Solution

X	Y	rank X	rank Y	d	d^2
53	47	8	1	7	49
98	25	1	8	−7	49
95	32	2	6	−4	16
81	37	3	5	−2	4
75	30	4	7	−3	9
71	40	5	3	2	4
59	39	6	4	2	4
55	45	7	2	5	25
					$\sum d^2 = 160$

$$\rho_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6 \times 160}{8(64-1)} = -0.9048$$

There is very high negative correlation between X and Y.

Solved Problem 4.37 ***AU, May/June 2006***

Calculate the coefficients of rank correlation from the following data.

X	:	48	34	40	12	16	16	66	25	16	57
Y	:	15	15	24	8	13	6	20	9	9	15

Solution

X	Y	rank X	rank Y	d	d^2
48	15	3	4	−1	1
34	15	5	4	1	1
40	24	4	1	3	9
12	8	10	9	1	1
16	13	8	6	2	4
16	6	8	10	−2	4
66	20	1	2	−1	1
25	9	6	7.5	−1.5	2.25
16	9	8	7.5	0.5	0.25
57	15	2	4	−2	4
					$\sum d^2 = 27.50$

In X-series, the value 16 is repeated three times.

In Y-series, the value 15 is repeated three times and the value 9 is repeated two times.

$$\therefore \quad \rho_s = \frac{6\left[\sum d^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) + \frac{1}{12}(m_3^3 - m_3)\right]}{n(n^2-1)}$$

$$= 1 - \frac{6\left[27.50 + \frac{1}{12}(3^3-3) + \frac{1}{12}(3^3-3) + \frac{1}{12}(2^3-2)\right]}{10(100-1)}$$

$$= 1 - \frac{6 \times 32}{990}$$

$$\rho_s = 0.806$$

There is high positive correlation.

Solved Problem 4.38 ***Anna Univ MBA: May/June 2006***

Calculate the coefficient of rank correlation from the following table.

X	:	48	34	40	12	16	66	25	16	57	
Y	:	15	15	24	8	13	6	20	9	9	15

Solution

X	R_1	Y	R_2	$d = R_1 - R_2$	d^2
48	8	15	7	1	1
34	6	15	7	−1	1
40	7	24	10	−3	1
12	1	8	2	−1	1
16	3	13	5	−2	4
16	3	6	1	2	4
66	10	20	9	1	1
25	5	9	3.5	1.5	2.25
16	3	9	3.5	−0.5	0.25
57	9	15	7	2	4
					$\sum d^2 = 27.5$

Here $n = 10, \sum d^2 = 27.5$

Coefficient of rank correlation

$$\rho_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$\rho_s = 1 - \frac{6 \times 27.5}{10(10^2-1)}$$

$$= 1 - \frac{165}{990}$$

$$= 1 - 0.1667$$

$$\rho_s = 0.8333$$

4.7 Test for Rank Correlation Coefficient

For small values of $n(\leq 30)$, the distribution of r_s (sample correlation coefficient) is not normal, and unlike other small sample statistic, we have encountered, it is not appropriate to use the t-distribution. Instead we use Spearman's rank correlation values which are given in a tabular form (Table) for $\alpha = 0.20,\ 0.10,\ 0.05,$ $0.02, 0.01$ and 0.002. For testing hypothesis about the rank correlation coefficient we determine the acceptance and rejection regions for such hypothesis.

For a given level of significance α and the number of pairs of observations n, we reject the null hypothesis H_0, if $|r_s| >$ tabulated value of r for given n and α, otherwise we accept H_0.

When n is greater than $30(n > 30)$, the sampling distribution of r_s is approximately normal, with a mean of zero and a standard deviation of $\frac{1}{\sqrt{n-1}}$. Thus, under the null hypothesis, the test statistic is

$$Z = \frac{r_s - 0}{\frac{1}{\sqrt{n-1}}} \sim N(0,1)$$

$$Z = r_s.(\sqrt{n-1})$$

If $|Z| \leq Z_\alpha$, we accept H_0 at a given level of significance α, otherwise we reject H_0.

Solved Problem 4.39

The following are the year of experience (X) and the average customer satisfaction (Y) for 10 service providers. Is there a significant rank correlation between measures? Use the 0.05 level of significance.

X	:	6.3	5.8	6.1	6.9	3.4	1.8	9.4	4.7	7.2	2.4
Y	:	5.3	8.6	4.7	4.2	4.9	6.1	5.1	6.3	6.8	5.2

Solution

Null hypothesis

$H_0 : \rho_s = 0$ i.e there is no significant rank correlation between the two measures.

Alternative hypothesis

$H_1 : \rho_s \neq 0$ i.e there is a significant rank correlation between the two measures.

Level of significance

$\alpha = 5\%$

Test statistic

Let us find the ranks for X and Y.

X	Y	R_1	R_2	$d = R_1 - R_2$	d^2
6.3	5.3	4	5	-1	1
5.8	8.6	6	1	5	25
6.1	4.7	5	9	-4	16
6.9	4.2	3	10	-7	49
3.4	4.9	8	8	0	0
1.8	6.1	10	4	6	36
9.4	5.1	1	7	-6	36
4.7	6.3	7	3	4	16
7.2	6.8	2	2	0	0
2.4	5.2	9	6	3	9
					$\sum d^2 = 188$

Therefore the sample rank correlation coefficient

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

$$= 1 - \frac{6 \times 188}{10(99)} = -0.139$$

The expected or critical value at 5% level of significance with $n = 10$ is 0.6364.

Conclusion

Since $|r_s| < 0.6364$, we accept H_0 and conclude that there is no significant rank correlation between the two measures.

Solved Problem 4.40

Test the hypothesis that X and Y are independent against the alternative that they are dependent if for a sample of size $n = 50$ pairs of observations we find that $r_s = -0.29$. Use $\alpha = 0.05$.

Solution

We are given

$$n = 50; r_s = -0.29$$

Null hypothesis

$H_0 : X$ and Y are independent.

Alternative hypothesis

$H_1 : X$ and Y are dependent.

Level of significance $\alpha = 0.05$

Test statistic (for $n > 30$)

$$Z = \frac{r_s - 0}{1/\sqrt{n-1}} \sim N(0,1)$$

$$= \frac{-0.29 - 0}{1/\sqrt{50-1}} = \frac{-0.29}{1/7} = -2.03$$

$$\therefore \quad |Z| = 2.03$$

The value of Z_α at $\alpha = 5\%$ is 1.96.

Conclusion

Since $|Z| > |Z_\alpha|$, we reject our null hypothesis H_0 and conclude that X and Y are dependent.

Solved Problem 4.41

A consumer panel tested 9 makes a microwave ovens for overall quality. The ranks assigned by the panel and the suggested retail prices were as follows