

## STAT 518 --- Chapter 5: Methods Based on Ranks

- The rank of an observation is its position in the ordered sample (ordered from smallest to largest).

- For example, in the sample

the corresponding ranks are

- Many of the methods in Chapter 5 rely on the ranks of the sample observations rather than the actual data values.

- This tends to reduce the effect of \_\_\_\_\_, which makes rank-based methods ideal for data coming from \_\_\_\_\_ distributions.

### Section 5.7: One-Sample or Matched-Pairs Data

- The t-test is a common parametric procedure to test whether the mean of a population equals some specified number.

- When we have paired data, we can test whether the two random variables have the same mean using the paired t-test.

- These classical tests require the sample (or the sample of differences) to come from a \_\_\_\_\_ distribution.

- If the normality assumption is questionable, but if our data (or differences) come from a symmetric distribution, and are measured on at least an interval scale, then the Wilcoxon Signed Ranks Test may be used.
- Note that earlier we learned the \_\_\_\_\_ Test as a procedure that could be used in this situation.
- The \_\_\_\_\_ Test did not require a \_\_\_\_\_ distribution.

### Checking for Normality/Symmetry

- A quick graphical check for whether data are normally distributed is the normal Q-Q plot.
- Sample quantiles are plotted against expected quantiles if the data were truly normal.

### Using the R function `qqnorm`:

A roughly straight pattern →

Bow-shaped curved pattern →

S-shaped snaking pattern →

- See examples on the course web page.

## Wilcoxon Signed Ranks Test

- Suppose we have  $n'$  paired observations  $(X_1, Y_1), (X_2, Y_2), \dots, (X_{n'}, Y_{n'})$ .
- The differences  $D_i$  are defined as
- The absolute differences are then
- We remove any observations for which  $D_i = 0$ , leaving  $n$  untied pairs, where  $n \leq n'$ .
- We rank the  $n$  remaining pairs from 1 to  $n$  from the smallest absolute difference to the largest.
- If several pairs have the same absolute difference, we use midranks (average the ranks that would have been assigned) for these.
- For  $i = 1, \dots, n$ , the signed ranks are
- The test statistic is

- If  $T^+$  is large, it is evidence that the  $Y_i$  tends to be \_\_\_\_\_ than the  $X_i$ .
- If  $T^+$  is small, it is evidence that the  $Y_i$  tends to be \_\_\_\_\_ than the  $X_i$ .
- The null distribution (when  $E(D_i)=0$  for all  $i$ ) has been calculated.
- Table A12 gives lower quantiles of the null distribution when none of the absolute differences are tied and when  $n \leq 50$ .
- Upper quantiles can be found via the formula:
- If there are many ties or if  $n > 50$ , the test statistic using all the signed ranks

is used.

- This  $T$  has an approximately standard normal null distribution, so Table A1 is used.

| <u>3 Sets of Hypotheses</u> |              |              |
|-----------------------------|--------------|--------------|
| Two-tailed                  | Lower-tailed | Upper-tailed |
| $H_0:$                      | $H_0:$       | $H_0:$       |
| $H_1:$                      | $H_1:$       | $H_1:$       |

- **The corresponding decision rules in each case are:**
- **If the large-sample tests are used, the decision rule is based on  $T$  and the normal quantiles from Table A1.**
- **P-values can be approximated by interpolating within Table A12 or by the normal approximation.**
- **In practice, software is the best way to get p-values.**

**Example 1: 12 sets of twins were scored for “aggressiveness” on a personality test. The aggressiveness data are given on page 355. Can we conclude that the firstborn twins have greater median aggressiveness than the second twins? (Use  $\alpha = .05$ .)**

- **See the R function `wilcox.test` for performing this test in R.**

**Example 2: Problem 1 on page 365 describes a study in which the reaction time of 20 randomly sampled individuals was measured, both before and after an alcoholic drink for each person. Is there evidence that the alcohol alters reaction time? (Use  $\alpha = .05$ .)**

### **One-sample data**

- **Suppose instead of paired data  $(X_i, Y_i)$ , we had a single sample  $Y_1, Y_2, \dots, Y_n$ .**
- **Suppose we wish to test whether the median of  $Y$  equals some constant number  $m$ .**
- **Our set of hypotheses would be one of:**

- We can let all  $X_i = m$ , i.e., form the pairs  $(m, Y_1), (m, Y_2), \dots, (m, Y_n)$  and carry out the signed-rank test as before.
- If  $T^+$  is large, it is evidence that the  $Y_i$ 's tend to be \_\_\_\_\_ than  $m$ .
- If  $T^+$  is small, it is evidence that the  $Y_i$ 's tend to be \_\_\_\_\_ than  $m$ .

**Example 3:** Pollution measurements were taken on a site downstream from an industrial plant. Is there evidence that the median pollution level is less than 34 ppm?

### Confidence Interval for the Median Difference

- Consider the sample  $D_1, D_2, \dots, D_n$  (where  $D_i = Y_i - X_i$ , or where the  $D_i$ 's are merely a single sample).
- If the  $D_i$ 's are mutually independent, have the same median, have symmetric distributions, and are measured on at least an interval scale, the following method produces a confidence interval for the median of  $D_i$ 's with coverage probability at least  $1 - \alpha$ .

- **Consider all possible averages**

**(including when  $i = j$ )**

- **Find  $\hat{\mu}_{i,j}$  from Table A12. Then the**

**form the lower and upper bounds for the confidence interval for the median difference.**

- **This is fairly computationally intensive, and the CI is best found using software: See the `WSRCi` function on the course web page.**

**Example 2 again: Find a 95% CI for the median difference in reaction time with alcohol and without.**

**Example 3 again: Find a 95% CI for the median pollution level.**



## Some Notes

- While the t-test (or paired t-test) requires \_\_\_\_\_ distributed data, the Wilcoxon signed-rank test requires only \_\_\_\_\_ data.
- Therefore the signed-rank test may be used with data having various types of distribution:
- In particular, the t-test lacks power when \_\_\_\_\_ exist in the data.
- The sign test is appropriate for any data measured on an ordinal or stronger scale.
- If the  $Y_i$  and  $X_i$  have identical symmetric distributions except for a shift in the means, the Wilcoxon signed-rank test has AT WORST an A.R.E. of relative to the t-test.
- At BEST, the Wilcoxon signed-rank test has an A.R.E. of \_\_\_\_\_ relative to the t-test.

## Efficiency of the Signed Rank Test

**Population**    **A.R.E.(signed-rank vs. t)**    **A.R.E.(signed-rank vs. sign)**

**Normal**

**Uniform (light tails)**

**Double exponential  
(heavy tails)**

- **However, note that the A.R.E. is specific to large samples.**
- **For small samples from the double exponential distribution, the signed-rank test is \_\_\_\_\_ efficient than the sign test.**