Introduction to Statistics and Data Science using *eStat*

Chapter 10 Nonparametric Testing Hypothesis

10.2 Nonparametric Test for Comparing Locations of Two Populations

10.2.1 Wilcoxon Rank Sum Test

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

10.2.1 Wilcoxon Rank Sum Test

[Example 10.2.1] A professor teaches the Statistics courses to students in the Department of Economics and the Department of Management.
In order to compare exam scores of students in the two departments, 7 students were randomly sampled from the Economics Department and 6 students from the Management Department as follows:

Department of Economics	87 75 65 95 90 81 93
Department of Management	57 85 90 83 87 71

1) Draw a histogram of the data to verify that the testing hypothesis can be performed using a parametric method.
2) Apply the Wilcoxon rank sum test with the significance level of 5%.
3) Check the result of the Wilcoxon rank sum test using 『eStat』.

10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.1>

- **Hypothesis** $H_o : \mu_1 = \mu_2 \quad H_1 : \mu_1 \neq \mu_2$

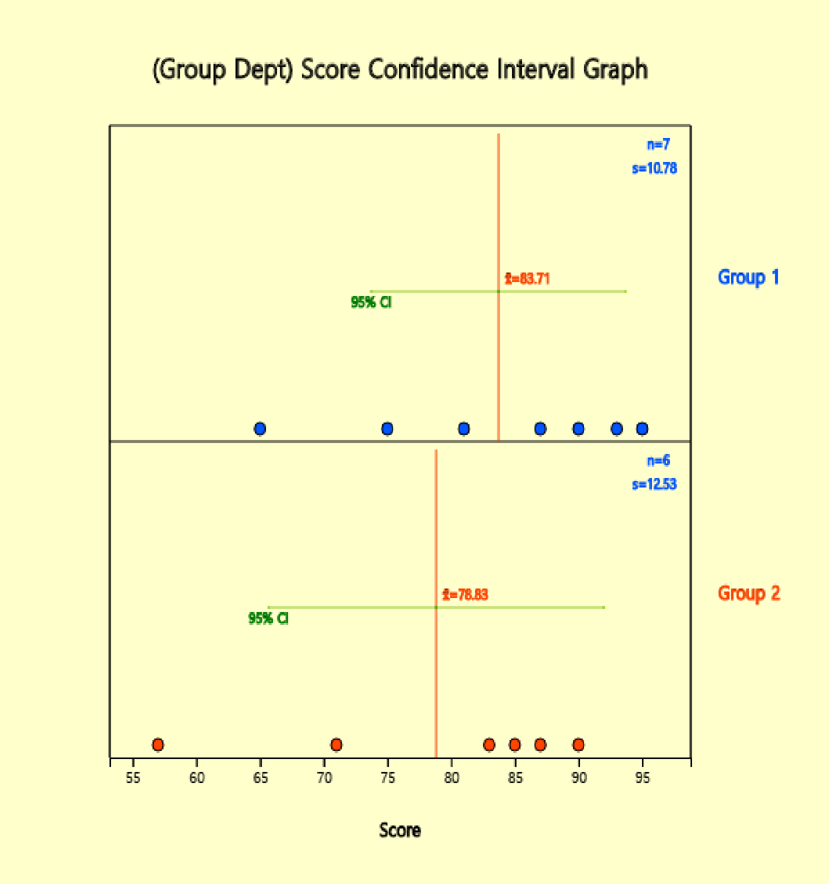
File: EX100201_ScoreByDept.csv

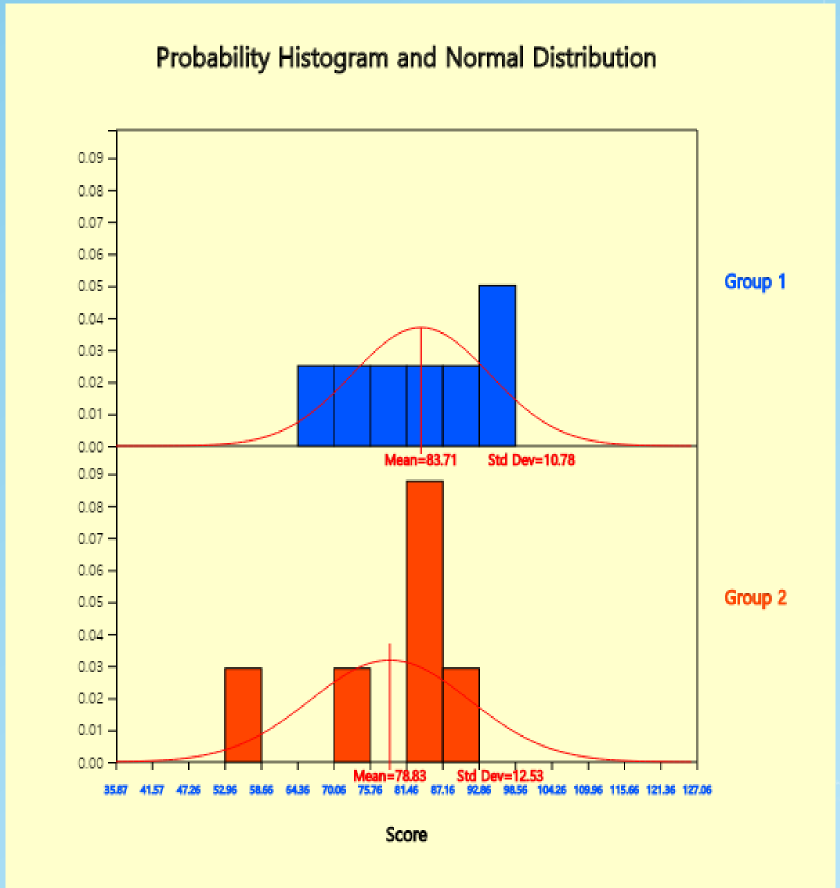
Analysis Var: 2: Score — by Group: 1: Dept

(Selected data: Raw Data) (or Paired Var)

SelectedVar: V2 by V1,

	Dept	Score	V3	V4
1	1	87		
2	1	75		
3	1	65		
4	1	95		
5	1	90		
6	1	81		
7	1	93		
8	2	57		
9	2	85		
10	2	90		
11	2	83		
12	2	87		
13	2	71		





10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.1>

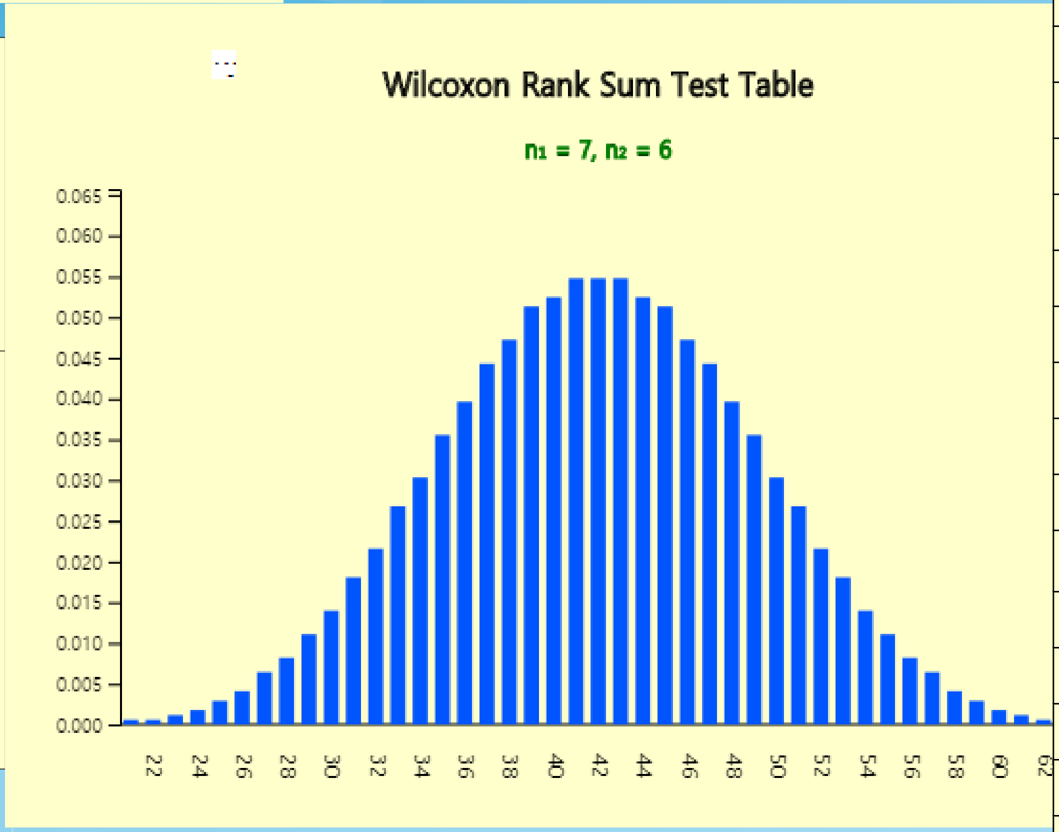
- Hypothesis $H_o : M_1 = M_2 \quad H_1 : M_1 \neq M_2$

Sorted Data of Sample 1	Sorted Data of Sample 2	Ranks of Sample 1	Ranks of Sample 2
	57		1
65		2	
	71		3
75		4	
81		5	
	83		6
	85		7
87	87	8.5	8.5
90	90	10.5	10.5
93		12	
95		13	
	Sum of ranks	$R_1 = 55$	$R_2 = 36$

10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.1>

All possible permutation of ranks	Sum of ranks, R_2
{1,2,3,4,5,6}	21
{1,2,3,4,5,7}	22
…	…
{8,9,10,11,12,13}	63



Rank Sum	n_1 = 7	n_2 = 6	
x	P(X=x)	P(X≤x)	P(X≥x)
21	0.0006	0.0006	1
22	0.0006	0.0012	0.9994
23	0.0012	0.0023	0.9988
24	0.0017	0.0041	0.9977
25	0.0029	0.007	0.9959
26	0.0041	0.0111	0.993
27	0.0064	0.0175	0.9889
28	0.0082	0.0256	0.9825
29	0.0111	0.0367	0.9744
…	…	…	…
55	0.0111	0.9744	0.0367
56	0.0082	0.9825	0.0256
57	0.0064	0.9889	0.0175
58	0.0041	0.993	0.0111
59	0.0029	0.9959	0.007
60	0.0017	0.9977	0.0041
61	0.0012	0.9988	0.0023
62	0.0006	0.9994	0.0012
63	0.0006	1	0.0006

- **Since P(X ≤ 28) = 0.0256, P(X ≥ 56) = 0.0256,**
- **decision rule: 'If $R_2 \le 27.5$ or $R_2 \ge 56.5$, then reject H_o'**
- **In this problem R_2 = 36, we can not reject H_o.**

10.2 Nonparametric Test for Location Parameters of Two Populations

<Answer of Example 10.2.1>

Wilcoxon Rank Sum Test : Location Parameter M_1, M_2

Menu

[Hypothesis] $H_0 : M_1 = M_2$

(◉) $H_1 : M_1 \neq M_2$ ( ) $H_1 : M_1 > M_2$ ( ) $H_1 : M_1 < M_2$

[Test Type] Rank Sum Test

Significance Level α = (◉) 5% ( ) 1%

[Sample Data] *If $n=n_1+n_2 \leq 25$ Wilcoxon Rank Sum Test, $n>25$ Normal Approximation Test*

Sample 1: 87 75 65 95 90 81 93

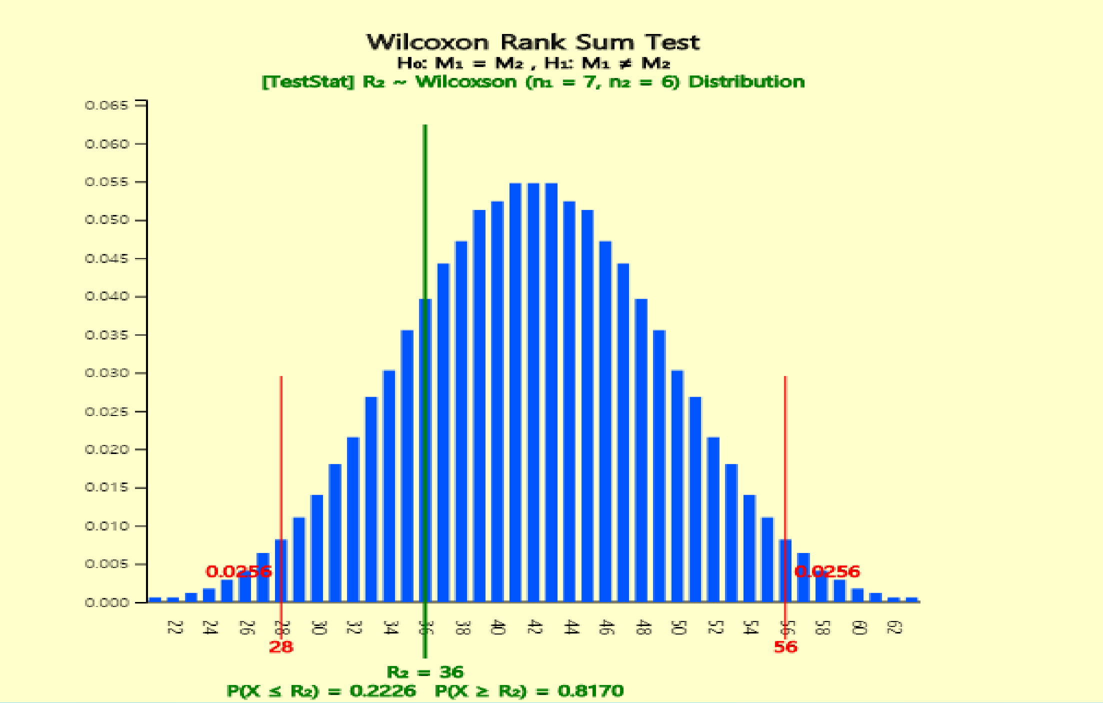
Sample 2: 57 85 90 83 87 71

[Sample Statistics]

Sample Size n_1 = 7 n_2 = 6

Rank Sum R_1 = 55 R_2 = 36

Execute



Wilcoxon Rank Sum Test	Analysis Var	Score			
Statistics	Observation	Mean	Std Dev	Rank Sum	
1 (Group 1)	7	83.714	10.781	55.00	
2 (Group 2)	6	78.833	12.529	36.00	
Total	13	81.462	11.399	91.00	
Missing Observations	0				
Hypothesis					
$H_0 : M_1 - M_2 = D$	D	[TestStat]	Group 2 Rank Sum R_2	$P(X \leq R_2)$	$P(X \geq R_2)$
$H_1 : M_1 - M_2 \neq D$	0.00	Group 2 Rank Sum(R_2)	36.00	0.2226	0.8170

10.2 Nonparametric Test for Location Parameters of Two Populations

- $w_2(n_1, n_2)$: Wilcoxon rank sum distribution of R_2 with sample size n_1 and n_2

Table 10.2.4 Wilcoxon rank sum test

Type of Hypothesis	Decision Rule Test Statistic: R_2 = 'Sum of ranks assigned samples of Y'
1) $H_0 : M_1 = M_2$ $H_1 : M_1 > M_2$	If $R_2 > w_2(n_1, n_2)_{\alpha}$, then reject H_0, else accept H_0
2) $H_0 : M_1 = M_2$ $H_1 : M_1 < M_2$	If $R_2 < w_2(n_1, n_2)_{1-\alpha}$, then reject H_0, else accept H_0
3) $H_0 : M_1 = M_2$ $H_1 : M_1 \neq M_2$	If $R_2 < w_2(n_1, n_2)_{1-\alpha/2}$ or $R_2 > w_2(n_1, n_2)_{\alpha/2}$, then reject H_0, else accept H_0

❖ If there is a tie in the combined sample, assign the average rank.

10.2 Nonparametric Test for Location Parameters of Two Populations

Table 10.2.5 Wilcoxon rank sum test (large sample case)

Type of Hypothesis	Decision Rule Test Statistic: R_2 = 'Sum of ranks assigned samples of Y'
1) $H_0 : M_1 = M_2$ $H_1 : M_1 > M_2$	If $\frac{R_2 - E(R_2)}{\sqrt{V(R_2)}} > z_\alpha$, then reject H_0, else accept H_0
2) $H_0 : M_1 = M_2$ $H_1 : M_1 < M_2$	If $\frac{R_2 - E(R_2)}{\sqrt{V(R_2)}} < -z_\alpha$, then reject H_0, else accept H_0
3) $H_0 : M_1 = M_2$ $H_1 : M_1 \neq M_2$	If $\left\| \frac{R_2 - E(R_2)}{\sqrt{V(R_2)}} \right\| > z_{\alpha/2}$, then reject H_0, else accept H_0

- $\mathbf{E}(R_2) = \frac{n_1(n_1+n_2+1)}{2}$, $\mathbf{V}(R_2) = \frac{n_1 n_2(n_1+n_2+1)}{12}$

Thank you