Examples about paried-sample comparison

Example 2.3: A new type of Beta-Blockers invented by Merke is at the stage of human clinical trial, in which 10 patients with hypertension are recruited to take the medicine. Their systolic blood pressure before and after trial are recorded:

$$\begin{array}{lll} (145,136), & (173,160), & (146,140) \\ (154,149), & (166,168), & (143,149) \\ (157,151), & (183,177), & (184,179), \\ (177,175) & & \end{array}$$

Does the data contain significant (at significant leval 1%) evidence indicating the drug is effective in reducing the blood pressure for hypothesis patients?

Example 2.4: (Following Ex2.3) Suppose 100 patients are recruited and, among them, 80 (20) have blood pressure reduced (increased) after the trial. Does this data indicate the drug is effective?

Estimation and Confidence interval for median m

Suppose $X_1, X_2, \cdots, X_n$i.i.d. $\sim F$ is continuous with median m. The point estimation for m by sample median:

$$\hat{m} = \begin{cases} X_{(\frac{n+1}{2})} & \text{,if n is odd} \\ \frac{1}{2}[X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}] & \text{,if n is even} \end{cases}$$

where $X_{(1)} \leq X_{(2)} \leq \cdots X_{(n)}$ are the ordered $\{X_1, X_2, \cdots, X_n\}$.

Theorem: (Confidence interval) Let $S = \sum_{i=1}^{n} I(X_i > m)$, the total number of $\{X_1, X_2, \cdots, X_n\}$ that are greater than m. Let k_α denote the positive integer such that:

$$P(S < k_\alpha) = \frac{\alpha}{2}$$

Then, $\left[X_{(k_\alpha)}, X_{(n-k_\alpha+1)}\right]$ is a confidence interval for m at confidence level $1-\alpha$. $(0 < \alpha < 1/2)$

Remark: For $0 < \alpha < 1/2$,

$$P(Bin(n,1/2) < k_\alpha) = \alpha/2 \Rightarrow \sum_{i=0}^{k_\alpha-1} \left(\begin{array}{c} n \\ i \end{array}\right) \left(\begin{array}{c} 1 \\ 2 \end{array}\right)^n = \alpha/2$$

If the above equality doesn't hold for any integer k_α, we set k_α as the integer at $\sum_{i=0}^{k_\alpha-1} \left(\begin{array}{c} n \\ i \end{array}\right) \left(\begin{array}{c} 1 \\ 2 \end{array}\right)^n$ is the closest to $\alpha/2$.

The Wilcoxon Signed Rank Test

The sign test uses only the signs of $X_i - m_0$, not the magnitude of $X_i - m_0$. The Wilcoxon signed rank test uses both the signs and the magnitude of $X_i - m_0$, under a little bit more assumption about the population distribution.

• For observations $X_1, X_2, \cdots, X_n$ with median m_0, the signed rank R_i is:

$$R_i = \text{rank of } |X_i - m_0| \times \text{sign of } X_i - m_0$$

Example 2.5: let $n = 4, m_0 = 1$. There are four observations X_1, X_2, X_3, X_4. Consider

$$\begin{cases} H_0 : m = m_0 \\ H_a : m > m_0 \end{cases}$$

X_i	X_1	X_2	X_3	X_4
obs	2.1	-1	-2.2	-0.3
$X_i - m_0$	1.1	-2	-3.2	-1.3
$\|X_i - m_0\|$	1.1	2	3.2	1.3
R of $\|X_i - m_0\|$				
Signs $X_i - m_0$				
R_i				

Example 2.6 Let $m_0 = 1$

X_i	X_1	X_2	X_3	X_4
obs	3.1	2.1	-1.3	3.3
$X_i - m_0$	2.1	1.1	-2.3	2.3
$\|X_i - m_0\|$	2.1	1.1	2.3	2.3
R of $\|X_i - m_0\|$				
Signs of $X_i - m_0$				
Signed rank R_i				

X_i	X_1	X_2	X_3	X_4	X_5
obs	3.1	4.2	-1.6	3.6	3.6
$X_i - m_0$	2.1	3.2	-2.6	2.6	2.6
$\|X_i - m_0\|$	2.1	3.2	2.6	2.6	2.6
R of $\|X_i - m_0\|$					
Signs $X_i - m_0$					
Signed R R_i					

Let $I_i = I(X_i > m_0) = I(R_i \text{ is positive})$, the Wilcoxon signed rank test statistic is

$$\begin{aligned} W = & \quad \text{the sum of all positive signed ranks} \\ = & \quad \sum_{i=1}^{n} R_i I_i \\ = & \quad \frac{1}{2}\sum_{i=1}^{n} R_i + \frac{n(n+1)}{4} \end{aligned}$$