Wilcoxon Signed Rank Discussion

C. Durso

Wilcoxon Signed Rank Test

- Set-up
- $x_1, x_2, \ldots x_n$ is a sample from a *symmetric* distribution with mean=median=μ
  - $P(X < \mu - \xi) = P(X > \mu + \xi)$
  - $P\left(X_i = X_j\right) = 0$
- Null hypothesis: $\mu = \mu_0$

The Test Statistic

- Rank $|x_1 - \mu_0|, |x_2 - \mu_0|, \ldots |x_n - \mu_0|$
- Let $r_1, r_2, \ldots r_n$ be the associated ranks
- Let $s_i = 1$ if $x_i - \mu_0 > 0$ and let $s_i = 0$ if $x_i - \mu_0 < 0$
- The test statistic w is $\sum_{i=1}^{n} s_i r_i$, the sum of the absolute ranks of the positive values of $x_i - \mu_0$

Alternate Version

- Equivalently, some implementations set $sgn_i = 1$ if $x_i - \mu_0 > 0$ and set $sgn_i = -1$ if $x_i - \mu_0 < 0$ then compute $\widetilde{w} = \sum_{i=1}^{n} sgn_i r_i$
- Note $\widetilde{w} = 2w - \sum_{i=1}^{n} r_i = 2w - \frac{n(n+1)}{2}$

Random Variable

- Let W be a random variable with the distribution of the test statistic w under the null hypothesis H_0.
- The expected value for W under the null hypothesis is $\sum_{k=1}^{n} k(P(s_i = 1)) = \sum_{k=1}^{n} \frac{k}{2} = \frac{n(n+1)}{4}$.

Inference

- If w is greater than this, the value $2P(W \geq w)$ is the probability of a value as extreme as w under H_0.
- If w is less than this, the value $2P(W \leq w)$ is the probability of a value as extreme as w under H_0.

Sign Test Discussion

C. Durso

Sign Test

Set-up

- $x_1, x_2, \ldots x_n$ is a sample from a distribution with median=μ
- Null hypothesis: $\mu = \mu_0$
- Only need information on truth values of $x_i - \mu_0 > 0$ and $x_i - \mu_0 < 0$

Common Application

- Paired measurements, pre- and post-intervention, $u_1, u_2, \ldots u_n$ and $v_1, v_2, \ldots v_n$
- Null hypothesis to capture “intervention has no effect”:
  - The events $u_i < v_i$ and $u_i > v_i$ are equally likely for all $i \in 1, \ldots n$
  - The events $u_i < v_i$ and $u_j < v_j$ are independent if $i \neq j$
- Apply the sign test to $x_1 = v_1 - u_1, x_2 = v_2 - u_2, \ldots x_n = v_n - u_n$

The Test Statistic

- Consider $(x_1 - \mu_0), (x_2 - \mu_0), \ldots (x_n - \mu_0)$
- Let $s_i = 1$ if $x_i - \mu_0 > 0$ and let $s_i = 0$ if $x_i - \mu_0 < 0$
- The test statistic w is $\sum_{i=1}^{n} s_i$, the count of the positive values of $x_i - \mu_0$

Random Variable

- The random variable W that has the distribution of w under the null hypothesis H_0 has the binomial distribution with size equal to the number m of non-zero values of $x_i - \mu_0$ and the probability parameter equal to $\frac{1}{2}$.
- The expected value of W is $\frac{m}{2}$.
- Let F be the cumulative distribution of $Binomial(m, 0.5)$.

Inference, Two-Sided

- If $w \leq \frac{m}{2}$, then $2F(w)$ is the probability of a value as extreme as w under H_0.
- If $w > \frac{m}{2}$, then $2[1 - F(w-1)] = 2(P\ (W \geq w))$ is the probability of a value as extreme as w under H_0.

Inference, One-Sided

- If domain knowledge implies $\mu \le \mu_0$ or $P(X > 0) \le 0.5$ and $w \le \frac{m}{2}$, then $F(w)$ is the probability of a value as extreme as w under H_0.
- If domain knowledge implies $\mu \ge \mu_0$ or $P(X > 0) \ge 0.5$ and $w \ge \frac{m}{2}$, then $1 - F(w-1) = P(W \ge w)$ is the probability of a value as extreme as w under H_0.

Two-Sample t-test, Equal Variances

C. Durso

Application

- Given independent samples of a continuous measurement from two populations, examine whether there is evidence that the two populations have different means for that measurement.
- Estimate how much the population measurement means differ.

Set-up

- Two sets of numerical values, $\{x_1, x_2, \ldots x_{n_X}\}$ and $\{y_1, y_2, \ldots y_{n_Y}\}$
- Respectively, samples from $normal(\mu_X, \sigma^2)$ and $normal(\mu_Y, \sigma^2)$
- Goals:
  - Confidence interval for $\mu_X - \mu_Y$
  - Test of null hypothesis $\mu_X = \mu_Y$

Statistic Terms

- Represent the mean of the x-values by $\bar{x}$ and the mean of the y-values by $\bar{y}$
- Approximate σ^2 by $\frac{\sum_{i=1}^{n_X}(x_i-\bar{x})^2+\sum_{i=1}^{n_Y}(y_i-\bar{y})^2}{n_X+n_Y-2}$
  - Call this S^2
- The variance of the random variable $\bar{X}-\bar{Y}$ equals $\frac{\sigma^2}{n_X}+\frac{\sigma^2}{n_Y}$. Approximate it by

$$S^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)$$

Statistic for Difference of Means

Theorem

Given two sets of numerical values, $\{x_1, x_2, \dots x_{n_X}\}$ and $\{y_1, y_2, \dots y_{n_Y}\}$, iid samples from $normal(\mu_X, \sigma^2)$ and $normal(\mu_Y, \sigma^2)$ respectively, the statistic

$$\frac{\bar{x} - \bar{y}}{\sqrt{S^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}}$$

has a Student's t distribution with $n_X + n_Y - 2$ degrees of freedom

Confidence Interval

The corresponding $100(1-p)\%$ confidence interval for $\mu_X - \mu_Y$ is

$$\left(\bar{x} - \bar{y} - t_{\frac{p}{2}}\sqrt{S^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}, \bar{x} - \bar{y} + t_{\frac{p}{2}}\sqrt{S^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}\right)$$

where $t_{\frac{p}{2}}$ satisfies the property that the probability of the event $\left(-\infty, -t_{\frac{p}{2}}\right)$ equals $\frac{p}{2}$ for a random variable with the Student's t distribution with $n_X + n_Y - 2$ degrees of freedom.

Hypothesis Test

To obtain the p-value corresponding to a two-sided test of the null hypothesis that $\{x_1, x_2, \ldots x_{n_X}\}$ and $\{y_1, y_2, \ldots y_{n_Y}\}$ are samples from $normal(\mu, \sigma^2)$, evaluate $2P\left(T < -\left|\frac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}}\right|\right)$ where T has the Student's t distribution with $n_X + n_Y - 2$ degrees of freedom.

Considerations

- Unless there is a priori reason to believe that the variances of the two populations are equal, Welch’s test is preferred.
- The two populations must be approximately normally distributed.

F-test Discussion

C. Durso

F-distribution Basics

- The F-distribution is a two-parameter family $F(df1, df2)$. For our purposes, the key feature of the family is that, if $df1$ and $df2$ are positive integers, $F(df1, df2)$ is the random variable obtained as follows:
  - Take $S1$ to be a sum of the squares of $df1$ independent standard normal random variables.
  - Take $S2$ to be a sum of the squares of $df2$ independent standard normal random variables.
  - Set Y to be the random variable $\left(\frac{S_1}{df1}\right) / \left(\frac{S_2}{df2}\right)$.
- The random variable Y has the distribution $F(df1, df2)$.

F-test of Equality of Variance

- If $x_1, x_2, \ldots x_n$ and $y_1, y_2, \ldots y_m$ are iid samples from $normal(\mu_1, \sigma^2)$ and $normal(\mu_2, \sigma^2)$, respectively, then the statistic $\frac{\sum(x_i - \bar{x})^2}{n-1} / \frac{\sum(y_j - \bar{y})^2}{m-1}$ has the distribution $F(n-1, m-1)$.
- Thus, if $x_1, x_2, \ldots x_n$ and $y_1, y_2, \ldots y_m$ are iid samples from $normal(\mu_1, \sigma_1^2)$ and $normal(\mu_2, \sigma_2^2)$, respectively, the statistic $\frac{\sum(x_i - \bar{x})^2}{n-1} / \frac{\sum(y_j - \bar{y})^2}{m-1}$ can be used to test the null hypothesis that $\sigma_1^2 = \sigma_2^2$.
- This test can be very sensitive to non-normality of the data. Levene's test or the Brown-Forsythe test are more common in practice.

Motivation Transformation

- If $x_1, x_2, \ldots x_n$ and $y_1, y_2, \ldots y_m$ are iid samples from $normal(\mu_1, \sigma^2)$ and $normal(\mu_2, \sigma^2)$, define $w_1, w_2, \ldots w_n = \frac{x_1 - \mu_1}{\sigma}, \frac{x_2 - \mu_1}{\sigma}, \ldots \frac{x_n - \mu_1}{\sigma}$ and $z_1, z_2, \ldots z_m = \frac{y_1 - \mu_2}{\sigma}, \frac{y_2 - \mu_2}{\sigma}, \ldots \frac{y_m - \mu_2}{\sigma}$.
- It is now the case that $w_1, w_2, \ldots w_n$ and $z_1, z_2, \ldots z_m$ are iid samples from the standard normal distribution.

Motivation Equality of Statistics

- The statistic $\frac{\sum(x_i-\bar{x})^2}{n-1} / \frac{\sum(y_j-\bar{y})^2}{m-1}$ equals the version using the standard normal samples: $\frac{\sum(w_i-\bar{w})^2}{n-1} / \frac{\sum(z_j-\bar{z})^2}{m-1}$.
- The μ_i's and the σ^2's cancel.

Motivation Simplification

- It turns out that if $w_1, w_2, \dots w_n$ is a sample from the standard normal distribution, then $\Sigma_{i=1}^{n}(w_i - \overline{w})^2$ is a sum of $n-1$ squares of $n-1$ iid values drawn from $normal(0,1)$.
- To illustrate, consider $n = 2$:

$$(w_1 - \overline{w})^2 + (w_2 - \overline{w})^2 = \left(w_1 - \frac{w_1 + w_2}{2}\right)^2 + \left(w_2 - \frac{w_1 + w_2}{2}\right)^2$$

$$= \left(\frac{w_1 - w_2}{2}\right)^2 + \left(\frac{w_2 - w_1}{2}\right)^2 = 2\left(\frac{w_1 - w_2}{2}\right)^2$$

$$= \left(\frac{w_1 - w_2}{\sqrt{2}}\right)^2$$

- This last term is the square of $n - 1 = 1$ sample(s) from the standard normal distribution.

Welch’s Two-Sample t-test

C. Durso

Application

- As in the case of equal variances, given two independent samples of a continuous measurement from two populations, examine whether there is evidence that the two populations have different means for that measurement.
- Estimate how much the population measurement means differ.

Set-up

- Two sets of numerical values, $\{x_1, x_2, \dots x_{n_X}\}$ and $\{y_1, y_2, \dots y_{n_Y}\}$
- Respectively, samples from $normal(\mu_X, \sigma_X^2)$ and $normal(\mu_Y, \sigma_Y^2)$
- Goals:
  - Confidence interval for $\mu_X - \mu_Y$
  - Test of null hypothesis $\mu_X = \mu_Y$

Behrens-Fisher Problem

- The problem of inference about the difference in means of two populations with possibly different normal distributions based on moderate-sized samples from each is called the Behrens-Fisher problem.
- There is no definitive solution.
- Welch’s t-test approximation is one approach.

Statistic Terms

- Represent the mean of the x-values by $\bar{x}$ and the mean of the y-values by $\bar{y}$.
- Approximate σ_X^2 by $\frac{\sum_{i=1}^{n_X}(x_i-\bar{x})^2}{n_X-1}$.
  - Call this S_X^2
- Approximate σ_Y^2 by $\frac{\sum_{i=1}^{n_Y}(y_i-\bar{y})^2}{n_Y-1}$.
  - Call this S_Y^2
- The variance of the random variable $\bar{X}-\bar{Y}$ equals $\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}$. Approximate it by $\left(\frac{S_X^2}{n_X}+\frac{S_Y^2}{n_Y}\right)$.

Approximate Degrees of Freedom

We will approximate the distribution of $\frac{\bar{x}-\bar{y}}{\sqrt{\left(\frac{s_X^2}{n_X}+\frac{s_Y^2}{n_Y}\right)}}$ by a distribution with

$$\nu = \frac{\left(\frac{s_X^2}{n_X}+\frac{s_Y^2}{n_Y}\right)^2}{\left(\frac{s_X^2}{n_X}\right)^2\left(\frac{1}{n_X-1}\right)+\left(\frac{s_Y^2}{n_Y}\right)^2\left(\frac{1}{n_Y-1}\right)}$$ degrees

of freedom.

Statistic for Difference of Means

Approximation

Given two sets of numerical values, $\{x_1, x_2, \dots x_{n_X}\}$ and $\{y_1, y_2, \dots y_{n_Y}\}$, iid samples from $normal(\mu_X, \sigma^2)$ and $normal(\mu_Y, \sigma^2)$, respectively, the statistic

$$\frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)}}$$

is approximately distributed as a Student's t distribution with ν degrees of freedom.

Confidence Interval

The corresponding $100(1-p)\%$ confidence interval for $\mu_X - \mu_Y$ is

$$\left(\bar{x} - \bar{y} - t_{\frac{p}{2}}\sqrt{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)}, \bar{x} - \bar{y} + t_{\frac{p}{2}}\sqrt{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)}\right)$$

where $t_{\frac{p}{2}}$ satisfies the property that the probability of the event $\left(-\infty, -t_{\frac{p}{2}}\right)$ equals $\frac{p}{2}$ for a random variable with the Student's t distribution with ν degrees of freedom.

Hypothesis Test

The p-value for a two-sided test of the null hypothesis that $\{x_1, x_2, \ldots x_{n_X}\}$ and $\{y_1, y_2, \ldots y_{n_Y}\}$ are samples from normal populations with equal means is

$2P\left(T < -\left|\frac{\bar{x}-\bar{y}}{\sqrt{\left(\frac{S_X^2}{n_X}+\frac{S_Y^2}{n_Y}\right)}}\right|\right)$ where T has the Student's t distribution with ν degrees of freedom.

Considerations

- Unless there is a prior reason to believe that the variances of the two populations are equal, Welch's test is preferred.
- The two populations must be approximately normally distributed.