



INFERRENTIAL STATISTICS - SPEARMAN'S RANK CORRELATION COEFFICIENT

Correlation designs and using the Spearman's Rank statistical test

Psychodynamic Approach

Correlation designs

As part of the methodology for the Cognitive Approach, we learned of the three experimental designs: independent groups, repeated measures and matched pairs (see **M5 Experimental Design**). A further design, although not an experimental design, is the **correlation design**. This design of study involves comparing two different sets of data for the same set of participants. Each participant does two measures, and those measures are recorded and compared. This comparison is done by testing the relationship between the two sets of results statistically.

The relationships correlation designs might identify are **positive correlations** (i.e. as one measure goes up, the other measure also goes up) and **negative correlations** (i.e. as one measure goes up, the other measure goes down). When there is no relationship identified, the term 'no correlation' is used.

The data used for a correlation design must be numerical, and both the measures must come from the same participant. There is *no independent variable* and *no dependent variable*. There are just two variables, none of more significance than the other. The hypothesis of such a design will not be about a 'difference between' but will be hypothesising a relationship between the two measures.

Strengths of correlation designs	Weaknesses of correlation designs
<p>good for finding relationships at the start of an investigation; also unexpected relationships; once two sets of data have been collected from the same participants, a relationship can be statistically tested</p> <p>the data yielded is more secure, as there are no participant variables to affect it</p>	<p>the findings only show a relationship between those sets of data, not a definite connection, it does not allow room for the concepts of chance or another factor causing the relationship to arise</p> <p>if the data are artificial or unconnected, it is not valid</p>

Scattergraphs

Correlation data are normally displayed graphically using a **scattergraph**. The scores for each measure for each participant are used along the x-axis and the y-axis to plot a point for each participant, and then a line of best fit may be drawn to test for a correlation.

What is referred to as the 'eyeball test' is used to see if there is a correlation in a scattergraph. Simply looking at the results plotted on the chart and seeing the line of best fit should tell you if there is a correlation or not. A good measure is to compare the number of scores on or close to the line of best fit with the number of anomalous values. The more values there are on the scattergraph which don't seem to fit, the less likely there is to actually be a correlation.

Spearman's Rank correlation coefficient

The **Spearman's Rank** test (Spearman's Rank correlation coefficient) is a statistical test used to mathematically calculate if there is a relationship between two sets of data. We will use the example of IQ and income of some participants.

Participant	IQ	Income (£)
1	118	35,000
2	103	12,000
3	98	10,000
4	124	18,000
5	109	15,000
6	115	20,000
7	130	30,000
8	110	12,000

In our study, eight participants were given an IQ test and their scores were recorded. The participants then had their relative salaries noted. The following instructions outline how to use the Spearman's test.

Step 1: Rank the first variable, the lowest rank for the lowest score
In this example, this means ranking the lowest IQ with a 1 and the highest IQ with an 8. When two participants share the same score, take an average rank. For example, if three of them have an IQ of 100, occupying ranks 3, 4 and 5, allocate all of them the rank 4 (average)

Step 2: Rank the second variable, the same way as the first

Simply rank the incomes with the lowest income receiving a 1 and the highest receiving an 8

Participant	IQ	Income (£)	IQ rank (Step 1)	Income rank (Step 2)
1	118	35,000	6	8
2	103	12,000	2	2.5
3	98	10,000	1	1
4	124	18,000	7	5
5	109	15,000	3	4
6	115	20,000	5	6
7	130	30,000	8	7
8	110	12,000	4	2.5

Step 3: Calculate the difference between the ranks for each participant

Take the value of Rank 2 away from the value of Rank 1 (in this example take income rank from IQ rank). This will give the difference between the two ranks. If the resultant number is negative, don't forget to record the minus sign

Step 4: Square the differences between the ranks

Square the value you obtain for each participant from Step 3

Step 5: Total the figures from Step 4

Find the total by adding up all the values obtained from Step 4. The Greek letter sigma, Σ , is used to show add total

Participant	IQ	Income (£)	IQ rank (Step 1)	Income rank (Step 2)	IQ rank – income rank	(IQ rank – income rank) ²
1	118	35,000	6	8	-2	4
2	103	12,000	2	2.5	-0.5	0.25
3	98	10,000	1	1	0	0
4	124	18,000	7	5	2	4
5	109	15,000	3	4	-1	1
6	115	20,000	5	6	-1	1
7	130	30,000	8	7	1	1
8	110	12,000	4	2.5	1.5	2.25
					Total (Σ) =	13.5

Step 6: Multiply the value of Step 5 by 6

In our example, $13.5 \times 6 = 81$

Step 7: Find the value of N

N is the number of pairs of observations you have, so this will simply be the number of participants, in our case 8

Step 8: Square the value of N and subtract 1 from that number

In our example, $8 \times 8 = 64$ and $64 - 1 = 63$

Step 9: Multiply the number from Step 8 by N

In our example, $63 \times 8 = 504$

Step 10: Divide the value of Step 6 by the value of Step 9

In our example, $81 \div 504 = 0.160714285$

Step 11: Calculate rho, the result of the Spearman's test, by doing: $1 - \text{Step 10}$

In our example $1 - 0.160714285 = 0.839285714$

A statistical table can then be used to see if there is a correlation between the two sets of data. Statistical tables cannot be simply generated by thinking about them, they take hundreds of mathematical studies to calculate. They are published in masses in books so people can refer to them in order to see if there is a correlation in their data.

[youtube.com/c/MegaLecture/](https://www.youtube.com/c/MegaLecture/)

+92 336 7801123
aspsychology101.wordpress.com

The table below is an extract from a statistical table for rho:

		Level of significance for one-tailed test			
		0.05	0.025	0.01	0.005
N =	4	1.000			
	5	0.900	1.000	1.000	
	6	0.829	0.886	0.943	1.000
	7	0.714	0.786	0.893	0.929
	8	0.643	0.738	0.833	0.881
	9	0.600	0.700	0.783	0.833
	10	0.564	0.648	0.745	0.794
	11	0.536	0.618	0.709	0.755
	12	0.503	0.587	0.671	0.727

You will come across levels of significance in more detail in M13 Inferential Statistics – Mann-Whitney U Test, but for now just bear in mind that choosing a level of significance basically means choosing an acceptable level at which you can reject the null hypothesis and that the results are due to chance, and accept the alternative hypothesis.

Using our example study, we have a rho value of 0.893285714. We used 8 participants in the study, giving us an N value of 8 so that row has been highlighted in the table.

The critical value shown for the appropriate level of significance in the table has to be LESS THAN rho for the hypothesis to be proven. If the rho value is less than the critical value in the table, the null hypothesis is retained, as it cannot be rejected. This means your rho value must be MORE THAN the value in the table (called the critical value)

So back to our example. We can definitely accept the alternative hypothesis as proven for significance of 0.05 (given that $0.643 < 0.893$). This means we can say that the results obtained showing a correlation are less than or equal to five per cent due to chance. We can also say this for 0.025, 0.01 and 0.005 as the rho value is larger than all of the critical values for $N = 8$. Because we can do so for a significance level of 0.005, we can actually say that the relationship is less than or equal to half a per cent due to chance.