Chp 14 : Nonparametrics Statistics :

I. : Sign Test : Single Population Inferences.

Two-Tailed Test : $\begin{cases} H_0 : \eta = \eta_0 \Rightarrow P(X \le \eta_0) = \frac{1}{2}. \\ H_A : \eta \ne \eta_0 \end{cases}$

Test ~~Statistic~~ $S = \max(\underbrace{S_1}_{\# < \eta_0}, \underbrace{S_2}_{\# > \eta_0})$.

P-value $= 2P(X \ge S)$ where $X \overset{H_0}{\sim} \text{Binomial}(n, \frac{1}{2})$.

One-Tailed Test : $\begin{cases} H_0 : \eta = \eta_0 \\ H_A : \eta > \eta_0 \quad (\text{or } H_a : \eta < \eta_0) \end{cases}$

Test Statistics : $S = \underbrace{S_2}_{\# > \eta_0} \quad (\text{or } \underbrace{S_1}_{\# < \eta_0})$.

(The rule : choose S_2 if it is expected to see a larger number of values $> \eta_0$ under H_a; otherwise S_1.)

P-value $= P(X \ge S)$ where $X \overset{H_0}{\sim} \text{Binomial}(n, \frac{1}{2})$.

Reject H_0 if P-value $\le \alpha$ - significance level.

Example : A paper company requires that the median height of pine trees exceed 40 feet before they are harvested. A sample of 9 trees in one large plot is selected and it resulted in the following data. Determine if the median height of trees exceeds 40 feet. ($\alpha = 0.05$).

STA 4165

52 58 64 41 39 45 48 46 5

>40(?)

Yes Yes Yes Yes No Yes Yes Yes Yes

$S_1 = \#$ of $>40 = \#$ of "Yes" $= 8$

$S_2 = \# < 40 = \#$ of "No" $= n - S_1 = 9 - 8 = 1$.

<a> $\begin{cases} H_0: \eta \le 40 \\ H_a: \eta > 40 \end{cases}$ $\Rightarrow$ results in a larger number of trees with height >40.

$\Rightarrow$ Test Statistic should be S_1.

P-value $= P(X \ge S_1) = P(X \ge 8)$ where $X \overset{H_0}{\sim} Bin(9, \frac{1}{2})$

$= 1 - P(X<8)$

$= 1 - P(X \le 7)$

$= 1 - .980$

$= .020$

or

$P(X=9) + P(X=8)$

$\binom{9}{8}\frac{1}{2}^{8}\frac{1}{2}^{(9-8)} + \binom{9}{9}\frac{1}{2}^{9}\frac{1}{2}^{(9-9)}$

$\frac{10}{2^9} \approx 0.02$

(hand calculation)

Since P-value $= .02 < \alpha$. Reject H_0.

<b> Consider a two-sided test: $\begin{cases} H_0: \eta = 40 \\ H_a: \eta \ne 40 \end{cases}$

$S = \max(S_1, S_2) = \max(8, 1) = 8$

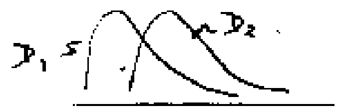
P-value $= 2P(X \ge 8) = 2 \times .02 = 0.04 < 0.05$

$\Rightarrow$ Reject H_0.

II : Wilcoxon Rank Sum Test : Two independent samples.

Two-Tailed Test:
$$\begin{cases} H_0: D_1 \text{ and } D_2 \text{ are identical} \\ H_a: D_1 \text{ is shifted either to the left or to the right of } D_2. \end{cases}$$

Test Statistic $T =$
$$\begin{cases} T_1 & \text{if } n_1 < n_2 \\ T_2 & \text{if } n_2 < n_1 \end{cases}$$



Rejection Region: $T \le T_L$ or $T \ge T_U$. using table XII.

One-Tailed Test:
$$\begin{cases} H_0: D_1 \,\&\, D_2 \text{ identical} \\ H_a: D_1 \text{ is shifted to the right (left) of } D \end{cases}$$

Test Statistic: $T = \begin{cases} T_1 & \text{if } n_1 < n_2 \\ T_2 & \text{if } n_2 < n_1 \end{cases}$

(Rule: Under H_a, we expect to see T_2 small, and large T
Therefore, rejection Region:

$T_1 \ge T_U$ and $T_2 \le T_L$.) :

Example: The annual percentage turnover rates for five U.S. and five Japanese plants are shown in the table.

U.S. Plants		Japanese Plants	
7.11%	9	3.52%	4
6.06%	7	2.02%	2
8.00%	10	4.91%	6
6.87%	8	3.22%	3
4.77%	5	1.92%	1
$T_{US} = 39$		$T_{JP} = 16$	
$n_1 = 5$		$n_2 = 5$.	

<a>: Test if the turnover rate of the US plants has a identic population distribution to the turnover rate of the Japanese plan

(Two-Tailed Test). $\begin{cases} H_0: \text{identical distributions.} \\ H_a: \text{shifted to the right/left.} \end{cases}$

Since $n_1 = n_2$, $T = T_{US}$ or T_{JP}.

From Table XII, reject H_0 if. $T \le 18$ or $T \ge 37$

$\begin{cases} n_1 = n_2 = 5. \\ \alpha = 0.05 \text{ (two-tailed)}. \end{cases}$ $\Rightarrow$ Reject H_0.

<b> Test if the U.S. plants has a population distribution of turnover rates that is shifted to the right of the Japanese Plants.

(One-Tailed Test) $\begin{cases} H_0: \text{identical} \\ H_a: \text{US} > \text{Japanese (right)} \end{cases}$

Since $n_1 = n_2$, $\Downarrow$

$T = T_{US}$ or T_{JP}. expect to see $\underline{T_{US} \text{ large}}$ and $\underline{T_{JP} \text{ small}}$

From Table XII. ($n_1 = n_2 = 5$. $\alpha = 0.05$ one-tailed), it gives

$T_L = 19$

$T_U = 36$

Suppose that $T = T_{US}$. then RR: reject H_0 if $\underline{T_{US} \ge T_U}$ 39

$\Rightarrow$ Reject H_0.

Suppose that $T = T_{JP}$ is used. then reject H_0 if $\underline{T_{JP} \le T_L}$ 16

$\Rightarrow$ Reject H_0

III: Wilcoxon Signed Rank Sum Test: Paired Samples.

Two-Tailed Test: $\begin{cases} H_0: D_1 \,\&\, D_2 \text{ identical} \\ H_a: \text{NOT identical.} \end{cases}$

Test Statistic T = minimum(T_+, T_-).

RR: $T \le T_0$ ~ given in Table XIII.

One-Tailed Test: $\begin{cases} H_0: D_1 \,\&\, D_2 \text{ identical} \\ H_a: D_1 \text{ is shifted to the right of } D_2 \; (D_1 > \\ \quad \cdots\cdots \text{ (left)} \cdots \; (D_1 < \end{cases}$

Test Statistic: $T = \begin{cases} T_+ & \text{if } T_+ \text{ is expected to be smaller than } T_- \text{ under } H_a. \\ T_- & \text{otherwise.} \end{cases}$

Reject H_0 if $T \le T_0$ ~ given in Table XIII.

Example: 10 customers were randomly selected to rate two product

Customer	Product A	Product B	Difference A − B	|A−B|	Rank
1	7	9	−2	2	4.5
2	4	5	−1	1	2
3	8	8	0	(Deleted)	NA
4	9	8	1	1	2
5	3	6	−3	3	6
6	6	10	−4	4	7.5
7	8	9	−1	1	2
8	10	8	2	2	4.5
9	9	4	5	5	9
10	5	9	−4	4	7.5

$n' = 10$

Positive rank sum $T_+ = 15.5$

Negative rank sum $T_- = 29.5$

n = # of untied pairs = 9

<a>: Test if there is significant diff between ratings of A & B.

$\begin{cases} H_0: \text{ identical} \\ H_a: \text{ different} \end{cases}$

$$T = \min(T_+, T_-) = \min(15.5, 29.5) = 15.5$$

Reject H_0 if $T \leq 6$ given in Table XIII. ($n=9$, $\alpha = 0.005$ two-tailed)

$\Rightarrow$ can NOT reject H_0 since $\underbrace{15.5}_{T\text{ observed}} > \underbrace{6}_{\text{critical value}}$.

<b>: Test if ratings of A are higher than B.

$\begin{cases} H_0: \text{ identical} \\ H_a: A > B \end{cases}$ $\Rightarrow$ $A - B > 0$ $\Rightarrow$ expect to see small T_-.

then $T = T_- = 29.5$

Critical value $= 8$ by Table XIII. ($n = 9$, $\alpha = 0.05$ one-tailed)

Reject H_0 if $T \leq 8$

$\Rightarrow$ can NOT reject H_0 since $T = 29.5 > 8$.

<u>Assumptions applicable to all three Tests:</u>

1. Random Sample (obs. are independent of each other);
2. Continuous probability distribution.