Lecture 11: Nonparametric Inference

The Wilcoxon Tests

Let X be a random variable of the continuous type and let m denote the median of X. To test $H_0: m = m_0$ vs $H_a: m \neq m_0$, we use the sign test.

If we denote the observations from this distribution by $X_1, X_2, \ldots, X_n$ and let Y be the number of negative differences among $X_1 - m_0, X_2 - m_0, \ldots, X_n - m_0$, then $Y \sim b(n, 1/2)$ under H_0 and is the test statistic for the sign test. For Y too large or too small, we reject H_0.

Example 1

Let X be the length of time in seconds between two calls entering a call center. Let m be the unique median of this distribution. We test $H_0: m = 6.2$ vs $H_a: m < 6.2$

Let Y = # of lengths of time between calls in a random sample of size 20 that are less than 6.2. The critical region $C = \{y : y \geq 14\}$ has level $\alpha = 0.0577$.

A sample of size 20 yielded the following data:

6.8	5.7	6.9	5.3	4.1	9.8
1.7	7.0	2.1	19.0	18.9	16.9
10.4	44.1	2.9	2.4	4.8	18.9
4.8	7.9				

Here $y = 9 < 14 \Rightarrow H_0$ is not rejected.

Now consider the ~~k~~ pairs (X_1, Y_1), $(X_2, Y_2), \ldots, (X_n, Y_n)$, where X and Y are dependent continuous random variables.

To test $H_0: p = 1/2$ vs H_a: any deviation from H_0, we let $W =$ # pairs for which $X_k - Y_k > 0$. Under H_0, $W \sim b(n, 1/2)$ and the test can be based on W.

Example: if X = length of right foot and
Y = length of left foot of an indiv.,
then $p = P(X > Y) = 1/2$ implies that
either foot of the individual is
equally likely to be longer.

The sign test does not take into account the magnitude of the differences $X_1 - m_0, X_2 - m_0, \ldots, X_n - m_0$, and hence is less popular than the Wilcoxon signed rank test.

The Wilcoxon signed rank test accounts for both the sign and magnitude of the differences.

Let $H_0: m = m_0$ (known constant).

Take a random sample $X_1, X_2, \ldots, X_n$ and rank the absolute values $|X_1 - m_0|, |X_2 - m_0|, \ldots, |X_n - m_0|$ in ascending order of magnitude.

So, for $k = 1, 2, \ldots, n$, let

$$R_k = \text{rank of } |X_k - m_0|$$

among $|X_1 - m_0|, |X_2 - m_0|, \ldots, |X_n - m_0|$.
Then $R_1, R_2, \ldots, R_n$ becomes just a permutation of the first n positive integers $1, 2, \ldots, n$. For each R_k, associate the sign of the difference $X_k - m_0$, such that if $X_k - m_0 > 0$, then we use R_k, but if $X_k - m_0 < 0$, then we use $-R_k$. Then the Wilcoxon statistic W is the sum of these n signed ranks, hence the name Wilcoxon signed rank statistic.

Example 2

Suppose the lengths of $n = 10$ sunfish

are: 5.0, 3.9, 5.2, 5.5, 2.8, 6.1, 6.4, 2.6, 1.7, 4.3

We wish to test $H_0: m = 3.7$ vs $H_a: m > 3.7$

So $x_k - m_0$: 1.3, 0.2, 1.5, 1.8, −0.9, 2.4, 2.7, −1.1, −2.0, 0.6

$|x_k - m_0|$: 1.3, 0.2, 1.5, 1.8, 0.9, 2.4, 2.7, 1.1, 2.0, 0.6

Ranks: 5, 1, 6, 7, 3, 9, 10, 4, 8, 2

Signed ranks: 5, 1, 6, 7, −3, 9, 10, −4, −8, 2

$\Rightarrow W = 5 + 1 + 6 + 7 - 3 + 9 + 10 - 4 - 8 + 2 = 25$

If $H_0: m = m_0$ is true, the half of the differences would be negative and thus half of the signs would be negative

Hence $H_0: m = m_0$ is supported when $W = w$ is close to zero and not supported ($H_a: m > m_0$ favoured) when $W = w$ is too large. (larger deviations $|X_k - m_0|$ associated with observations for which $x_k - m_0 > 0$).

Critical regions: $H_a: m > m_0 \Rightarrow \{w: w \geq c_1\}$

$H_a: m < m_0 \Rightarrow \{w: w \leq c_2\}$

$H_a: m \neq m_0 \Rightarrow \{w: w \leq c_3 \text{ or } w \geq c_4\}$.

c_1, c_2, c_3 and c_4 are determined from the distribution of W under H_0. When H_0 is true, $P(X_k < m_0) = P(X_k > m_0) = 1/2$, $k = 1, 2, \ldots, n$.

$\Rightarrow R_k$ will be associated with a negative sign with a prob of $1/2$.

We assume that the distribution of X_k is <u>symmetric</u> about m_0. Since $X_1, X_2, \ldots, X_n$ are mutually independent, the assignments of these n signs are independent as well. We can write

$$W = \sum_{k=1}^{n} \text{sign}(k) R_k = \sum_{k=1}^{n} \text{sgn}(k) R_k$$

Because of symmetry, W has the same distribution as the r.v.

$$V = \sum_{k=1}^{n} V_k, \text{ where } V_1, V_2, \ldots, V_n \text{ are indep.}$$

and $P(V_k = k) = P(V_k = -k) = 1/2$, $k = 1, 2, \ldots, n$.

Now, $E(V_k) = -k(1/2) + k(1/2) = 0$

$$E(W) = E(V) = \sum_{k=1}^{n} E(V_k) = 0$$

$$\text{Var}(V_k) = E(V_k^2) = (-k)^2(1/2) + (k)^2(1/2)$$

$$= k^2$$

Hence, $\text{Var}(W) = \text{var}(V) = \sum_{k=1}^{n} \text{Var}(V_k)$

$$= \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

From the Central Limit Theorem,

$$Z = \frac{W - 0}{\sqrt{n(n+1)(2n+1)/6}} \approx N(0,1) \text{ under } H_0.$$

So, for large n, $P(W \geq c \mid H_0) = P(Z \geq z_\alpha \mid H_0)$

Example 3

Let m be the median of a symmetric distribution of the continuous type. To test $H_0: m = 160$ vs $H_a: m > 160$ at $\alpha = 0.05$ based on a random sample of size $n = 16$, we reject H_0 when $z = \frac{w}{\sqrt{16(17)(33)/6}} \geq 1.645$

or when $w \geq 1.645\sqrt{\frac{16(17)(33)}{6}} = 63.626$

If the observed values are:

176.9, 158.3, 152.1, 158.8, 172.4, 169.8, 159.7
162.7, 156.6, 174.5, 184.4, 165.2, 147.8, 177.8,
160.1, 160.5

Then $w = 1-2+3-4-5+6+\ldots+16 = 60$

Since $60 < 63.626$, H_0 is not rejected

at $\alpha = 0.05$

To obtain the p-value, we make a

unit continuity correction and compute

$$\text{p-value} = P(W \geq 60) = P\left(\frac{W-0}{\sqrt{16(17)33/6}} \geq \frac{59-0}{\sqrt{16(17)(33/6)}}\right)$$

$$\approx P(Z \geq 1.525) = 0.0636 > 0.05$$

<u>Note</u>: Ties are assigned the average of corresponding ranks, while $x_k = m_0$ is deleted and reduced sample size used.

Exercise

It is claimed that the median weight m of certain loads of candy is 40,000 pounds.

a) Use the following 13 observations and the Wilcoxon statistic to test the null hypothesis

$H_0: m = 40{,}000$ vs $H_a: m < 40{,}000$
at $\alpha = 0.05$:

41,195	39,485	41,229	36,840	38,050
40,890	38,345	34,930	39,245	31,031
40,780	38,050	30,906		

b) What is the approximate p-value of this test?

c) Use the sign test to test the same hypothesis

d) Calculate the p-value from the sign test and compare it with the p-value obtained from the Wilcoxon test.

The Wilcoxon test can also be used to test for equality of two medians, i.e the equality of two continuous distributions. Let F and G be two continuous distributions with the same shape and spread but possibly different locations (medians). Then there exists a constant c such that

$$F(x) = G(x+c) \quad \text{for all } x.$$

Let $X_1, X_2, \ldots, X_{n_1}$ and $Y_1, Y_2, \ldots, Y_{n_2}$ be two random samples from $F(x)$ and $G(y)$, respectively.

We can combine the two samples and assign ranks $1, 2, \ldots, n_1+n_2$ to the ordered values

Let w = sum of ranks for $y_1, y_2, \ldots, y_{n_2}$.

Assign average ranks to ties.

If the distribution of Y is shifted to the right of X (ie. $F(x) = G(x+c)$), then y-values would tend to be larger than the values of X and w would be larger than expected when $F(z) = G(z)$.

If m_X and m_Y are respective medians, then the critical region for testing $H_0: m_X = m_Y$ vs $H_a: m_X < m_Y$ would be of the form $\{w: w \geq c\}$. Similarly, if $H_a: m_X > m_Y$, $C = \{w: w \leq c\}$.

We use the normal approximation to derive the distribution of W.

When $F(z) = G(z)$,

$$E(W) = \frac{n_2(n_1+n_2+1)}{2} = \mu_w$$

and $$Var(W) = \frac{n_1 n_2(n_1+n_2+1)}{12} = \sigma_w^2$$

and $$Z = \frac{W - \mu_w}{\sigma_w} \approx N(0,1)$$

Example 3:

The weights of the contents of $n_1 = 8$ and $n_2 = 8$ tins of cinnamon packaged by companies A and B, respectively, selected at random, yielded the following observations of X and Y:

x: 117.1 121.3 127.8 121.9 117.4 124.5
119.5 115.1

y: 123.5 125.3 126.5 127.9 122.1 125.6
129.8 117.2

$C = \{w : w \geq c\}$. Since $n_1 = n_2 = 8$,

$$z = \frac{w - \mu_w}{\sigma_w} = \frac{w - 8(8+8+1)/2}{\sqrt{(8)(8)(8+8+1)/12}} \geq 1.645$$

or $w \geq 1.645 \sqrt{\frac{(8)(8)(17)}{12}} + 4(17) = 83.66$

at $\alpha = 0.05$.

It is easy to see that $w = 3+8+9+11+12+13+15+16$
(exercise) $= 87 > 83.66$

Thus, H_0 is rejected. The p-value of the test is computed as:

$$\text{p-value} = P(W \geq 87)$$

$$= P\left(\frac{W-68}{\sqrt{90.667}} \geq \frac{86.5-68}{\sqrt{90.667}}\right)$$

$$\approx P(Z \geq 1.943) = 0.0260 < 0.05$$

Here, a half-unit correction for continuity has been made.