**A9Wb Testing the significance of Spearman's rank correlation coefficient, rs**

In this Section, we need to undertake a hypothesis test to test whether the true correlation relationship between the population Y, X values is significant, based upon the sample data values y, x. Having found a correlation using Spearman, it is necessary to test it to discover whether it is significant.

The hypothesis test takes the default form that there is no correlation. The alternative hypothesis is that there is a positive or negative correlation. Not only do significance levels affect the critical value but so do the number of values in the sample. The smaller the sample, the higher the correlation must be for it to be significant. In this hypothesis test we are assessing the possibility that ρs = 0.

If the sample size is n ≥ 6, then r is distributed as a t distribution with the number of degrees of freedom df = n – 2. It can be shown that the relationship between rs, ρs, and n, is given by equation (1):

(1)

**Example 1**

In this example 12 marketing plans for a new university student support centre have been assessed and ranked by the head of the business and digital technology departments.

|  |  |  |
| --- | --- | --- |
| Marketing plan | Business Dean rank, Br | Technology Dean rank, Tr |
| 1 | 5 | 7 |
| 2 | 1 | 1 |
| 3 | 4 | 5 |
| 4 | 7 | 4 |
| 5 | 6 | 6 |
| 6 | 8 | 10 |
| 7 | 9 | 8 |
| 8 | 12 | 11 |
| 9 | 2 | 3 |
| 10 | 3 | 2 |
| 11 | 10 | 12 |
| 12 | 11 | 9 |

Table 1

1. Calculate Spearman’s rank-order correlation coefficient, rs.
2. To determine whether the correlation between variables is significant.

**Step 1 - State hypothesis**

Null hypothesis H0: ρs = 0 no population correlation

Alternative hypothesis H1: ρs ≠ 0 population correlation exists

WHAT THE HYPOTHESES MEAN IN WORDS

Null Hypothesis H0: The population correlation coefficient IS NOT significantly different from zero. There IS NOT a significant linear relationship(correlation) between x and y in the population.

Alternate Hypothesis H1: The population correlation coefficient IS significantly DIFFERENT FROM zero. There IS A SIGNIFICANT LINEAR RELATIONSHIP (correlation) between x and y in the population.

**Step 2 - Select test**

We already know that this is testing the significance of Spearman’s rank correlation coefficient and because ρs ≠ 0, we will use a two-tail test.

**Step 3 - Set the level of significance** (α = 0.05)

**Step 4 - Extract relevant statistic**

Use Excel to compute the test statistics

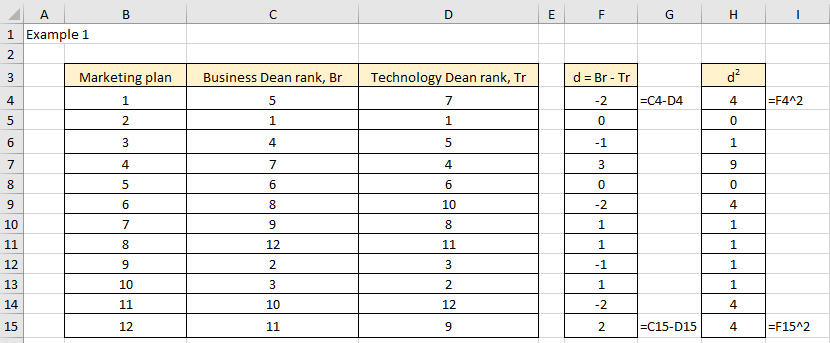


Figure 1

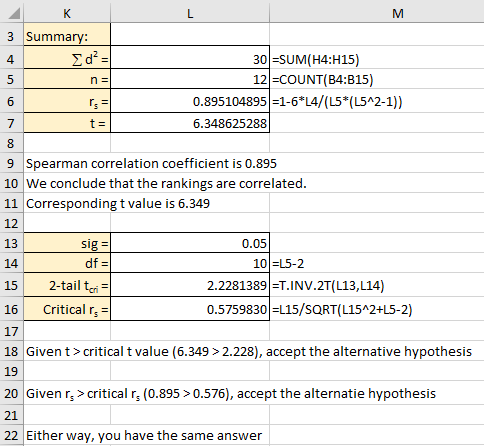


Figure 2

From Excel

If H0 true, then ρs = 0 and equation (1) simplifies to equation (2).

(2)

Calculate the value of rs.

rs = 0.895.

Calculate value of t using equation (2)

t = 6.349

We note that the alternative hypothesis is ≠ and therefore we have no implied direction for the value of ρs. All we know is that there could be a significant correlation and that ρs > 0 or ρs < 0. In this case we have two directions where ρ would be deemed significant and this is called a two-tailed test.

Critical values of rs

Small sample size (n ≤ 10)

The critical value of rs may be found either from a table of values or by calculation depending upon the size of the sample, n. The tables with critical values for rs can be found on the web. Table 1 represents critical rs values for n ≤ 10 (we only reproduced a small section of the table here):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N | 6 | 7 | 8 | 9 | 10 |
| Significance Level 5% | 0.829 | 0.759 | 0.738 | 0.666 | 0.632 |

Table 1 Critical values of Spearman’s correlation coefficient

Large sample size approximation (n > 10)

If the sample size exceeds 10, the test statistic is approximated by a t statistic with n - 2 degrees of freedom, as shown in equation (2).

The two-tail critical value of t

From Excel, t = t (0.05/2, df = n – 2) = t (0.025, 10) = ± 2.228.

The critical rs value can be found by re-arranging equation (2) to make rs the subject of the equation

 (3)

From Excel,

rs = 0.576

**Step 5 - Make decision**

Comparing t values

Given that 6.349 > 2.228, the test statistic falls in the critical region. Therefore, we reject the null hypothesis H0.

Comparing rs values

Given that 0.895 > 0.576, the test statistic falls in the critical region. Therefore, we reject the null hypothesis H0.

In both cases, we accept the alternative hypothesis.



At the level of significance of 5%, there is evidence to suggest a correlation between the ranks of the two variables does exist.

Normal approximation

Note that for n > 20, rs may be treated as normal (0, 1), where

 (4)

For example, if the significance level is 5% two tail and n = 40, then zcri = ± 1.96 and the critical value of rs = ± 0.314.