BSc III Year

Paper Code: 394 (Non parametric Methods and Numerical Analysis)

Non parametric tests

Unit 1

NONPARAMETRIC STATISTICS

In general, a statistical technique is categorized as NPS if it has at least one of the following characteristics:

1. The method is used on nominal data
2. The method is used in ordinal data
3. The method is used in interval scale or ratio scale data but there is no assumption regarding the probability distribution of the population where the sample is selected.

❑ Sign Test

❑ Mann-Whitney Test

❑ Kruskal Wallis Test

❑ Wilcoxon Signed Rank Test

❑ Spearman's Rank Correlation Test

Sign Test

- The sign test is used to test the null hypothesis and whether or not two groups are equally sized.

- In other word, to test of the population proportion for testing $p = 0.5$ in a small sample (usually $n \leq 20$)

- It based on the direction of the + and − sign of the observation and not their numerical magnitude.

- It also called the binomial sign test with the null proportion is 0.5 (Uses the binomial distribution as the decision rule).

A binomial experiment consist of n identical trial with probability of success, p in each trial. The probability of x success in n trials is given by

$$P(X = x) = {}^{n}C_{x} p^{x} q^{n-x}; \qquad x = 0, 1, 2.... n$$

$$\text{where } {}^{n}C_{x} = \binom{n}{x} = \frac{n!}{(n-x)!x!}$$

$$X \sim B(n, p) \text{ where } p = 0.5$$

- There are two types of sign test :
  1. One sample sign test
  2. Paired sample sign test

<u>One Sample Sign Test</u>

Procedure:
1. Put a + sign for a value greater than the mean value
    Put a - sign for a value less than the mean value
    Put a 0 as the value equal to the mean value

2. Calculate:
    i.   The number of + sign, denoted by x
    ii.   The number of sample, denoted by n (discard/ignore the data with value 0)

3. Run the test
    i.   State the null and alternative hypothesis
    ii.   Determine level of significance, α
    iii. Reject H_0 if $p - \text{value} < \alpha$
    iv.   Determining the $p - $ value for the test for $n,\ x$ and $p = 0.5$, from binomial probability table base on the type of test being conducted

	Sign of H_0	Sign of H_1	p - value
Two tail test	=	$\neq$	if $x < \dfrac{n}{2}$: $2P(X \leq x)$ if $x > \dfrac{n}{2}$: $2P(X \geq x)$
Right tail test	$\leq$	>	$P(X \geq x) = 1 - P(X \leq x-1)$
Left tail test	$\geq$	<	$P(X \leq x)$

v.   Make a decision

Example:

The following data constitute a random sample of 15 measurement of the octane rating of a certain kind gasoline:

| 99.0 | 102.3 | 99.8 | 100.5 | 99.7 | 96.2 | 99.1 | 102.5 | 103.3 | 97.4 | 100.4 |
| | | | | 98.9 | 98.3 | 98.0 | 101.6 | | | |

Test the null hypothesis $\tilde{\mu} \leq 98.0$ against the alternative hypothesis $\tilde{\mu} > 98.0$ at the 0.01 level of significance.

Solution:

$\tilde{\mu}_0 = 98.0$

| 99.0 | 102.3 | 99.8 | 100.5 | 99.7 | 96.2 | 99.1 | 102.5 | 103.3 | 97.4 | 100.4 |
| + | + | + | + | + | - | + | + | + | - | + |

| 98.9 | 98.3 | 98.0 | 101.6 |
| + | + | 0 | + |

Number of + sign, x = 12
Number of sample, n = 14 $\longrightarrow$ (15 -1)
$p = 0.5$

1. $H_0 : \mu \leq 98.0$
   $H_1 : \mu > 98.0$

2. $\alpha = 0.01,$ Reject H_o if $p-\text{value} < 0.01$

3. From binomial probability table for $x = 12,\, n = 14$ and $p = 0.5$

   $X \sim b(14, 0.5),\; p-\text{value} = P(X \geq 12) = 1 - P(X \leq 11) = 1 - 0.9935 = 0.0065$

4. Since $p-\text{value} = 0.0065 < 0.01 = \alpha,$ thus we reject H_0 and accept H_1 and conclude that the median octane rating of the given kind of gasoline exceeds 98.0

Paired Sample Sign Test

Procedure:

1. Calculate the difference, $d_i = x_i - y_i$ and record the sign of d_i

2. i. Calculate the number of + sign and denoted as x

   ii. The number of sample, denoted by n (discard/ignore data with value 0) *probability is 0.5 ($p = 0.5$)

3. Run the test

   i. State the null hypothesis and alternative hypothesis

   ii. Determine the level of significance α

   iii. Reject H_0 if p value $< \alpha$

   iv. Determining the p value for the test for n, x and p = 0.5 from binomial probability table base on type of test being conducted.

	Sign of H_0	Sign of H_1	p - value
Two tail test	=	$\neq$	if $x < \dfrac{n}{2}$: $\quad 2P(X \leq x)$ if $x > \dfrac{n}{2}$: $\quad 2P(X \geq x)$
Right tail test	$\leq$	>	$P(X \geq x) = 1 - P(X \leq x - 1)$
Left tail test	$\geq$	<	$P(X \leq x)$

v.   Make a decision

<u>Example:</u>

10 engineering students went on a diet program in an attempt to loose weight with the following results:

Name	Weight before	Weight after
Abu	69	58
Ah Lek	82	73
Sami	76	70
Kassim	89	71
Chong	93	82
Raja	79	66
Busu	72	75
Wong	68	71
Ali	83	67
Tan	103	73

Is the diet program an effective means of losing weight? Do the test at significance level $\alpha = 0.10$

Solution:

Let the sign + indicates weight before − weight after > 0 and − indicates weight before − weight after < 0

Thus

Name	Weight before	Weight after	Sign
Abu	69	58	+
Ah Lek	82	73	+
Sami	76	70	+
Kassim	89	71	+
Chong	93	82	+
Raja	79	66	+
Busu	72	75	-
Wong	68	71	-
Ali	83	67	+
Tan	103	73	+

1. The + sign indicates the diet program is effective in reducing weight

$$H_0 : p \leq 0.5$$
$$H_1 : p > 0.5$$

2. $\alpha = 0.10$. So we reject H_0 if p value < 0.10

3. Number of + sign, $x = 8$
   Number of sample, $n = 10$
   $p = 0.5$

$$P(X \geq 8) = 1 - P(X \leq 7) = 1 - 0.9453 = 0.0547$$

4. Since $p - \text{value} = 0.0547 < 0.10 = \alpha$. So we can reject H_0
   and we can conclude that there is sufficient evidence that the diet program
   is an effective programme to reduce weight.

Exercise:

A paint supplier claims that a new additive will reduce the drying time of its acrylic paint. To test his claim, 8 panels of wood are painted with one side of each panel with paint containing the new additive and the other side with paint containing the regular additive. The drying time, in hours, were recorded as follows:

Panel	Drying Times	
	New Additive	Regular Additive
1	6.4	6.6
2	5.8	5.8
3	7.4	7.8
4	5.5	5.7
5	6.3	6.0
6	7.8	8.4
7	8.6	8.8
8	8.2	8.4

Use the sign test at the 0.05 level to test the hypothesis that the new additive have the same drying time as the regular additive.

Mann-Whitney Test

- To determine whether a difference exist between two populations
- Sometimes called as Wilcoxon rank sum test
- Two independent random samples are required from each population. Let $m_1 \text{ and } m_2$ be the random samples of sizes n_1 and n_2 where $n_1 \leq n_2$ from population X and Y respectively

1. Null and alternative hypothesis

	Two tail test	Left tail test	Right tail test
H_0	$m_1 = m_2$	$m_1 \geq m_2$	$m_1 \leq m_2$
H_1	$m_1 \neq m_2$	$m_1 < m_2$	$m_1 > m_2$
Rejection area	$T \notin [T_L, T_U]$	$T < c_v$	$T > c_v$

$$c_v = [T_L, T_U] = \text{critical value} \qquad T_U = \text{upper critical value} \qquad T_L = \text{lower critical value}$$

Test statistic T:

- Designate the smaller size of the two sample as sample 1. If the sample are equal, either one or more may be designated as sample 1

- Rank the combined data value as if they were from a single group. The smallest data value gets a rank 1 and so on. In the event of tie, each of the tied get the average rank that the values are occupying.

- List the ranks for data values from sample 1 and find the sum of the rank for sample 1. Repeat the same thing to sample 2.

- Find $T_1 = \sum R_1$, the rank sum for the observation in sample 1. This is the test statistics for a left-tailed test.

- Find $T_1* = n_1 (n_1 + n_2 + 1) - T_1$, the sum of the ranks of the observations from population 1 if the assigned ranks had been reversed from large to small. This is the test statistics for a right-tailed test.

- The test statistic for a two-tailed test is T, $\text{Min}(T_1, T_1*)$.

Critical value of T

- The Mann-Whitney test/Wilcoxon rank sum table list lower and upper critical value for the test with n_1 and n_2 as the number of observations in the respective sample.

- The rejection region will be in either one or both tails depending on the null hypothesis being tested for n_1 and n_2 values.

- Compute the upper tail critical value, $T_U = n_1 \left(n_1 + n_2 + 1 \right) - T_L$.

- The value of T_L is read directly from the table of Mann-Whitney.

<u>Example:</u>

Data below show the marks obtained by electrical engineering students in an examination:

Gender	Marks
Male	60
Male	62
Male	78
Male	83
Female	40
Female	65
Female	70
Female	88
Female	92

Can we conclude the achievements of male and female students identical at significance level $\alpha = 0.1$

Solution:

Gender	Marks	Rank
Male	60	2
Male	62	3
Male	78	6
Male	83	7
Female	40	1
Female	65	4
Female	70	5
Female	88	8
Female	92	9

1. H_0 : Male and Female achievement are the same

   H_1 : Male and Female achievement are not the same

2. We have $n_1 = 4,\ n_2 = 5,\ T_1 = \sum R_1 = 2 + 3 + 6 + 7 = 18$

   $T_1^* = 4(4 + 5 + 1) - 18 = 22$

   $T = min(18, 22)$

   $T = 18$

3. From the table of Wilcoxon rank sum test for

$$\frac{\alpha}{2} = 0.05 \ , \ n_1 = 4, n_2 = 5,$$

so critical value, $\mathrm{T}_L = 13, \quad \mathrm{T}_U = 4(4+5+1)-13 = 27$

4. Reject H_0 if $T \notin [13, 27]$

5. Since $18 \in [13, 27]$, thus we fail to reject H_0 and conclude that the achievements of male and female are not significantly different.

Exercise:

Using high school records, Johnson High school administrators selected a random sample of four high school students who attended Garfield Junior High and another random sample of five students who attended Mulbery Junior High. The ordinal class standings for the nine students are listed in the table below. Test using Mann-Whitney test at 0.05 level of significance.

Garfield J. High		Mulbery J. High	
Student	Class standing	Student	Class standing
Fields	8	Hart	70
Clark	52	Phipps	202
Jones	112	Kirwood	144
TIbbs	21	Abbott	175
		Guest	146

Kruskal Wallis Test

- An extension of the Mann-Whiteny test or a.k.a Wilcoxon rank sum test of the previous section
- It compares more than two independent samples
- It is the non-parametric counterpart to the one way analysis of variance
- However, unlike one way ANOVA, it does not assume that sample have been drawn from normally distributed populations with equal variances

The null hypothesis and alternative hypothesis:

$$H_0 : m_1 = m_2 = \ldots = m_k \ \left(\text{the population median are equal}\right)$$

$$H_1 : \text{at least one } m_i \text{ differs from the others} \left(\text{the population median are not equal}\right)$$

Test statistic H

- Rank the combined data values if they were from a single group. The smallest data value gets a rank of 1, the next smallest, 2 and so on. In the event of tie, each of the tied values gets their average rank
- Add the rank from data values from each of the k group, obtaining

$$\sum R_1, \sum R_2, ..., \sum R_k$$

- The calculate value of the test statistics is:

$$H = \frac{12}{n(n+1)} \left[\sum_{i=1}^{k} \frac{(R_i)^2}{n_i} \right] - 3(n+1)$$

$n_i =$ the repective sample sizes for the k samples

$n = n_1 + n_2 + ... + n_k$

Critical value of H:

- The distribution of H is closely approximated by Chi-square distribution whenever each sample size at least 5, for α = the level of significance for the test, the critical H is the chi-square value for $df = k - 1$ and the upper tail area is α .

- We will reject H_0 if calculated $H >$ critical value $= \chi^2_{\alpha, df}$

Example:

Each of three aerospace companies has randomly selected a group of technical staff workers to participate in a training conference sponsored by a supplier firm. The three companies have sent 6, 5 and 7 employees respectively. At the beginning of the session. A preliminary test is given, and the scores are shown in the table below. At the 0.05 level, can we conclude that the median scores for the three population of technical staff workers could be the same?

Test score		
Firm 1	Firm 2	Firm 3
67	64	75
57	73	61
62	72	76
59	68	71
70	65	78
67		74
		79

Solution:

Test score					
Firm 1	Rank	Firm 2	Rank	Firm 3	Rank
67	7.5	64	5	75	15
57	1	73	13	61	3
62	4	72	12	76	16
59	2	68	9	71	11
70	10	65	6	78	17
67	7.5			74	14
				79	18
$\sum R_1$	32	$\sum R_2$	45	$\sum R_3$	94

1. $H_0 : m_1 = m_2 = m_3$

H_1 : at least one m_i differs from the others (the population medians are not equal)

2. $\alpha = 0.05$

$df = k - 1 = 3 - 1 = 2$

From χ^2 distribution table for $\alpha = 0.05$ and $df = 2$, critical value $= 5.991$ and we reject H_0 if $H >$ critical value

3. Calculated H :

$$H = \frac{12}{n(n+1)} \left[\sum_{i=1}^{k} \frac{(R_i)^2}{n_i} \right] - 3(n+1)$$

$$= \frac{12}{18(18+1)} \left[\frac{32^2}{6} + \frac{45^2}{5} + \frac{94^2}{7} \right] - 3(18+1) = 7.49$$

4. Since $H = 7.49 >$ critical value $= 5.99$, thus we rejected H_0 and conclude that the three population do not have the same median

Exercise:

Four groups of students were randomly assigned to be taught with four different techniques, and their achievement test scores were recorded. At the 0.05 level, are the distributions of test scores the same, or do they differ in location?

1	2	3	4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88

Wilcoxon Signed Rank Test

- Can be applied to two types of sample: one sample or paired sample
- For one sample, this method tests whether the sample could have been drawn from a population having a hypothesized value as its median
- For paired sample, to test whether the two populations from which these samples are drawn identical.

The Wilcoxon Signed rank test for one sample

- Null and alternative hypothesis:

Case	H_0	H_1	Rejection region
Two tail	H_0: median $R(d) = m_0$	H_1 : median $R(d) \neq m_0$	$\min\left(T^+, T^-\right) \leq a, \dfrac{\alpha}{2}$
Right tail	H_0 : median $R(d) \leq m_0$	H_1 : median $R(d) > m_0$	$T^- \leq a$
Left tail	H_0 : median $R(d) \geq m_0$	H_1 : median $R(d) < m_0$	$T^+ \leq a$

- Test procedure:
  i. For each of the observed values, find the difference between each value and the median; $d_i = x_i - m_0$

  where $m_0 =$ median value that has been specified

  ii. Ignoring the observation where $d_i = 0$, rank the $|d_i|$ values so the smallest $|d_i|$ will have a rank of 1. Where two or more differences have the same value find their mean rank, and use this.

  iii. For observation where $x_i > m_0$, list the rank as $+R(d_i)$ column and $x_i < m_0$ list the rank as $-R(d_i)$ column

iv. Then, sum the ranks of the positive differences, T^+ and sum the ranks of the negative differences T^-

.

$$T^+ = \sum +R(d_i), \quad T^- = \sum -R(d_i)$$

iv. The test statistic, W is the depends on the alternative hypothesis:
- For a two tailed test the test statistic $W = \min\left(T^+, T^-\right)$
- For a one tailed test where the H_1: median $R(d_i) > m_0$ the test statistic, $W = T^-$
- For a one tailed test where the H_1: median $R(d_i) < m_0$ the test statistic $W = T^+$

- Critical region:
  Compare the test statistic, W with the critical value in the tables; the null hypothesis is rejected if $W \leq$ critical value, a

- Make a decision

Example:

An environmental activist believes her community's drinking water contains at least the 40.0 parts per million (ppm) limit recommended by health officials for a certain metal. In response to her claim, the health department samples and analyzes drinking water from a sample of 11 households in the community. The results are as in the table below. At the 0.05 level of significance, can we conclude that the community's drinking water might equal or exceed the 40.0 ppm recommended limit?

Household	Observed concentration x_i
A	39
B	20.2
C	40
D	32.2
E	30.5
F	26.5
G	42.1
H	45.6

I	42.1
J	29.9
K	40.9

Solution:

$m_0 = 40$

Household	Observed concentration x_i	$d_i = x_i - m_0$	$\lvert d_i \rvert$	Rank, $R(d_i)$	$+R(d_i)$	$-R(d_i)$
A	39	-1	1	2		2
B	20.2	-19.8	19.8	10		10
C	40	0	—	—		
D	32.2	-7.8	7.8	6		6
E	30.5	-9.5	9.5	7		7
F	26.5	-13.5	13.5	9		9
G	42.1	2.1	2.1	3.5	3.5	
H	45.6	5.6	5.6	5	5	
I	42.1	2.1	2.1	3.5	3.5	
J	29.9	-10.1	10.1	8		8
K	40.9	0.9	0.9	1	1	
				Σ	$T^+ = 13$	$T^- = 42$

1. H_0: median of $R(d) \geq 40$

   H_1: median of $R(d) < 40$

   (One tail test)

   $\alpha = 0.05, \ n = 10$

2. Based on the alternative hypothesis, the test statistic $T^+ = \sum +R(d_1) = 13$

3. From table of Wilcoxon signed rank for one tail test,

   $\alpha = 0.05, \ n = 10,$ critical value, $a = 11$

   We will reject H_0 if $T^+ \leq a$

5. Since $T^+ = 13 > a = 11$, thus we failed to reject H_0 and conclude that the city's water supply might have at least 40.0 ppm of the metal

Student satisfaction surveys ask students to rate a particular course, on a scale of 1 (poor) to 10 (excellent). In previous years the replies have been symmetrically distributed about a median of 4. This year there has been a much greater on-line element to the course, and staff want to know how the rating of this version of the course compares with the previous one.
14 students, randomly selected, were asked to rate the new version of the course and their ratings were as follows:

1 3 6 4 8 2   3 6   5 2 3 4 1   2

Is there any evidence at the 5% level that students rate this version any differently?

The Wilcoxon Signed rank test for paired sample

- Null and alternative hypothesis:

Case	H_0	H_1	Rejection region
Two tail	H_0: median $R(d) = 0$	H_1: median $R(d) \neq 0$	$\min \left(T^+, T^- \right) \leq a$
Right tail	H_0: median $R(d) \leq 0$	H_1: median $R(d) > 0$	$T^- \leq a$
Left tail	H_0: median $R(d) \geq 0$	H_1: median $R(d) < 0$	$T^+ \leq a$

- Test procedure:
    i.   For each of the observed values, calculate $d_i = x_i - y_i$
    ii.   Ignoring observation where $d_i = 0$, rank the $|d_i|$ values so the smallest $|d_i|$ will have a rank of 1. Where two or more differences have the same value find their mean rank, and use this.
    iii. For observation where $x_i - y_i > 0$, list the rank as $+R(d_i)$ column and $x_i - y_i < 0$ list the rank as $-R(d_i)$ column

-

Then, sum the ranks of the positive differences, T^+ and sum the ranks of the negative differences T^-

$$T^+ = \sum +R(d_i), \quad T^- = \sum -R(d_i)$$

iv. The test statistic, W is the depends on the alternative hypothesis:

- For a two tailed test the test statistic $W = \min\left(T^+, T^-\right)$

- For a one tailed test where the $H_1:$ median $R(d) > 0$ the test statistic, $W = T^-$

- For a one tailed test where the $H_1:$ median $R(d) < 0$ the test statistic $W = T^+$

- Critical region:

Compare the test statistic, W with the critical value in the tables; the null hypothesis is rejected if $W \leq$ critical value, a

- Make a decision

Example:

Two computer software packages are being considered for use in the inventory control department of a small manufacturing firm. The firm has selected 12 different computing task that are typical of the kinds of jobs. The results are shown in the table below. At the 0.10 level, can we conclude that the median difference for the population of such task might be zero?

Computing task	Time required for software packages	
	x_i	y_i
A	24	23.1
B	16.7	20.4
C	21.6	17.7
D	23.7	20.7
E	37.5	42.1
F	31.4	36.1
G	14.9	21.8
H	37.3	40.3
I	17.9	26
J	15.5	15.5
K	29	35.4
L	19.9	25.5

Solution:

Computing task	Time required for software packages		$d_i = x_i - y_i$	$\lvert d_i \rvert$	Rank, $R(d_i)$	$+R(d_i)$	$-R(d_i)$
	x_i	y_i					
A	24	23.1	0.9	0.9	1	1	
B	16.7	20.4	-3.7	3.7	4		4
C	21.6	17.7	3.9	3.9	5	5	
D	23.7	20.7	3	3	2.5	2.5	
E	37.5	42.1	-4.6	4.6	6		6
F	31.4	36.1	-4.7	4.7	7		7
G	14.9	21.8	-6.9	6.9	10		10
H	37.3	40.3	-3	3	2.5		2.5
I	17.9	26	-8.1	8.1	11		11
J	15.5	15.5	0	0	—		—
K	29	35.4	-6.4	6.4	9		9
L	19.9	25.5	-5.6	5.6	8		8
					$\sum$	$T^+ = 8.5$	$T^- = 57.5$

1. H_0: median of $R(d) = 0$

   H_1: median of $R(d) \neq 0$ (two tail test)

2. Based on the alternative hypothesis, the test is $\min(T^+, T^-) = \min(8.5, 57.5) = 8.5$
3. $\alpha = 0.10, \ n = 12 - 1 = 11$
4. From table of Wilcoxon signed rank for two tail test,

   $\alpha = 0.10, \ n = 11, \ \text{then } a = 14$

   We will reject H_0 if $\min(T^+, T^-) \leq a$

5. Since $\min(8.5, 57.5) = 8.5 \leq 14$, thus we reject H_0 and conclude that the software packages are not equally rapid in handling computing tasks like those in the sample, or the population median for $d_i = x_i - y_i$ is not equal to zero and that package x is faster than package y in handling computing task like ones sample.

Exercise:

Spearman's Rank Correlation Test

- We have seen the correlation coefficient r measure the linear relationship between two continuous variable X and Y

- Spearman's Rank Correlation Test is used to measure the strength and the direction of the relationship between two variables which are at least ordinal data.

- A measure of correlation for ranked data based on the definition of Pearson Correlation where there is no tie or few ties called Spearman rank Correlation Coefficient, denoted by

$$r_s = 1 - \frac{6T}{n(n^2 - 1)}$$

where $T = \sum_{i=1}^{n}(d_i^2) = \sum_{i=1}^{n}\left[R(X_i) - R(Y_i)\right]^2$

$R(X_i)$ is the rank assigned to x_i

$R(Y_i)$ is the rank assigned to y_i

d_i is the difference between the ranks assigne to x_i and y_i

n is the number of pairs of data

- A value of +1 or -1 indicated perfect association between X and Y
- The plus sign with value $r_s > 0.5$ indicates strong positive correlation between the x and y, and $r_s < 0.5$ indicates weak positive correlation between the x and y
- The minus sign with value $r_s > 0.5$ indicates strong negative correlation between the x and y, and $r_s < 0.5$ indicates weak negative correlation between the x and y
- When r_s is zero or close to zero, we would conclude that the variable are uncorrelated

The data below show the effect of the mole ratio of sebacic acid on the intrinsic viscosity of copolyesters.

Mole ratio	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3
Viscosity	0.45	0.20	0.34	0.58	0.70	0.57	0.55	0.44

Find the Spearman rank correlation coefficient to measure the relationship of mole ratio of sebacic acid and the viscosity of copolyesters.

Solutions:

X: mole ratio of sebacic
Y: viscosity of copolyesters

Mole ratio	Viscosity	$R(x_i)$	$R(y_i)$	$d_i = R(x_i) - R(y_i)$	d_i^2
1.0	0.45	8	4	4	16
0.9	0.20	7	1	6	36
0.8	0.34	6	2	4	16
0.7	0.58	5	7	-2	4
0.6	0.70	4	8	-4	16
0.5	0.57	3	6	-3	9
0.4	0.55	2	5	-3	9
0.3	0.44	1	3	-2	4
					T = 110

Thus
$$r_s = 1 - \frac{6T}{n(n^2 - 1)} = 1 - \frac{6(110)}{8(64 - 1)} = -0.3095$$

which shows a weak negative correlation between the mole ratio of sebacic acid and the viscosity of copolyesters

Exercise:

The following data were collected and rank during an experiment to determine the change in thrust efficiency, y as the divergence angle of a rocket nozzle, x changes:

Rank X	1	2	3	4	5	6	7	8	9	10
Rank Y	2	3	1	5	7	9	4	6	10	8

Find the Spearman rank correlation coefficient to measure the relationship between the divergence angle of a rocket nozzle and the change in thrust efficiency.