Introduction to Statistics and Data Science using *eStat*

Chapter 10 Nonparametric Testing Hypothesis

10.2 Nonparametric Test for Comparing Locations of Two Populations
10.2.2 Paired Sample

Jung Jin Lee
Professor of Soongsil University, Korea
Visiting Professor of ADA University, Azerbaijan

10.1 Nonparametric Test for Location of Single Population
- 10.1.1 Sign Test
- 10.1.2 Wilcoxon Signed Rank Sum Test

10.2 Nonparametric Test for Comparing Locations of Two Populations
- 10.2.1 Independent Samples: Wilcoxon Rank Sum Test
- 10.2.2 Paired Samples: Wilcoxon Signed Rank Sum Test

10.3 Nonparametric Test for Comparing Locations of Several Populations
- 10.3.1 Completely Randomized Design: Kruskal-Wallis Test
- 10.3.2 Randomized block design: Friedman Test

10.2 Nonparametric Test for Location Parameters of Two Populations

10.2.2 Paired Sample - Wilcoxon Signed Rank Sum Test

Table 10.2.6 Data of differences for paired samples

Pair number	Sample of population 1 (x_{i1})	Sample of population 2 (x_{i2})	Difference $d_i = x_{i1} - x_{i2}$
1	x_{11}	x_{12}	$d_1 = x_{11} - x_{12}$
2	x_{21}	x_{22}	$d_2 = x_{21} - x_{22}$
...	...	...	...
n	x_{n1}	x_{n2}	$d_n = x_{n1} - x_{n2}$

10.2 Nonparametric Test for Location Parameters of Two Populations

10.2.2 Paired Sample - Wilcoxon Signed Rank Sum Test

[Example 10.2.2] The following is the survey result of eight samples from young couples. The husband's age and wife's age of each couple are recorded.

(28, 28) (30, 29) (34, 31) (29, 32) (28, 29) (31, 33) (39, 35) (34, 29)

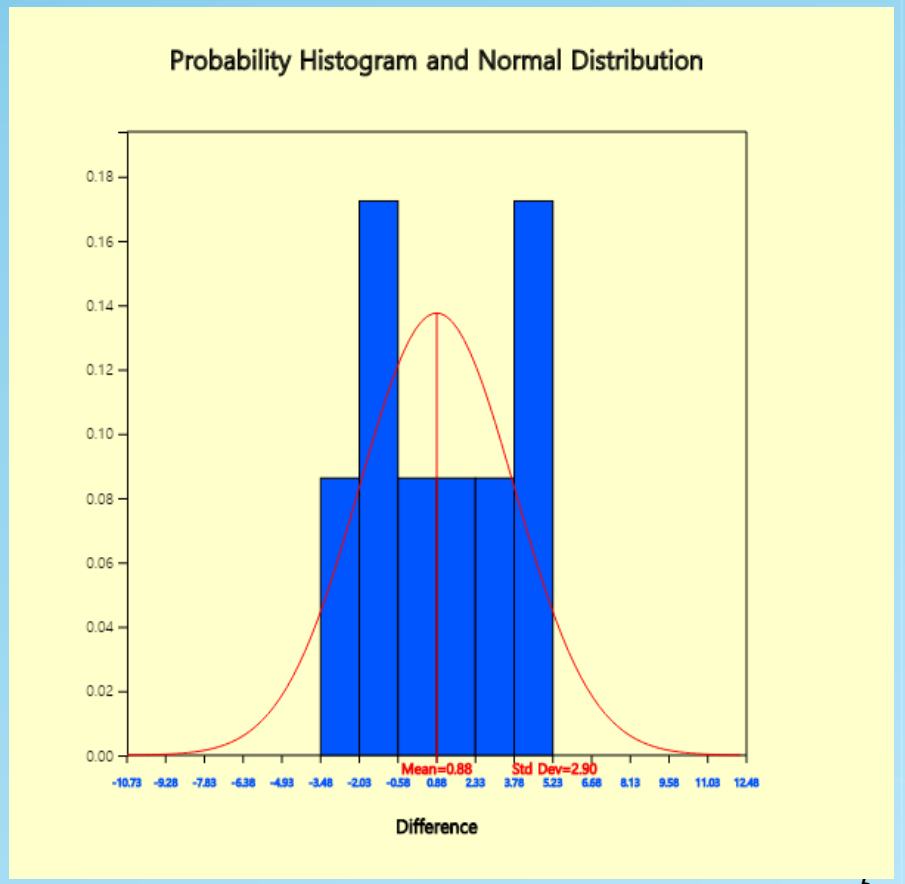
1) Calculate data of differences in each pair and draw their histogram to check whether a parametric test is applicable or not.
2) Apply the Wilcoxon signed rank sum test to see whether the husband's age is greater than the wife's age with the significance level of 0.05.
3) Check the result of the above signed rank sum test using 『eStat』.

10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.2>

- **Age difference between husband and wife**

Pair Number	Husband (x_i)	Wife (y_i)	Difference $d_i = x_i - y_i$
1	28	28	0
2	30	29	1
3	34	31	3
4	29	32	-3
5	28	29	-1
6	31	33	-2
7	39	35	4
8	34	29	5



10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.2>

- **Hypothesis** $H_o: M_1 = M_2$ $H_1: M_1 \neq M_2$

  $H_o: M_d = 0$ $H_1: M_d \neq 0$

Difference data	1	3	-3	-1	-2	4	5
Sign data	+	+	-	-	-	+	+
|data – 0|	1	3	3	1	2	4	5
Rank of |data – 0|	1.5	4.5	4.5	1.5	3	6	7
Rank sum of '+' sign ($R_+ = 19$)	1.5 + 4.5 + 6 + 7						

Signed Rank Sum	n = 7		
x	P(X = x)	P(X≤x)	P(X≥x)
0	0.0078	0.0078	1.0000
1	0.0078	0.0156	0.9922
2	0.0078	0.0234	0.9844
3	0.0156	0.0391	0.9766
…	…	…	…
25	0.0156	0.9766	0.0391
26	0.0078	0.9844	0.0234
27	0.0078	0.9922	0.0156
28	0.0078	1.0000	0.0078

- **Since P(X ≤ 2) = 0.02234, P(X ≥ 26) = 0.0234,**
- **decision rule: 'If $R_+ \leq 2.5$ or $R_+ \geq 25.5$, then reject H_o'**
- **In this problem $R_+ = 19$, we can not reject H_o.**

10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.2>

File EX100202_

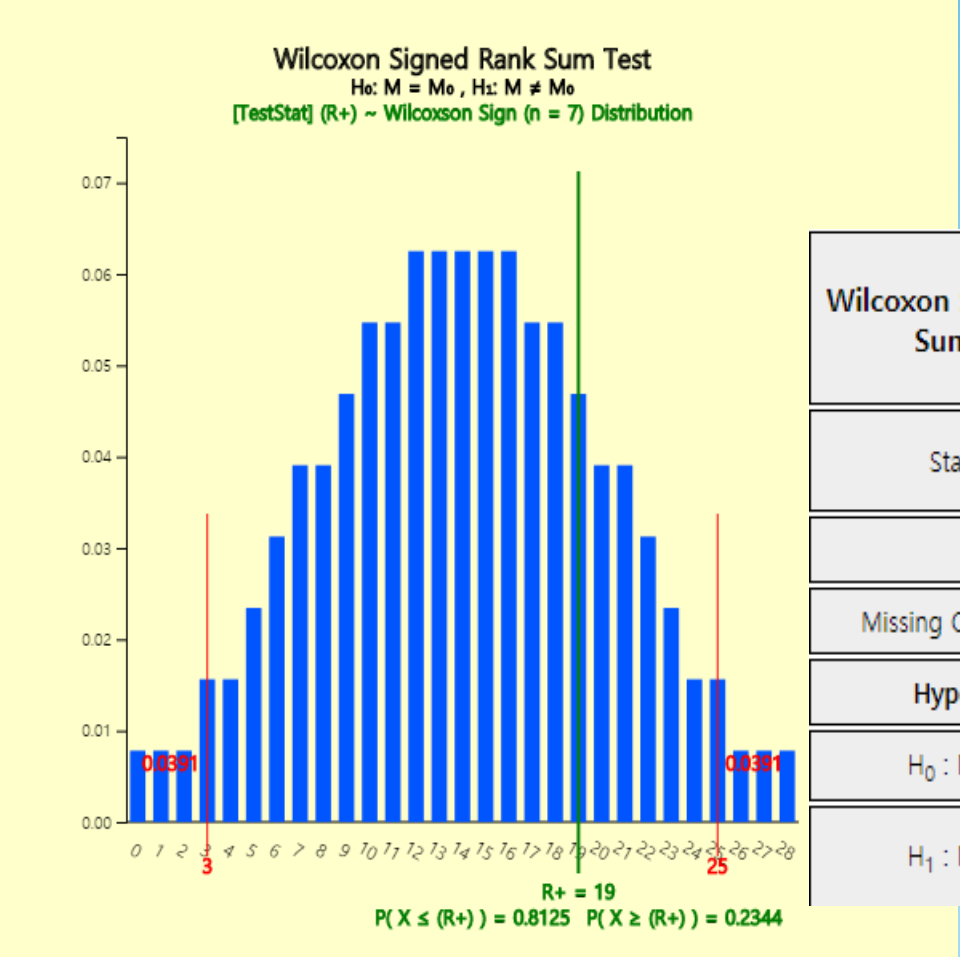
Analysis Var

1: Difference

(Selected data: Raw Data)

SelectedVar V1

	Differenc	V2
1	0	
2	1	
3	3	
4	-3	
5	-1	
6	-2	
7	4	
8	5	



Wilcoxon Signed Rank Sum Test	Analysis Var	Difference			
Statistics	Total Observation	Test Observation	-Group Observation	+Group Observation	
	8	7	3	4	
Missing Observations	0				
Hypothesis					
$H_0 : M = M_0$	μ_0	[TestStat]	Rank Sum R+	P(X ≤ R+)	P(X ≥ R+)
$H_1 : M \neq M_0$	0.00	+Group Rank Sum(R+)	19.00	0.8125	0.2344

10.2 Nonparametric Test for Location Parameters of Two Populations

Table 10.2.9 Wilcoxon signed rank sum test for paired samples

Type of Hypothesis	Decision Rule Test Statistic: R_+ = 'sum of ranks on $\lvert d_i \rvert$ with + sign'
1) $H_0 : M_d = 0$ $H_1 : M_d > 0$	If $R_+ > w_+(n)_\alpha$, then reject H_0, else accept H_0
2) $H_0 : M_d = 0$ $H_1 : M_d < 0$	If $R_+ < w_+(n)_{1-\alpha}$, then reject H_0, else accept H_0
3) $H_0 : M_d = 0$ $H_1 : M_d \neq 0$	If $R_+ < w_+(n)_{1-\alpha/2}$ or $R_+ > w_+(n)_{\alpha/2}$, then reject H_0, else accept H_0

- $w_+(n)$: Distribution of + rank sum of $|x_i - M_0|$
- If any of the observed values has the same value as M_0 , they are not used in test.

Thank you