs\ \ H_1:\ \Delta \neq 0 \quad (1)$$

$$H_0:\ \ p = 0.5\ vs\ H_1:\ p \neq 0.5\ (2)$$

Sign Test

Under Normal-Approximation: $n^*p(1-p) \geq 5$

- Reject H_0 if:

$$C \geq \left(\frac{n^*}{2} + \frac{1}{2} + z_{1-\alpha/2}\sqrt{\frac{n^*}{4}}\right) \text{ or } C \leq \left(\frac{n^*}{2} - \frac{1}{2} - z_{1-\alpha/2}\sqrt{\frac{n^*}{4}}\right)$$

- If $C > \frac{n^*}{2}, p - value = 2\times\left[1 - \Phi\left(\frac{C-\frac{n^*}{2}-\frac{1}{2}}{\sqrt{\frac{n^*}{4}}}\right)\right]$
- If $C < \frac{n^*}{2}, p - value = 2\times\left[\Phi\left(\frac{C-\frac{n^*}{2}+\frac{1}{2}}{\sqrt{\frac{n^*}{4}}}\right)\right]$
- If $C = \frac{n^*}{2}, p - value = 1$

Sign Test

Using exact methods:

- If $C > \frac{n^*}{2}, p - value = 2 \times \sum_{k=C}^{n^*} \binom{n^*}{k} \left(\frac{1}{2}\right)^{n^*}$

- If $C < \frac{n^*}{2}, p - value = 2 \times \sum_{k=0}^{C} \binom{n^*}{k} \left(\frac{1}{2}\right)^{n^*}$

- If $C = \frac{n^*}{2}, p - value = 1$

Sign Test vs Wilcoxon Signed-Rank Test

- The Sign test does not use the magnitudes of the differences
- If the number of positive differences is high, this is evidence that the distribution is asymmetric around zero
- Wilcoxon Signed-Rank test is also used for paired data, but it is based on the ranks of the absolute values of the differences
- Wilcoxon Signed-Rank test is the non-parametric equivalent of the Paired t-test.
  - Also assumes an underlying continuous symmetric distribution

Wilcoxon Signed-Rank Test

- Absolute values of the differences are put in ascending order of the magnitudes: ranks 1 to n^* (zero differences are ignored)
  - If several differences have the same value, assign the average rank

- Let T_+ be the sum of the ranks for positive values

  T_- be the sum of the ranks for negative values

- Under the null hypothesis, we would expect T_+ and T_- not to differ

Wilcoxon Signed-Rank Test

- If no ties, the test statistic is:

$$T = \frac{\left|T_+ - \frac{n^*(n^*+1)}{4}\right| - \frac{1}{2}}{\sqrt{n^*(n^*+1)(2n^*+1)/24}}$$

- If ties, the test statistic is:

$$T = \frac{\left|T_+ - \frac{n^*(n^*+1)}{4}\right| - \frac{1}{2}}{\sqrt{\frac{n^*(n^*+1)(2n^*+1)}{24} - \sum_{i=1}^{g}({t_i}^3 - t_i)/48}},$$

where t_i refers to the number of differences with the same absolute value in the i^{th} tied group and g is the number of tied groups.

Wilcoxon Signed-Rank Test

The hypotheses to be tested are:

- H_0: the median difference between the two groups is zero
- H_1: the median is not zero

Under Normal-Approximation: $n^* \geq 16$.

- Reject H_0 if $T > z_{1-\alpha/2}$, with $p - value = 2\times[1- \Phi(T)]$.

For exact methods, there are tables giving the critical values or software can be easily used.

Wilcoxon Rank-Sum Test

- Wilcoxon Rank-Sum test is the non-parametric equivalent of the Two-Sample Independent t-test.
- If the medians for the two populations are different, then one population should tend to have larger values (ranks in the pooled sample) than the other population.
- Wilcoxon-Rank Sun test is based on the ranks of the combined samples.
- Also assumes underlying continuous symmetric distributions of the populations

Wilcoxon Rank-Sum Test

- Absolute values of the differences are ranked together and given ranks from 1 to n^* (zero differences are ignored)
  - Any group of tied ranks is given the midrank of the group
- Let T_1 be the sum of the ranks for one group (X)

  T_2 be the sum of the ranks for the other group (Y)
- Focus on T_1
  - The smallest value which T_1 can take arises when all the X values are less than all the Y values: $T_1 = \left(\frac{1}{2}\right) n_1(n_1 + 1)$
  - The maximum value which T_1 can take arises when all the X values are greater than all the Y values: $T_1 = n_1 n_2 + \left(\frac{1}{2}\right) n_1(n_1 + 1)$
  - The null expectation of $T_1\ is\ \left(\frac{1}{2}\right) n_1(n_1 + n_2 + 1)$

Wilcoxon Rank-Sum Test

- The test statistic will be sum of the ranks in the first population.
- If no ties, the test statistic is:

$$T = \frac{\left|T_1 - \frac{n_1(n_1 + n_2 + 1)}{2}\right| - \frac{1}{2}}{\sqrt{(n_1 n_2/12)\ (n_1 + n_2 + 1)}}$$

- If ties, the test statistic is:

$$T = \frac{\left|T_1 - \frac{n_1(n_1 + n_2 + 1)}{2}\right| - \frac{1}{2}}{\sqrt{(n_1 n_2/12)\ [(n_1 + n_2 + 1) - \sum_{i=1}^{g}(t_i{}^2 - 1)/(n_1 + n_2)(n_1 + n_2 + 1)]}}$$

where t_i refers to the number of observations with the same absolute value in the i^{th} tied group and g is the number of tied groups.

Wilcoxon Rank-Sum Test

Hypotheses to be tested are:

- H_0: the medians of the two populations are equal
- H_1: the medians of the two populations are not equal

Under Normal-Approximation: $n_1 \ and \ n_2 \geq 10$.

- Reject H_0 if $T > z_{1-\alpha/2}$, with $p - value = 2\times[1\text{-}\ \Phi(\mathrm{T})]$.

For exact methods, there are tables giving the critical values or software can be easily used.

Non-parametric test for ONE-WAY ANOVA

- The non-parametric equivalent for ONE-WAY ANOVA is called Kruskal-Wallis (KW) test
- KW test is used for testing 3+ independent groups
- KW is testing the equality of population medians among 3+ groups
  - Assumes an identically-shaped and scaled distribution for each group, except for the difference in medians.
- Identical to ONE-WAY ANOVA where the data is replaced by their ranks.

Final thoughts

- The non-parametric tests are less powerful than the parametric ones
- The t-tests are pretty robust to mild deviations of normality
- The normality assumption is best tested using *residuals* generated from regression models (later in this course)
  - QQplots
  - Tests for normality (e.g., Shapiro-Wilk test)
- Still, given the direct correspondence between residuals and observed values, one can also use histograms with overlaid normal density curve to assess data normality
  - Careful: normality should be tested and satisfied for each group (when independent)!

Readings

Rosner, *Fundamentals of Biostatistics*, Chapter 9

- Sections: 9.1 – 9.4