# Data Analysis and Statistical Methods
# Statistics 651

http://www.stat.tamu.edu/~suhasini/teaching.html

Lecture 22 (MWF) Nonparametric tests

Suhasini Subba Rao

# Review of comparison methods

- Independent samples: This when we have two completely independent samples drawn from two populations, and we want to compare their population means.
  - **The independent sample t-test** If the sample sizes are relatively large and there are few outliers we can test equality of the means (one-sided versions) using the independent sample t-test.
  - **Wilcoxon sum rank test** If the sample sizes are small and there appears to outliers we can test equility of their distributions (means) using the Wilcoxon sum rank test (we do not require normality of the sample means only that the distributions of both populations have identical shapes).

- Matched/paired samples: This is when we have 'paired' observations, each of the pairs coming from different populations. Eg. the running

time of a runner at high and low altitudes. In this case the pairs are dependent, we cannot use the above tests because they are dependent. If unsure of a pairing we can check for dependence by plotting the pairs against each other (eg. high altitude time against low altitude time for each individual). If a pairing seems reasonable err on the cautious side and use a paired t-test.

- **The paired t-test** If sample size is relatively large and their are not many outliers use a paired t-test.
- **The Sign test** Today's class.
- **The Wilcoxon sign rank test** Today's class.

# Example 1: Runners at altitude

Runners were compared at a high and low altitude. For each runner, the running time was measured at a high altitude and then again at a low altitude. 12 runners were used.

| Runner | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| High | 9.4 | 9.8 | 9.9 | 10.3 | 8.9 | 8.8 | 9.8 | 8.2 | 9.4 | 9.9 | 12.2 | 9.3 |
| Low | 8.7 | 7.9 | 8.3 | 8.4 | 9.2 | 9.1 | 8.2 | 8.1 | 8.9 | 8.2 | 8.9 | 7.5 |

Do you think altitude has an effect on running time? Let $\mu_Y$ denote the mean time at a low altitude and $\mu_X$ denote the mean time at a high altitude. Use $\alpha = 0.05$.

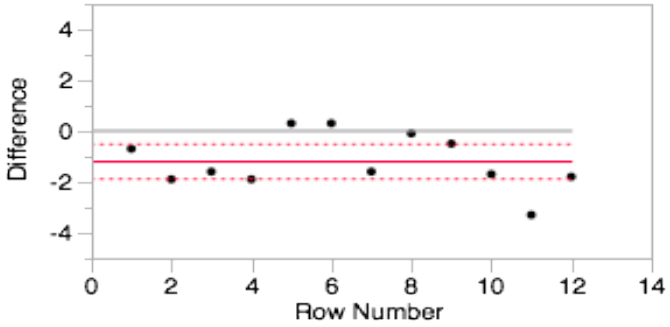
We want to test $H_0 : \mu_Y - \mu_X = \mu_d \geq 0$ against the alternative $H_A : \mu_Y - \mu_X = \mu_d < 0$.

# Solution using the matched paired t-test

The JMP output for the matched t-test is given below.

**Matched Pairs**

**Difference: Low-High**




| | | | |
|---|---|---|---|
| Low | 8.45 | t-Ratio | -3.88501 |
| High | 9.65833 | DF | 11 |
| Mean Difference | -1.2083 | Prob > \|t\| | 0.0025* |
| Std Error | 0.31102 | Prob > t | 0.9987 |
| Upper 95% | -0.5238 | Prob < t | 0.0013* |
| Lower 95% | -1.8929 | | |
| N | 12 | | |
| Correlation | 0.07937 | | |

- The t-transform is

$$t = \frac{-1.21 - 0}{1.16/\sqrt{10}} = \frac{-1.21}{0.566} = -2.13.$$

- Looking up the t-tables with 9dfs at 5% level gives -1.79. Since -2.13 is less than -1.79, the area to the left of -2.13 is less than 5%, therefore the p-value is less than 5% (it is a one-sided test) and we can reject the null.

- Naturally, this is consistent with the JMP output which was giving a p-value of 0.13%.

# Assumptions to do the matched t-test

- Let us return to the differences in the running example:

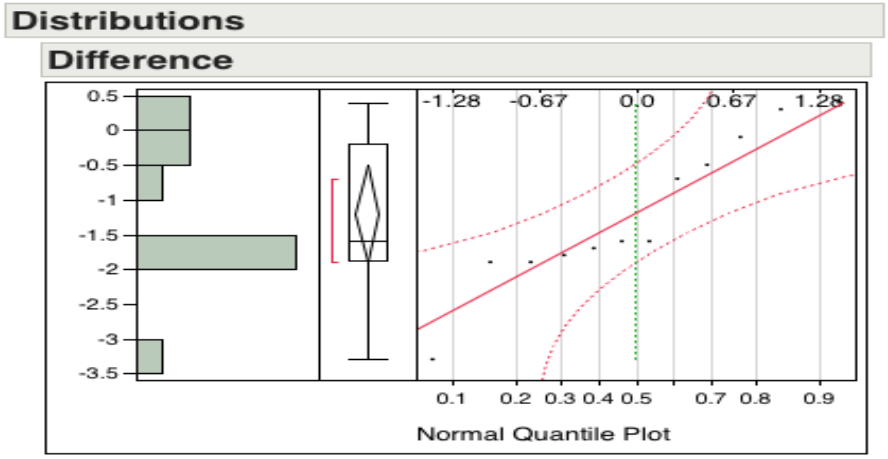
| Runner | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| High | 9.4 | 9.8 | 9.9 | 10.3 | 8.9 | 8.8 | 9.8 | 8.2 | 9.4 | 9.9 | 12.2 | 9.3 |
| Low | 8.7 | 7.9 | 8.3 | 8.4 | 9.2 | 9.1 | 8.2 | 8.1 | 8.9 | 8.2 | 8.9 | 7.5 |
| D=L - H | -0.7 | -1.9 | -1.6 | -1.9 | 0.3 | 0.3 | -1.6 | -0.1 | -0.5 | -1.7 | -3.3 | -1.8 |

- Our main assumption to do the matched paired t-test (besides independence between the pairings) is that the sample mean of the differences is close to normal. Equivalently, there aren't any huge outliers that will have an undue influence on the test.

- If the sample size is relatively large (thanks to the central limit theorem kicking in) the sample mean of the difference will be close to normal.

However, if the sample size is small and the $actual$ distribution of the data is not close to normal, then the matched paired t-test will not give accurate p-values (just like the results of an independent sample t-test when the sample size is small - recall the iron example in lecture 19).

# QQplot of the differences in the running data




**Summary Statistics**

| | |
|---|---|
| Mean | -1.2 |
| Std Dev | 1.0904545 |
| Std Err Mean | 0.3147871 |
| Upper 95% Mean | -0.507158 |
| Lower 95% Mean | -1.892842 |
| N | 12 |

It is hard to tell with such a small sample size, but the data does not look that normal (there is possibly an outlier). Below we look at alternative tests that are not sensitive to outliers.

# Nonparametric test 1: The Sign Test

We first discuss the Sign test, which will motivate the Wilcoxon Sign-Rank test (which tends to have more power than the Sign-Rank test). Let us look again at the difference data:

| Runner | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 8.7 | 7.9 | 8.3 | 8.4 | 9.2 | 9.1 | 8.2 | 8.1 | 8.9 | 8.2 | 8.9 | 7.5 |
| High | 9.4 | 9.8 | 9.9 | 10.3 | 8.9 | 8.8 | 9.8 | 8.2 | 9.4 | 9.9 | 12.2 | 9.3 |
| $D_i$ | -0.7 | -1.9 | -1.6 | -1.9 | 0.3 | 0.3 | -1.6 | -0.1 | -0.5 | -1.7 | -3.3 | -1.8 |
| Sign($D_i$) | - | - | - | - | + | + | - | - | - | - | - | - |

A simple heuristic way of reasoning that running at high altitudes may increase running times is that there are 10 negative and 2 positives (note that this method is extremely robust to outliers). 10 negatives is 'a lot of negatives' if the running times are on average equal. Of course, this could be explained by random chance. But how likely will this happen by 'chance'. To understand this we make the problem more precise.

# Nonparametric test 1: The Sign test

- In the Sign test we assume that the distribution of the differences come from continuous random variables.

- In the Sign Test (if the distributions are symmetric), we test the same hypothesis as the matched paired t-test; $H_0 : \mu_L - \mu_H = \mu_d \geq 0$ vs against $H_A : \mu_L - \mu_H = \mu_d < 0$. For non-symmetric distributions the hypotheses are based on the median of the differences $H_0 : M_d \geq 0$ vs $H_A : M_d < 0$.

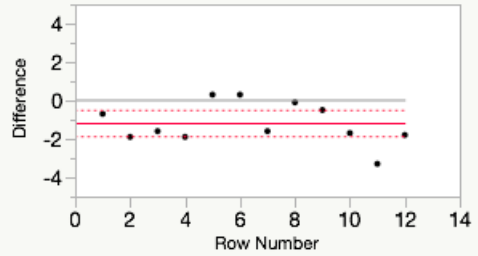
- Let $M$ denote the number of runners who run faster at a high altitude than low altitude (the number of positive signs).

- If the alternative were true, $M$ would be less than we would expect (since the mean is negative, pulling the data to the left of zero). We focus on $M$, because under the null it would be less than we expect.

- The p-value is the probability of obtaining $M$ positives out of $n$ or less when the null is true (the median is zero). The p-value is given in the JMP output.

**Matched Pairs**

**Difference: Low-High**




| | | | |
|---|---|---|---|
| Low | 8.45 | t-Ratio | -3.88501 |
| High | 9.65833 | DF | 11 |
| Mean Difference | -1.2083 | Prob > \|t\| | 0.0025* |
| Std Error | 0.31102 | Prob > t | 0.9987 |
| Upper 95% | -0.5238 | Prob < t | 0.0013* |
| Lower 95% | -1.8929 | | |
| N | 12 | | |
| Correlation | 0.07937 | | |

**Sign Test**

| | Low-High |
|---|---|
| Test Statistic M | -4.000 |
| Prob ≥ \|M\| | 0.0386* |
| Prob ≥ M | 0.9968 |
| Prob ≤ M | 0.0193* |

- The p-value corresponds to Prod$\leq$M in the JMP output, which is 1.93%.

- The plot in the output, gives the differences, which allows one to count the number of points/differences above and below the zero.

# Advantages and disadvantages of the sign test

A main disadvantage of the Sign test is that it does not take into account the magnitude of the difference. This has the advantage that it is very robust to outliers. But it is not sensitive to the alternative (it is difficult to reject the null even when the alternative is true).

- In the next few slides we demonstrate how this may be problematic.

- In the next few slides we apply the sign test to the Friday 13th data set considered in Lecture 21. We show that the sign test is not sensitive enough to reject the null.

# The Sign test and the Friday 13th Data

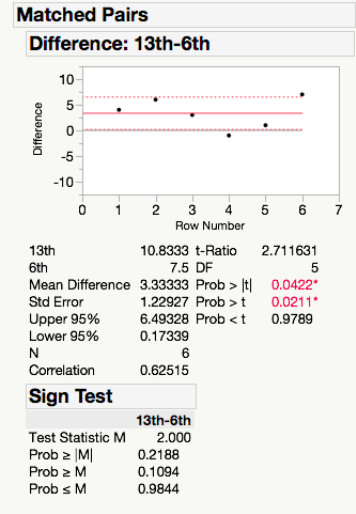
| | | | | | | |
|---|---|---|---|---|---|---|
| 6th | 9 | 6 | 11 | 11 | 3 | 5 |
| 13th | 13 | 12 | 14 | 10 | 4 | 12 |
| diff | 4 | 6 | 3 | -1 | 1 | 7 |
| Sign | + | + | + | - | + | + |

- We want to see if in general more accidents happen on the 13th, this means testing $H_0 : \mu_d \leq 0$ vs. $H_A : \mu_d > 0$

- Under under the alternative there will be more positive signs than we expect.

- For this data set there are $M = 5$ positive signs.

# The p-value using the sign test of 13th data



**Matched Pairs**

**Difference: 13th-6th**



| | | | |
|---|---|---|---|
| 13th | 10.8333 | t-Ratio | 2.711631 |
| 6th | 7.5 | DF | 5 |
| Mean Difference | 3.33333 | Prob > \|t\| | 0.0422* |
| Std Error | 1.22927 | Prob > t | 0.0211* |
| Upper 95% | 6.49328 | Prob < t | 0.9789 |
| Lower 95% | 0.17339 | | |
| N | 6 | | |
| Correlation | 0.62515 | | |

**Sign Test**

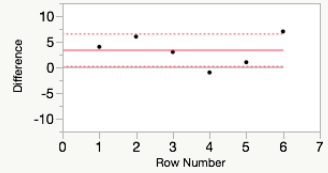
| | 13th-6th |
|---|---|
| Test Statistic M | 2.000 |
| Prob ≥ \|M\| | 0.2188 |
| Prob ≥ M | 0.1094 |
| Prob ≤ M | 0.9844 |

The p-value is over 10%. Thus using the sign test we see that there is no evidence in the data (at the 5% level) to suggest that more accidents happen on the 13th than 6th. This may be because the sign test is very conservative. And does not easily reject the null.

# Problems with the sign test

**Matched Pairs**

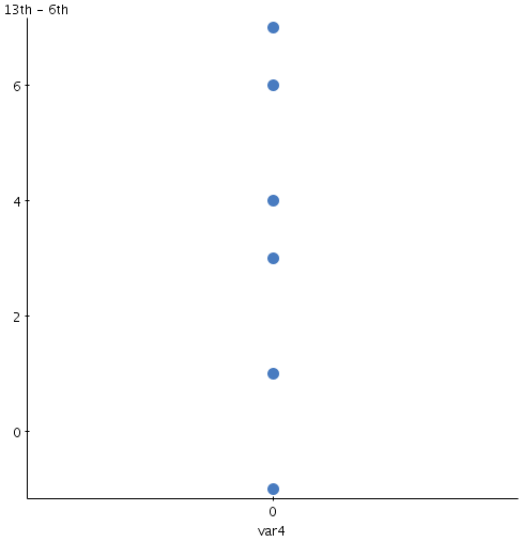
**Difference: 13th-6th**




| | | | |
|---|---|---|---|
| 13th | 10.8333 | t-Ratio | 2.711631 |
| 6th | 7.5 | DF | 5 |
| Mean Difference | 3.33333 | Prob > \|t\| | 0.0422* |
| Std Error | 1.22927 | Prob > t | 0.0211* |
| Upper 95% | 6.49328 | Prob < t | 0.9789 |
| Lower 95% | 0.17339 | | |
| N | 6 | | |
| Correlation | 0.62515 | | |

**Sign Test**

| | 13th-6th |
|---|---|
| Test Statistic M | 2.000 |
| Prob ≥ \|M\| | 0.2188 |
| Prob ≥ M | 0.1094 |
| Prob ≤ M | 0.9844 |




To understand why the sign-test could not reject the null we plot the differences. And pay special attention to the magnitude of the differences.

- It is true, there is one negative difference, but the difference is very close to zero. It is a "small" as compare with the positive differences.

- The sign test does not take into account the magnitude of the numbers just the signs. This makes it very robust against outliers, but makes it very difficult to obtain a small p-value (and thus reject the null).

- If we took the magnitude of the signs into account we return to the paired t-test, which is not robust against outliers.

- We require a compromise. This is the Wilcoxon sign rank test - it takes both the signs and the ranking of the differences into account.

# Nonparametric test 2: The Wilcoxon sign rank test

- For the rest of this lecture we concentrate on Wilcoxon Sign-Rank test, which is similar to the sign-test as it is based on counting signs together with the $rankings$ of the differences.

- The one disadvantage of the Wilcoxon sign-rank test over the Sign test is that more assumptions are required.

- The Wilcoxon Sign-rank test requires the underlying population distribution of the differences to be symmetric continuous random variables. This is difficult to check from the data. But your knowledge of how the data was generated may give some insight into whether this assumption holds true.

- In the next few slides we motivate the Wilcoxon Sign rank test through the Friday 13th data.

# The Wilcoxon Sign-Rank test (uses Table 6)

- We do not require normality of $D_i$, but the distribution of the differences $D_i$ **must be symmetric** about the median.

- Recall if a distribution is symmetric, then the mean and median are the same. Let $\mu_d$ denote the mean/median of the differences.

- For the nonparametric test the hypotheses are:
  - Two sided test $H_0 : \mu_d = 0$ vs. $H_A : \mu_d \neq 0$.
  - One sided test (pointing right). $H_0 : \mu_d \leq 0$ vs. $H_A : \mu_d > 0$.
  - One sided test (pointing left). $H_0 : \mu_d \geq$ vs. $H_A : \mu_d < 0$.

# General Recipe

On first reading, it may be easier to skip these slides and follow the Friday 13th example later in the slides.

- Calculate the difference between the pairs of samples $D_i = X_i - Y_i$.
- Delete all zero values and let $n^*$ be the number of non-zero values.
- List the absolute values of the differences.
- Rank from the smallest to largest all the numbers (after the sign is ignored). Same numbers are given joint place (see Friday 13th example, below).
- To each rank, allocate a sign (negative or positive), depending on whether the difference was previously given as negative or positive.

- Add all the negative ranks together, denote this $T_-$.
- Add all the positive ranks together, denote this $T_+$.
- Selecting $T_-$ or $T_+$ for the test We choose the rank which we expect to be 'smallest' if the alternative were true (see the plots on the right).

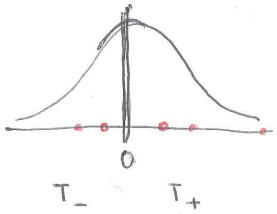
Case 1

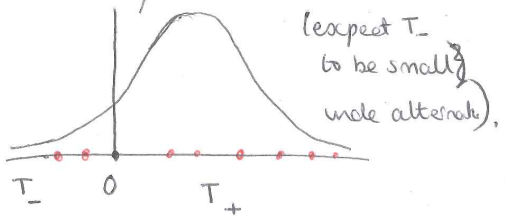
$H_0: \mu_d = 0$ vs $H_a: \mu_d \neq 0$. (Two-sided test)

Select. $T = \min(T_+, T_-)$. Reject null if T < Critical value (from Table 6)

Case 2

$H_0: \mu_d \leq 0$ vs $H_a: \mu_d > 0$ (one-sided)

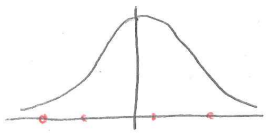


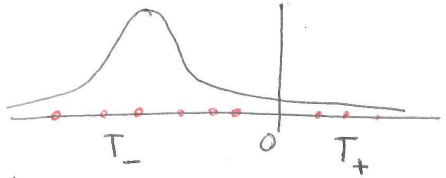




(expect $T_-$ to be small under alternative)

Case 3

$H_0: \mu_d \geq 0$ $H_A: \mu_d < 0$ (one-sided)






(expect $T_+$ to be small $T_-$ under the alternative)

# Summary of above plot

- If you are doing a two-sided test, choose the smaller of $T_-$ and $T_+$, and label this $T$ ($T = \min(T_-, T_+)$).

- If you are doing a one-sided test with $H_0 : \mu_d \leq 0$ against $H_A : \mu_d > 0$, this means you are trying to see if $T_-$ is too small (since the alternative is pointing to the right, if the alternative is true we will have very few negative ranks). You need to use $T = T_-$.

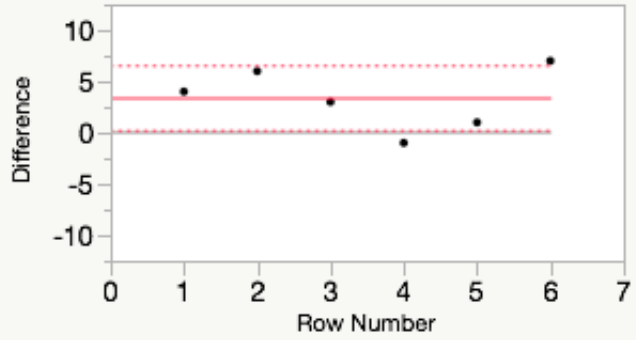
- If you are doing a one-sided test with $H_0 : \mu_d \geq 0$ against $H_A : \mu_d < 0$, this means you are trying to see if $T_+$ is too small (since the alternative is pointing to the left, if the alternative is true we will have very few positive ranks). You need to use $T = T_+$.

- Look up Table 6. The columns are the sample size of the pairs, depending on whether you are doing a two-sided or one-sided test and $\alpha$, is select the appropriate 'value'. If $T$ is less than this 'value', there is enough evidence to reject the null.

# Wilcoxon sign-rank test for the Friday 13th data

**Matched Pairs**

**Difference: 13th-6th**




| | | | |
|---|---|---|---|
| 13th | 10.8333 | t-Ratio | 2.711631 |
| 6th | 7.5 | DF | 5 |
| Mean Difference | 3.33333 | Prob > \|t\| | 0.0422* |
| Std Error | 1.22927 | Prob > t | 0.0211* |
| Upper 95% | 6.49328 | Prob < t | 0.9789 |
| Lower 95% | 0.17339 | | |
| N | 6 | | |
| Correlation | 0.62515 | | |

- Rank the absolute values of data from the smallest number to the largest. For example, we see that the deviation of the 4th and 5th row number is the least from zero and both have the same magnitude, they are given the joint lowest rank of 1.5 = (Rank 1+Rank 2)/2.

- It is done by hand below.

| 6th | 13th | Difference | sign | Abs. Rank | $+$ | - |
|---|---|---|---|---|---|---|
| 9 | 13 | 4 | $+$ | 4 | 4 | |
| 6 | 12 | 6 | $+$ | 5 | 5 | |
| 11 | 14 | 3 | $+$ | 3 | 3 | |
| 11 | 10 | -1 | - | 1.5 | | 1.5 |
| 3 | 4 | 1 | $+$ | 1.5 | 1.5 | |
| 5 | 12 | 7 | $+$ | 6 | 6 | |
| Total | | | | | $T_+ = 19.5$ | $T_- = 1.5$ |

- The test is $H_0 : \mu_d \leq 0$ vs. $H_A : \mu_d > 0$ (the mean number of differences between the Friday 13th and 6th is greater than zero).

# The Wilcoxon sign rank test and Table 6

- $T_+ =$ sum of positive ranks ($T_+ = 19.5$).
- $T_- =$ sum of negative ranks ($T_- = 1.5$).
- We choose the test statistic $T$ as the rank which we expect to the smallest under the alternative.
- Under the null, we would expect the sum of negative ranks to be small, therefore we choose the negative rank $T = T_- = 1.5$.
- Look up Table 6, for a particular $p = \alpha$ and $n^*$ (so in our case we use a one-sided test with $p = 0.05$ and $n^* = 6$), the value in the table is 2. Since $T = 1.5 < 2$, we can determine that this value is 'small' and we can reject $H_0$ at the 5% level and say there is evidence that more accidents tend to happen on 13th than 6th.

# Using Table 6 to do the test

| One-Sided | Two-Sided | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ |
|---|---|---|---|---|---|---|
| $p = .1$ | $p = .2$ | 2 | 3 | 5 | 8 | 10 |
| $p = .05$ | $p = .1$ | 0 | 2 | 3 | 5 | 8 |
| $p = .025$ | $p = .05$ | | 0 | 2 | 3 | 5 |
| $p = .01$ | $p = .02$ | | | 0 | 1 | 3 |
| $p = .005$ | $p = .01$ | | | | 0 | 1 |
| $p = .0025$ | $p = .005$ | | | | | 0 |
| $p = .001$ | $p = .002$ | | | | | |

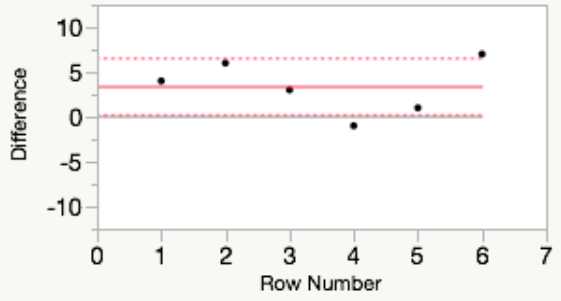
- We observe that for the one-sided test with $n = 2$, to reject the null at the 5% level, we require $T < 2$. For the Friday 13th Data set $T = 1.5 < T$. Therefore the p-value is less than 5%/

- This means that by using the Wilcoxon Sign Rank test we can reject the null (just as we did with the paired t-test).

# The JMP output for Friday 13th data

**Matched Pairs**

**Difference: 13th-6th**




| | | | |
|---|---|---|---|
| 13th | 10.8333 | t-Ratio | 2.711631 |
| 6th | 7.5 | DF | 5 |
| Mean Difference | 3.33333 | Prob > \|t\| | 0.0422* |
| Std Error | 1.22927 | Prob > t | 0.0211* |
| Upper 95% | 6.49328 | Prob < t | 0.9789 |
| Lower 95% | 0.17339 | | |
| N | 6 | | |
| Correlation | 0.62515 | | |

**Wilcoxon Signed Rank**

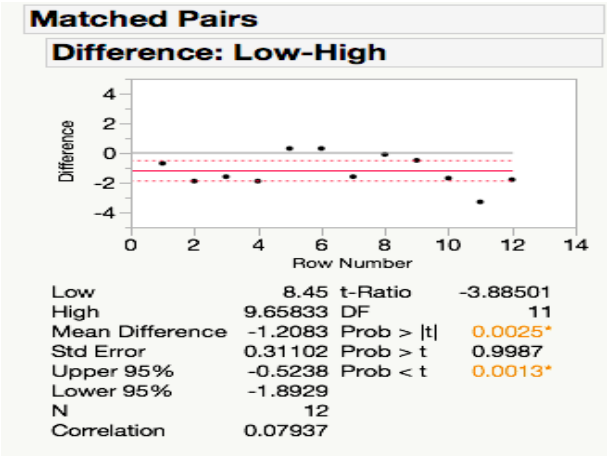
| | 13th-6th |
|---|---|
| Test Statistic S | 9.000 |
| Prob>\|S\| | 0.0938 |
| Prob>S | 0.0469* |
| Prob<S | 0.9531 |

The JMP output for the Wilcoxon test is a little different from what we have. It uses the some normal approximations, which only hold for large sample sizes. In this course we ignore the JMP output and do the calculatation by hand - using the tables to find the critical values.

# Example: Runners at altitude

Runners were compared at a high and low altitude. For each runner, the running time was measure at a high altitude and then again at a low altitude. 12 runners were used.

| Runner | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| High | 9.4 | 9.8 | 9.9 | 10.3 | 8.9 | 8.8 | 9.8 | 8.2 | 9.4 | 9.9 | 12.2 | 9.3 |
| Low | 8.7 | 7.9 | 8.3 | 8.4 | 9.2 | 9.1 | 8.2 | 8.1 | 8.9 | 8.2 | 8.9 | 7.5 |




| | | | |
|---|---|---|---|
| Low | 8.45 | t-Ratio | -3.88501 |
| High | 9.65833 | DF | 11 |
| Mean Difference | -1.2083 | Prob > \|t\| | 0.0025* |
| Std Error | 0.31102 | Prob > t | 0.9987 |
| Upper 95% | -0.5238 | Prob < t | 0.0013* |
| Lower 95% | -1.8929 | | |
| N | 12 | | |
| Correlation | 0.07937 | | |

We use the Wilcoxon sign rank test to test the hypothesis that running at high altitudes increases running times.

# Solution

- It is clear that the each pair of observations are dependent, therefore we evaluate differences

| Runner | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L-H | -0.7 | -1.9 | -1.6 | -1.9 | 0.3 | 0.3 | -1.6 | -0.1 | -0.5 | -1.7 | -3.3 | -1.8 |

- Assume that the distribution of the differences, $D_i$ = Low - High, is symmetric.

- We test $H_0 : \mu_d \geq 0$ (equivalently the mean is greater than or equal to zero) against the alternative $H_A : \mu_d < 0$ (equivalently the mean is less than zero).

# Solution: Wilcoxon Sign Rank test

It is clear that the each pair of observations are dependent, hence we need to consider the differences,

| Runner | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | SUM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L - H | -0.7 | -1.9 | -1.6 | -1.9 | 0.3 | 0.3 | -1.6 | -0.1 | -0.5 | -1.7 | -3.3 | -1.8 | |
| sign | - | - | - | - | + | + | - | - | - | - | - | - | |
| Rank | 5 | 8.5 | 6.5 | 8.5 | 2.5 | 2.5 | 6.5 | 1 | 4 | 10 | 12 | 11 | |
| $T_-$ | 5 | 8.5 | 6.5 | 8.5 | | | 6.5 | 1 | 4 | 10 | 12 | 11 | 73 |
| $T_+$ | | | | | 2.5 | 2.5 | | | | | | | 5 |

- We see that $T_- = 73$ and $T_+ = 5$.
- Since the alternative is pointing to the left we use the positive rank $T = T_+ = 5$ (the sum of ranks of the positive values, as this is likely to the small under the alternative) in the test.
- Look up Table 6, with n=12, use the one-sided test an $\alpha = 0.05$. We see that the VALUE $= 17$.

- Since $T_+ = 5 \leq 17$. The p-value is less than 5% for the Wilcoxon sign rank test. There is enough evidence to reject the null.

# The Assumptions for the Wilcoxon Sign test

The main assumptions to do the the Wilcoxon Sign rank test are

- The differences $D_i$ are independent of each other (there is no relationship between them).

- The distribution of the difference is symmetric $D_i$.

A similar test can be applied to the one sample case to test $H_0 : \mu = \mu_0$ against $H_A : \mu \neq \mu_0$ (or the one-sided equivalents), this test tends to be more robust than the one-sample t-test (but requires symmetry of the observations). We don't go through the details here.

# Reminder: What test to do when...

If we want to test whether two samples come from the same population against two different population:

- If there is no pairing in the data (sample sizes can be the same or different) and
  - the data is normal (check with QQplot on residual), then use the independent sample t-test.
  - the data is not normal, but the sample size is sufficient large that the distribution of the two sample means are normal, then use independent sample t-test.
  - the data not normal and sample size is small, then use the Wilcoxon rank sum (Mann-Whitney U) test.

- If there is a natural pairing across the two samples (both samples need to have the same sample size in this case)

- the data is normal, then use the paired t-test.
- the data is not normal, but the sample size is sufficient large that the distribution of the sample mean of the differences normal, then use paired t-test.
- the data not normal and sample size is small, then use a nonparametric test, such as the sign test or Wilcoxon sign rank test.

We should use the test which suits the data. Because loss of power can occur if we use the wrong test.

# The size of the sample

- Always bear in mind that the size of the sample does not matter, so long the data has been collected in a way to ensure that it is unbiased.

- The standard error will always measures the uncertainty associated with an estimator.

- The smaller the sample size the larger the standard error, which reflects the greater uncertainty in the estimator.