Nonparametric Location Tests: One-Sample

Nathaniel E. Helwig

Assistant Professor of Psychology and Statistics
University of Minnesota (Twin Cities)

Updated 04-Jan-2017

Copyright

Outline of Notes

1) Background Information:
- Hypothesis testing
- Location tests
- Order and rank statistics
- One-sample problem

2) Signed Rank Test (Wilcoxon):
- Overview
- Hypothesis testing
- Estimating location
- Confidence intervals

3) Sign Test (Fisher):
- Overview
- Hypothesis testing
- Estimating location
- Confidence intervals

4) Some Considerations:
- Choosing a location test
- Univariate symmetry
- Bivariate symmetry

Background Information

Neyman-Pearson Hypothesis Testing Procedure

Hypothesis testing procedure (by Jerzy Neyman & Egon Pearson[1]):

(a) Start with a null and alternative hypothesis (H_0 and H_1) about θ
(b) Calculate some test statistic T from the observed data
(c) Calculate p-value; i.e., probability of observing a test statistic as or more extreme than T under the assumption H_0 is true
(d) Reject H_0 if the p-value is below some user-determined threshold

Typically we assume observed data are from some known probability distribution (e.g., Normal, t, Poisson, binomial, etc.).

[1] Egon Pearson was the son of Karl Pearson (very influential statistician).

Confidence Intervals

In addition to testing H_0, we may want to know how confident we can be in our estimate of the unknown population parameter θ.

A symmetric $100(1-\alpha)\%$ confidence interval (CI) has the form:

$$\hat{\theta} \pm T^*_{1-\alpha/2}\sigma_{\hat{\theta}}$$

where $\hat{\theta}$ is our estimate of θ, $\sigma_{\hat{\theta}}$ is the standard error of $\hat{\theta}$, and $T^*_{1-\alpha/2}$ is the critical value of the test statistic, i.e., $P(T \leq T^*_{1-\alpha/2}) = 1 - \alpha/2$.

Interpreting Confidence Intervals:

- *Correct*: through repeated samples, e.g., 99 out of 100 confidence intervals would be expected to contain true θ with $\alpha = .01$
- *Wrong*: through one sample; e.g., there is a 99% chance the confidence interval around my $\hat{\theta}$ contains the true θ (with $\alpha = .01$)

Definition of "Location Test"

Allow us to test hypotheses about mean or median of a population.

There are **one-sample** tests and **two-sample** tests.

- *One-Sample*: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$
- *Two-Sample*: $H_0 : \mu_1 - \mu_2 = \mu_0$ vs. $H_1 : \mu_1 - \mu_2 \neq \mu_0$

There are **one-sided** tests and **two-sided** tests.

- *One-Sided*: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$ or $H_1 : \mu > \mu_0$
- *Two-Sided*: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

Problems with Parametric Location Tests

Typical parametric location tests (e.g., Student's *t* tests) focus on analyzing mean differences.

Robustness: sample mean is not robust to outliers

- Consider a sample of data $x_1, \ldots, x_n$ with expectation $\mu < \infty$
- Suppose we fix $x_1, x_2, \ldots, x_{n-1}$ and let $x_n \to \infty$
- Note $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \to \infty$, i.e., one large outlier ruins sample mean

Generalizability: parametric tests are meant for particular distributions

- Assume data are from some known distribution
- Parametric inferences are invalid if assumption is wrong

Order Statistics

Given a sample of data

$$X_1,\ X_2,\ X_3, \ldots, X_n$$

from some cdf F, the order statistics are typically denoted by

$$X_{(1)},\ X_{(2)},\ X_{(3)}, \ldots, X_{(n)}$$

where $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq \cdots \leq X_{(n)}$ are the ordered data

Note that…

- the 1st order statistic $X_{(1)}$ is the sample minimum
- the n-th order statistic $X_{(n)}$ is the sample maximum

Rank Statistics

Given a sample of data

$$X_1,\ X_2,\ X_3, \ldots, X_n$$

from some cdf F, the rank statistics are typically denoted by

$$R_1,\ R_2,\ R_3, \ldots, R_n$$

where $R_i \in [1, n]$ for all $i \in \{1, \ldots, n\}$ are the data ranks

If there are no ties (i.e., if $X_i \neq X_j\ \forall i, j$), then $R_i \in \{1, \ldots, n\}$

Order and Rank Statistics: Example (No Ties)

Given a sample of data

$$X_1 = 3,\ X_2 = 12,\ X_3 = 11,\ X_4 = 18,\ X_5 = 14,\ X_6 = 10$$

from some cdf F, the order statistics are

$$X_{(1)} = 3,\ X_{(2)} = 10,\ X_{(3)} = 11,\ X_{(4)} = 12,\ X_{(5)} = 14,\ X_{(6)} = 18$$

and the ranks are given by

$$R_1 = 1,\ R_2 = 4,\ R_3 = 3,\ R_4 = 6,\ R_5 = 5,\ R_6 = 2$$

Order and Rank Statistics: Example (With Ties)

Given a sample of data

$$X_1 = 3, \; X_2 = 11, \; X_3 = 11, \; X_4 = 14, \; X_5 = 14, \; X_6 = 11$$

from some cdf F, the order statistics are

$$X_{(1)} = 3, \; X_{(2)} = 11, \; X_{(3)} = 11, \; X_{(4)} = 11, \; X_{(5)} = 14, \; X_{(6)} = 14$$

and the ranks are given by

$$R_1 = 1, \; R_2 = 3, \; R_3 = 3, \; R_4 = 5.5, \; R_5 = 5.5, \; R_6 = 3$$

This is fractional ranking where we use average ranks:

- Replace $R_i \in \{2, 3, 4\}$ with the average rank $3 = (2 + 3 + 4)/3$
- Replace $R_i \in \{5, 6\}$ with the average rank $5.5 = (5 + 6)/2$

Order and Rank Statistics: Examples (in R)

Revisit example with no ties:

```
> x = c(3,12,11,18,14,10)
> sort(x)
[1]  3 10 11 12 14 18
> rank(x)
[1] 1 4 3 6 5 2
```

Revisit example with ties:

```
> x = c(3,11,11,14,14,11)
> sort(x)
[1]  3 11 11 11 14 14
> rank(x)
[1] 1.0 3.0 3.0 5.5 5.5 3.0
```

Summation of Integers (Carl Gauss)

Carl Friedrich Gauss was a German mathematician who made amazing contributions to all areas of mathematics (including statistics).

According to legend, when Carl was in primary school (about 8 y/o) the teacher asked the class to sum together all integers from 1 to 100.

- This was supposed to occupy the students for several hours

After a few seconds, Carl wrote down the correct answer of 5050!

- Carl noticed that $1 + 100 = 101$, $2 + 99 = 101$, $3 + 98 = 101$, etc.
- There are 50 pairs that sum to 101 $\Longrightarrow \sum_{i=1}^{100} i = 5050$

General Summation Formulas

Summation of Integers: $\sum_{i=1}^{n} i = n(n+1)/2$

- From pattern noticed by Carl Gauss

Summation of Squares: $\sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6$

- Can prove using difference approach (similar to Carl Gauss idea)

These formulas relate to test statistics that we use for rank data.

Problem(s) of Interest

For the one-sample location problem, we could have:

- *Paired-replicates data*: (X_i, Y_i) are independent samples
- *One-sample data*: Z_i are independent samples

We want to make inference about:

- *Paired-replicates data*: difference in location (treatment effect)
- *One-sample data*: single population's location

Typical Assumptions

Independence assumption:

- *Paired-replicates data*: $Z_i = Y_i - X_i$ are independent samples
- *One-sample data*: Z_i are independent samples

Symmetry assumption:

- *Paired-replicates data*: $Z_i \sim F_i$ which is continuous and symmetric around θ (common median)
- *One-sample data*: $Z_i \sim F_i$ which is continuous and symmetric around θ (common median)

Wilcoxon's Signed Rank Test

Assumptions and Hypothesis

Assumes both independence and symmetry.

The null hypothesis about θ (common median) is

$$H_0 : \theta = \theta_0$$

and we could have one of three alternative hypotheses:

- One-Sided Upper-Tail: $H_1 : \theta > \theta_0$
- One-Sided Lower-Tail: $H_1 : \theta < \theta_0$
- Two-Sided: $H_1 : \theta \neq \theta_0$

Test Statistic

Let R_i for $i \in \{1, \ldots, n\}$ denote the ranks of $|Z_i - \theta_0|$.

Defining the indicator variable

$$\psi_i = \begin{cases} 1 & \text{if } Z_i - \theta_0 > 0 \\ 0 & \text{if } Z_i - \theta_0 < 0 \end{cases}$$

the Wilcoxon signed rank test statistic T^+ is defined as

$$T^+ = \sum_{i=1}^{n} R_i \psi_i$$

where $R_i \psi_i$ is the positive signed rank of $Z_i - \theta_0$

Distribution of Test Statistic under H_0

Assume no ties, let B denote the number of $Z_i - \theta_0$ values that are greater than 0, and let $r_1 < r_2 < \cdots < r_B$ denote the (ordered) ranks of the positive $Z_i - \theta_0$ values

- Note that $T^+ = \sum_{i=1}^{B} r_i$

Under $H_0 : \theta = \theta_0$ we have that $Z_i - \theta_0 \sim \tilde{F}_i$, which is continuous and symmetric around 0.

All 2^n possible outcomes for $(r_1, r_2, \ldots, r_B)$ occur with equal probability.

- For given n, form all 2^n possible outcomes with corresponding T^+
- Each outcome has probability $\frac{1}{2^n}$ under H_0

Null Distribution Example

Suppose we have $n = 3$ observations (Z_1, Z_2, Z_3) with no ties.

The $2^3 = 8$ possible outcomes for $(r_1, r_2, \ldots, r_B)$ are

B	$(r_1, r_2, \ldots, r_B)$	$T^+ = \sum_{i=1}^{B} r_i$	Probability under H_0
0		0	$1/8$
1	$r_1 = 1$	1	$1/8$
1	$r_1 = 2$	2	$1/8$
1	$r_1 = 3$	3	$1/8$
2	$r_1 = 1, r_2 = 2$	3	$1/8$
2	$r_1 = 1, r_2 = 3$	4	$1/8$
2	$r_1 = 2, r_2 = 3$	5	$1/8$
3	$r_1 = 1, r_2 = 2, r_3 = 3$	6	$1/8$

Example probability calculation: $P(T^+ < 2) = \sum_{i=0}^{1} P(T^+ = i) = 0.25$

Hypothesis Testing

One-Sided Upper Tail Test:

- $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$
- Reject H_0 if $T^+ \geq t_\alpha$ where $P(T^+ > t_\alpha) = \alpha$

One-Sided Lower Tail Test:

- $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$
- Reject H_0 if $T^+ \leq \frac{n(n+1)}{2} - t_\alpha$

Two-Sided Test:

- $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$
- Reject H_0 if $T^+ \geq t_{\alpha/2}$ or $T^+ \leq \frac{n(n+1)}{2} - t_{\alpha/2}$

Large Sample Approximation

Under H_0, the expected value and variance of T^+ are

- $E(T^+) = \frac{n(n+1)}{4}$
- $V(T^+) = \frac{n(n+1)(2n+1)}{24}$

We can create a standardized test statistic T^* of the form

$$T^* = \frac{T^+ - E(T^+)}{\sqrt{V(T^+)}}$$

which asymptotically follows a $\mathrm{N}(0, 1)$ distribution.

Derivation of Large Sample Approximation

Note that we have $T^+ = \sum_{i=1}^n U_i$ where

- $U_i = R_i\psi_i$ are independent variables for $i = 1, \ldots, n$
- $P(U_i = i) = P(U_i = 0) = 1/2$

Using the independence of the U_i variables we have

- $E(T^+) = \sum_{i=1}^n E(U_i)$
- $V(T^+) = \sum_{i=1}^n V(U_i)$

Using the distribution of U_i we have

- $E(U_i) = i\frac{1}{2} + 0\frac{1}{2} = \frac{i}{2} \implies E(T^+) = \frac{1}{2}\sum_{i=1}^n i = \frac{n(n+1)}{4}$
- $V(U_i) = E(U_i^2) - [E(U_i)]^2 = \frac{i^2}{2} - \frac{i^2}{4} = \frac{i^2}{4} \implies$ $V(T^+) = \frac{1}{4}\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{24}$

Handling Zeros and Ties

If $Z_i = \theta_0$, then discard Z_i and redefine $\tilde{n}$ as the number of observations that do not equal θ_0.

If $Z_i = Z_j$ for two (non-zero) observations, then use the average ranking procedure to handle ties.

- T^+ is calculated in same fashion (using average ranks)
- No longer an exact level α test
- Need to adjust variance term for large sample approximation

Example 3.1: Description

Hamilton Depression Scale Factor IV measures suicidal tendencies.

- Higher scores mean more suicidal tendencies

Nine psychiatric patients were treated with a tranquilizer drug.

X and Y are pre- and post-treatment Hamilton Depression Scale Factor IV scores

Want to test if the tranquilizer significantly reduced suicidal tendencies

- $H_0 : \theta = 0$ versus $H_1 : \theta < 0$.
- θ is median of $Z = Y - X$

Example 3.1: Data

Nonparametric Statistical Methods, 3rd Ed. (Hollander et al., 2014)

Table 3.1 The Hamilton Depression Scale Factor IV Values

Patient i	X_i	Y_i
1	1.83	0.878
2	0.50	0.647
3	1.62	0.598
4	2.48	2.050
5	1.68	1.060
6	1.88	1.290
7	1.55	1.060
8	3.06	3.140
9	1.30	1.290

Source: D. S. Salsburg (1970).

Example 3.1: By Hand

Patient i	X_i	Y_i	Z_i	R_i	ψ_i
1	1.83	0.878	-0.952	8	0
2	0.50	0.647	0.147	3	1
3	1.62	0.598	-1.022	9	0
4	2.48	2.050	-0.430	4	0
5	1.68	1.060	-0.620	7	0
6	1.88	1.290	-0.590	6	0
7	1.55	1.060	-0.490	5	0
8	3.06	3.140	0.080	2	1
9	1.30	1.290	-0.010	1	0

Note. $Z_i = Y_i - X_i$ and R_i is rank of $|Z_i|$

$$T^+ = \sum_{i=1}^n R_i\psi_i = 3 + 2 = 5$$

Example 3.1: Using R

```
> pre = c(1.83,0.50,1.62,2.48,1.68,1.88,1.55,3.06,1.30)
> post = c(0.878,0.647,0.598,2.050,1.060,1.290,1.060,3.140,1.290)
> z = post - pre
> wilcox.test(z,alternative="less")

 Wilcoxon signed rank test

data:  z
V = 5, p-value = 0.01953
alternative hypothesis: true location is less than 0

> wilcox.test(post,pre,alternative="less",paired=TRUE)

 Wilcoxon signed rank test

data:  post and pre
V = 5, p-value = 0.01953
alternative hypothesis: true location shift is less than 0
```

An Estimator of θ

To estimate the median (or median difference) θ, first form the $M = n(n+1)/2$ average values

$$W_{ij} = (Z_i + Z_j)/2$$

for $i \leq j = 1, \ldots, n$, which are known as Walsh averages.

The estimate of θ corresponding to Wilcoxon's signed rank test is

$$\hat{\theta} = \text{median}(W_{ij};\; i \leq j = 1, \ldots, n)$$

which is the median of the Walsh averages.

Motivation: make mean of $Z_i - \hat{\theta}$ as close as possible to $n(n+1)/4$.

Symmetric Two-Sided Confidence Interval for θ

Define the following terms

- $M = n(n+1)/2$ is the number of Walsh averages
- $W^{(1)} \leq W^{(2)} \leq \cdots \leq W^{(M)}$ are the ordered Walsh averages
- $t_{\alpha/2}$ is the critical value such that $P(T^+ > t_{\alpha/2}) = \alpha/2$ under H_0
- $C_\alpha = M + 1 - t_{\alpha/2}$ is the transformed critical value

A symmetric $(1-\alpha)100\%$ confidence interval for θ is given by

$$\theta_L = W^{(C_\alpha)}$$
$$\theta_U = W^{(M+1-C_\alpha)} = W^{(t_{\alpha/2})}$$

One-Sided Confidence Intervals for θ

Define the following additional terms

- t_α is the critical value such that $P(T^+ > t_\alpha) = \alpha$ under H_0
- $C_\alpha^* = M + 1 - t_\alpha$ transformed critical value

An asymmetric $(1-\alpha)100\%$ upper confidence bound for θ is

$$\theta_L = -\infty$$
$$\theta_U = W^{(M+1-C_\alpha^*)} = W^{(t_\alpha)}$$

An asymmetric $(1-\alpha)100\%$ lower confidence bound for θ is

$$\theta_L = W^{(C_\alpha^*)}$$
$$\theta_U = \infty$$

Example 3.1: Estimate θ

Get $W^{(1)} \leq W^{(2)} \leq \cdots \leq W^{(M)}$ and $\hat{\theta}$ for previous example:

```
> require(NSM3)  # use install.packages("NSM3") to get NSM3
> owa(pre,post)
$owa
 [1] -1.0220 -0.9870 -0.9520 -0.8210 -0.8060 -0.7860 -0.7710
 [8] -0.7560 -0.7260 -0.7210 -0.6910 -0.6200 -0.6050 -0.5900
[15] -0.5550 -0.5400 -0.5250 -0.5160 -0.5100 -0.4900 -0.4810
[22] -0.4710 -0.4600 -0.4375 -0.4360 -0.4300 -0.4025 -0.3150
[29] -0.3000 -0.2700 -0.2550 -0.2500 -0.2365 -0.2215 -0.2200
[36] -0.2050 -0.1750 -0.1715 -0.1415 -0.0100  0.0350  0.0685
[43]  0.0800  0.1135  0.1470

$h.l
[1] -0.46
```

Example 3.1: Confidence Interval for θ

```
> wilcox.test(z,alternative="less",conf.int=TRUE)

 Wilcoxon signed rank test

 data:  z
 V = 5, p-value = 0.01953
alternative hypothesis: true location is less than 0
 95 percent confidence interval:
   -Inf -0.175
 sample estimates:
 (pseudo)median
         -0.46
```

Fisher's Sign Test

Assumptions and Hypothesis

Assumes only independence (no symmetry assumption).

The null hypothesis about θ (common median) is

$$H_0 : \theta = \theta_0$$

and we could have one of three alternative hypotheses:

- One-Sided Upper-Tail: $H_1 : \theta > \theta_0$
- One-Sided Lower-Tail: $H_1 : \theta < \theta_0$
- Two-Sided: $H_1 : \theta \neq \theta_0$

Test Statistic

Defining the indicator variable

$$\psi_i = \begin{cases} 1 & \text{if } Z_i - \theta_0 > 0 \\ 0 & \text{if } Z_i - \theta_0 < 0 \end{cases}$$

the sign test statistic B is defined as

$$B = \sum_{i=1}^{n} \psi_i$$

which is the number of positive $Z_i - \theta_0$ values.

Distribution of Test Statistic under H_0

If θ_0 is the true median, ψ_i has a 50% chance of taking each value:

- $P(\psi = 0|\theta = \theta_0) = P(\psi = 1|\theta = \theta_0) = 1/2$

Thus, the sign statistic follows a binomial distribution under H_0

- $B \sim \text{Binom}(n, 1/2)$

Hypothesis Testing

One-Sided Upper Tail Test:

- $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$
- Reject H_0 if $B \geq b_\alpha$ where $P(B > b_\alpha) = \alpha$

One-Sided Lower Tail Test:

- $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$
- Reject H_0 if $B \leq n - b_\alpha$

Two-Sided Test:

- $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$
- Reject H_0 if $B \geq b_{\alpha/2}$ or $B \leq n - b_{\alpha/2}$

Large Sample Approximation

Under H_0, $B \sim \text{Binom}(n, 1/2)$ so the expected value and variance are

- $E(B) = np = \frac{n}{2}$
- $V(B) = np(1 - p) = \frac{n}{4}$

We can create a standardized test statistic B^* of the form

$$B^* = \frac{B - E(B)}{\sqrt{V(B)}}$$

which asymptotically follows a $\mathrm{N}(0, 1)$ distribution.

Example 3.1: Revisited (By Hand)

Patient i	X_i	Y_i	Z_i	ψ_i
1	1.83	0.878	-0.952	0
2	0.50	0.647	0.147	1
3	1.62	0.598	-1.022	0
4	2.48	2.050	-0.430	0
5	1.68	1.060	-0.620	0
6	1.88	1.290	-0.590	0
7	1.55	1.060	-0.490	0
8	3.06	3.140	0.080	1
9	1.30	1.290	-0.010	0

$B = \sum_{i=1}^{n} \psi_i = 2$ and p-value = $P(B < 2 | H_0 \text{ is true}) = 0.0898$

```
> pbinom(2,9,1/2)
[1] 0.08984375
```

Example 3.1: Revisited (Using R, one-sample)

```
> library(BSDA)
> z = post - pre
> SIGN.test(z,alternative="less")

 One-sample Sign-Test

data:  z
s = 2, p-value = 0.08984
alternative hypothesis: true median is less than 0
95 percent confidence interval:
  -Inf 0.041
sample estimates:
median of x
      -0.49

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.9102   -Inf -0.010
Interpolated CI       0.9500   -Inf  0.041
Upper Achieved CI     0.9805   -Inf  0.080
```

Example 3.1: Revisited (Using R, paired-samples)

```
> library(BSDA)
> SIGN.test(post,pre,alternative="less")

 Dependent-samples Sign-Test

data:  post and pre
S = 2, p-value = 0.08984
alternative hypothesis: true median difference is less than 0
95 percent confidence interval:
  -Inf 0.041
sample estimates:
median of x-y
        -0.49

                    Conf.Level L.E.pt U.E.pt
Lower Achieved CI      0.9102   -Inf -0.010
Interpolated CI        0.9500   -Inf  0.041
Upper Achieved CI      0.9805   -Inf  0.080
```

A Different Estimator of θ

To estimate the median (or median difference) θ, calculate

$$\tilde{\theta} = \text{median}(Z_i;\ i = 1, \ldots, n)$$

which is the median of observed sample (or paired differences).

```
> median(z)
[1] -0.49
```

Motivation: make mean of $Z_i - \tilde{\theta}$ as close as possible to $n/2$.

Symmetric Two-Sided Confidence Interval for θ

Define the following terms

- $b_{\alpha/2}$ is the critical value such that $P(B > b_{\alpha/2}) = \alpha/2$ under H_0
- $C_\alpha = n + 1 - b_{\alpha/2}$ is the transformed critical value

A symmetric $(1-\alpha)100\%$ confidence interval for θ is given by

$$\theta_L = Z^{(C_\alpha)}$$
$$\theta_U = Z^{(n+1-C_\alpha)} = Z^{(b_{\alpha/2})}$$

where $Z^{(i)}$ is the i-th order statistic of the sample $\{Z_i\}_{i=1}^n$.

One-Sided Confidence Intervals for θ

Define the following additional terms

- b_α is the critical value such that $P(B > b_\alpha) = \alpha$ under H_0
- $C_\alpha^* = n + 1 - b_\alpha$ transformed critical value

An asymmetric $(1 - \alpha)100\%$ upper confidence bound for θ is

$$\theta_L = -\infty$$
$$\theta_U = Z^{(n+1-C_\alpha^*)} = Z^{(b_\alpha)}$$

An asymmetric $(1 - \alpha)100\%$ lower confidence bound for θ is

$$\theta_L = Z^{(C_\alpha^*)}$$
$$\theta_U = \infty$$

Example 3.1: Revisited Confidence Interval for θ

```
> zs = sort(z)
> zs
[1] -1.022 -0.952 -0.620 -0.590 -0.490
[6] -0.430 -0.010  0.080  0.147
> round(pbinom(0:9,9,1/2),4)
[1] 0.0020 0.0195 0.0898 0.2539 0.5000
[6] 0.7461 0.9102 0.9805 0.9980 1.0000
 > zs[7:8]
[1] -0.01  0.08
> zs[7]+(zs[8]-zs[7])*(0.95-0.9102)/(0.9805-0.9102)
[1] 0.04095306
```

The asymmetric 95% upper confidence bound is $(-\infty, 0.041)$.

Some Considerations

Which Location Test Should You Choose?

Answer depends on your data and what assumptions you are willing to make about the population distribution.

If observed data are normally distributed, then. . .

- t-test is most powerful test
- Wilcoxon's signed rank test is slightly less powerful than t test (4.5% efficiency loss)
- Fisher's sign test is less powerful than others (36.3% efficiency loss compared to t test)

If observed data are NOT normally distributed, then. . .

- Signed rank test is typically as or more efficient than t test
- Sign test should be preferred if data population is asymmetric

Assumptions and Hypotheses

Assumes $z_i \stackrel{\text{iid}}{\sim} F$ where θ is the median of F, i.e., $F(\theta) = 1/2$

The null hypothesis is that F is symmetric around θ, i.e.,

$$H_0 : F(\theta - b) + F(\theta + b) = 1 \quad \forall b$$

and we could have one of three alternative hypotheses:

- One-Sided Left-Skew: $H_1 : F(\theta + b) \geq 1 - F(\theta - b) \quad \forall b > 0$
- One-Sided Right-Skew: $H_1 : F(\theta + b) \leq 1 - F(\theta - b) \quad \forall b > 0$
- Two-Sided: $F(\theta - b) + F(\theta + b) \neq 1$ for any b

Test Statistic

For every triple of observations (Z_i, Z_j, Z_k), $1 \leq i < j < k \leq n$, define

$$f^*(Z_i, Z_j, Z_k) = \text{sign}(Z_i + Z_j - 2Z_k) + \text{sign}(Z_i + Z_k - 2Z_j) + \text{sign}(Z_j + Z_k - 2Z_i)$$

and note that there are $n(n-1)(n-2)/6$ distinct triples in the sample.

- (Z_i, Z_j, Z_k) is a left triple (skewed to left) if $f^*(Z_i, Z_j, Z_k) = -1$
- (Z_i, Z_j, Z_k) is a right triple (skewed to right) if $f^*(Z_i, Z_j, Z_k) = 1$
- If $f^*(Z_i, Z_j, Z_k) = 0$, then (Z_i, Z_j, Z_k) is neither left nor right

Define the unstandardized test statistic

$$\begin{aligned} T &= \sum_{1 \leq i < j < k \leq n} f^*(Z_i, Z_j, Z_k) \\ &= \{\text{\# of right triples}\} - \{\text{\# of left triples}\} \end{aligned}$$

Test Statistic (continued)

Define the standardized test statistic

$$V = T/\hat{\sigma} \overset{\text{asy}}{\sim} \mathrm{N}(0, 1)$$

where the variance estimate is given by

$$\hat{\sigma}^2 = \frac{(n-3)(n-4)}{(n-1)(n-2)} \sum_{t=1}^{n} B_t^2 + \frac{(n-3)}{(n-4)} \sum_{s=1}^{n-1} \sum_{t=s+1}^{n} B_{s,t}^2 + \frac{n(n-1)(n-2)}{6} - \left[1 - \frac{(n-3)(n-4)(n-5)}{n(n-1)(n-2)}\right] T^2$$

and the B_t and B_{st} terms are defined as

$$B_t = \{\text{\# right triples involving } Z_t\} - \{\text{\# left triples involving } Z_t\}$$
$$B_{st} = \{\text{\# right triples involving } Z_s, Z_t\} - \{\text{\# left triples involving } Z_s, Z_t\}$$

Hypothesis Testing

One-Sided Left-Skew Test:

- $H_0 : F$ is symmetric versus $H_1 : F$ is left-skewed
- Reject H_0 if $V \leq -Z_\alpha$ where $P(Z > Z_\alpha) = \alpha$

One-Sided Right-Skew Test:

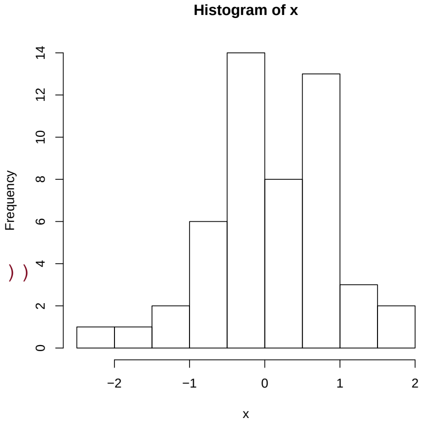
- $H_0 : F$ is symmetric versus $H_1 : F$ is right-skewed
- Reject H_0 if $V \geq Z_\alpha$

Two-Sided Test:

- $H_0 : F$ is symmetric versus $H_1 : F$ is not symmetric
- Reject H_0 if $|V| \geq Z_{\alpha/2}$

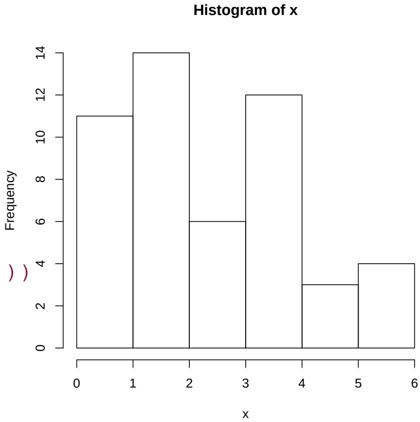
Example 1: Symmetric

```
> set.seed(1)
> x = rnorm(50)
> hist(x)
> require(NSM3)
> test = RFPW(x)
> c(test$obs.stat, test$p.val)
[1] -1.4572415  0.1450497

# two-sided
> 2*(1 - pnorm(abs(test$obs.stat)))
[1] 0.1450497
# left-skew
> pnorm(test$obs.stat)
[1] 0.07252486
# right-skew
> 1 - pnorm(test$obs.stat)
[1] 0.9274751
```



Example 2: Asymmetric

```
> set.seed(1)
> x = rchisq(50,df=3)
> hist(x)
> require(NSM3)
> test = RFPW(x)
> c(test$obs.stat, test$p.val)
[1] 1.70708892 0.08780553

# two-sided
> 2*(1 - pnorm(abs(test$obs.stat)))
[1] 0.08780553
# left-skew
> pnorm(test$obs.stat)
[1] 0.9560972
# right-skew
> 1 - pnorm(test$obs.stat)
[1] 0.04390276
```



Exchangeability

Components of a random vector (X, Y) are exchangeable if the vectors (X, Y) and (Y, X) have the same distribution.

- Permuting components does not change distribution
- Implies $F_X \equiv F_Y$ and $F_{X|Y} \equiv F_{Y|X}$ and $F_Z \equiv F_{-Z}$ with $Z = Y - X$
- $F_Z \equiv F_{-Z}$ implies that F_Z is symmetric about 0

More generally, if components of $(X + \theta, Y)$ are exchangeable, then

$$Z - \theta = Y - (X + \theta)$$

has the same distribution as

$$\theta - Z = (X + \theta) - Y$$

implies that F_Z is symmetric about θ

Assumptions and Hypotheses

Assumes $(x_i, y_i) \stackrel{\text{iid}}{\sim} F(x, y)$ where F is some bivariate distribution.

The null hypothesis is that F is exchangeable, i.e.,

$$H_0 : F(x, y) = F(y, x) \quad \forall x, y$$

and there is only one possible alternative hypothesis

$$H_1 : F(x, y) \neq F(y, x) \quad \text{for some } x, y$$

Test Statistic

For each pair (x_i, y_i) let $a_i = \min(x_i, y_i)$ and $b_i = \max(x_i, y_i)$, and define

$$r_i = \begin{cases} 1, & \text{if } x_i = a_i < b_i = y_i \\ 0, & \text{if } x_i = b_i \geq a_i = y_i \end{cases}$$

so that $r_i = 1$ if $x_i < y_i$ and $r_i = 0$ otherwise.

Next, define the n^2 values d_{ij}, for $i, j = 1, \ldots, n$, as

$$d_{ij} = \begin{cases} 1, & \text{if } a_j < b_i \leq b_j \text{ and } a_i \leq a_j \\ 0, & \text{otherwise} \end{cases}$$

Test Statistic (continued)

For each $j = 1, \ldots, n$ calculate the signed summation of d_{ij} as

$$T_j = \sum_{i=1}^{n} s_i d_{ij}$$

where $s_i = 2r_i - 1$. Note that $s_i = 1$ if $r_i = 1$ and $s_i = -1$ if $r_i = 0$.

Finally, calculate the observed test statistic

$$A_{\text{obs}} = \frac{1}{n^2} \sum_{j=1}^{n} T_j^2$$

Distribution of Test Statistic under H_0

In addition to observed $(r_1, \ldots, r_n)$, there are 2^{n-1} other possibilities.

- r_i can be 0 or 1, so there are 2^n total configurations
- Each configuration is equally likely under H_0

Let $A^{(1)} \leq A^{(2)} \leq \cdots \leq A^{(2^n)}$ denote the 2^n values of the test statistic.

- Need to form all possible $A^{(k)}$ values for make null distribution
- Note that d_{ij} is same for all $A^{(k)}$ values (by definition of d_{ij})

Hypothesis Testing

Two-Sided Test:

- $H_0 : F$ is exchangeable versus $H_1 : F$ is not exchangeable
- Reject H_0 if $A_{\mathrm{obs}} > A^{(m)}$ where $m = 2^n - \lfloor 2^n \alpha \rfloor$

If you are unlucky and $A_{\mathrm{obs}} = A^{(m)}$, use a randomized decision.

- Reject H_0 with probability $q = \frac{2^n \alpha - M_1}{M_2}$
- $M_1 = \sum_{k=1}^{2^n} 1_{\{A^{(k)} > A^{(m)}\}}$ is the # of $A^{(k)}$ values greater than $A^{(m)}$
- $M_2 = \sum_{k=1}^{2^n} 1_{\{A^{(k)} = A^{(m)}\}}$ is the # of $A^{(k)}$ values equal to $A^{(m)}$

Example 3.1: Exchangeability Test

```
> pre = c(1.83,0.50,1.62,2.48,1.68,1.88,1.55,3.06,1.30)
> post = c(0.878,0.647,0.598,2.050,1.060,1.290,1.060,3.140,1.290)
> require(NSM3)
> HollBivSym(pre,post)
[1] 0.6666667
> set.seed(1)
> pHollBivSym(pre,post)
Number of X values:  9 Number of Y values:  9
Hollander A  Statistic:  0.6667
Monte Carlo  (Using  10000 Iterations) upper-tail probability:  0.0321
```