

How do you calculate Point-Biserial Correlation in Excel?

Authored by
stats writer

April 19, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you calculate Point-Biserial Correlation in Excel?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137214>

Point-Biserial Correlation is a statistical measure used to determine the relationship between a continuous variable and a binary variable. In Excel, this can be calculated by first organizing the data into two columns, one for the continuous variable and one for the binary variable. Then, using the CORREL function, the correlation coefficient can be calculated. This coefficient, ranging from -1 to 1, represents the strength and direction of the relationship between the two variables. A positive value indicates a positive correlation, while a negative value indicates a negative correlation. Point-Biserial Correlation in Excel is a quick and simple way to analyze the association between a continuous and binary variable.

Calculate Point-Biserial Correlation in Excel

Point-biserial correlation is used to measure the relationship between a binary variable, x , and a continuous variable, y .

Similar to the Pearson correlation coefficient, the point-biserial correlation coefficient takes on a value between -1 and 1 where:

**-1 indicates a perfectly negative correlation between two variables
0 indicates no correlation between two variables
1 indicates a perfectly positive correlation between two variables**

This tutorial explains how to calculate the point-biserial correlation between two variables in Excel.

Example: Point-Biserial Correlation in Excel

Suppose we have the following binary variable, x, and a continuous variable, y:

	A	B	C	D	E	F	G
1	x	y					
2	0	12					
3	1	14					
4	1	17					
5	0	17					
6	0	11					
7	0	22					
8	1	23					
9	0	11					
10	1	19					
11	1	8					
12	0	12					
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							

To calculate the point-biserial correlation between x and y, we can simply use the =CORREL() function as follows:

	A	B	C	D	E	F	G
1	x	y					
2	0	12		0.218163	=CORREL(A2:A12, B2:B12)		
3	1	14					
4	1	17					
5	0	17					
6	0	11					
7	0	22					
8	1	23					
9	0	11					
10	1	19					
11	1	8					
12	0	12					
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							

The point-biserial correlation between x and y is 0.218163.

Since this number is positive, this indicates that when the variable x takes on the value "1" that the variable y tends to take on higher values compared to when the variable x takes on the value "0."

We can easily verify this by calculating the average value of y when x is 0 and when x is 1:

	A	B	C	D	E	F	G	H
1	x	y						
2	0	12		0.218163				
3	1	14						
4	1	17		Avg. y when x = 0	14.2	=AVERAGEIF(A2:A12, 0, B2:B12)		
5	0	17		Avg. y when x = 1	16.2	=AVERAGEIF(A2:A12, 1, B2:B12)		
6	0	11						
7	0	22						
8	1	23						
9	0	11						
10	1	19						
11	1	8						
12	0	12						
13								
14								
15								
16								
17								
18								
19								
20								

When $x = 0$, the average value of y is 14.2. When $x = 1$, the average value of y is 16.2. This confirms the fact that the point-biserial correlation between the two variables should be positive.

We can also use the following formulas to calculate the p-value for this correlation coefficient:

	A	B	C	D	E	F	G	H
1	x	y						
2	0	12		0.218163				
3	1	14						
4	1	17		Avg. y when x = 0	14.2	=AVERAGEIF(A2:A12, 0, B2:B12)		
5	0	17		Avg. y when x = 1	16.2	=AVERAGEIF(A2:A12, 1, B2:B12)		
6	0	11						
7	0	22		sample size	11	=COUNT(B2:B12)		
8	1	23		t-test statistic	0.6706	=D2*SQRT(E7-2)/SQRT(1-D2^2)		
9	0	11		p-value	0.5193	=T.DIST.2T(E8, E7-2)		
10	1	19						
11	1	8						
12	0	12						
13								
14								
15								
16								
17								
18								

The p-value turns out to be 0.5193. Thus, although the correlation coefficient between the two variables is slightly positive it turns out to not be a statistically significant correlation.

How to Calculate Spearman Rank Correlation in Excel

How to Calculate Partial Correlation in Excel

How to Find the P-value for a Correlation Coefficient in Excel