How can you calculate a Pearson Correlation Coefficient by hand?

Authored by

stats writer

April 22, 2024

The Pearson Correlation Coefficient is a statistical measure that determines the strength and direction of the linear relationship between two variables. It is often used to analyze and understand the relationship between two sets of data. To calculate the Pearson Correlation Coefficient by hand, one must follow a specific formula that involves finding the mean, standard deviation, and covariance of the two variables. These values are then used to calculate the correlation coefficient, which ranges from -1 to +1. A positive value indicates a positive correlation, while a negative value indicates a negative correlation. By calculating the Pearson Correlation Coefficient by hand, one can gain a deeper understanding of the relationship between two variables and make informed decisions based on the resulting value.

Calculate a Pearson Correlation Coefficient by Hand

A Pearson Correlation Coefficient measures the linear association between two variables.

It always takes on a value between -1 and 1 where:

-1 indicates a perfectly negative linear correlation between two variables0 indicates no linear correlation between two variables1 indicates a perfectly positive linear correlation between two variables

The formula to calculate a Pearson Correlation Coefficient, denoted *r*, is:

Source: Wikipedia

This tutorial provides a step-by-step example of how to calculate a Pearson Correlation Coefficient by hand for the following dataset:

x	y
6	45
12	47
13	39
17	58
22	68
25	76
27	75
29	74
30	78
32	81

Step 1: Calculate the Mean of X and Y

First, we'll calculate the mean of both the X and Y values:

	x	y
	6	45
	12	47
	13	39
	17	58
	22	68
	25	76
	27	75
	29	74
	30	78
	32	81
Mean	**21.3**	**64.1**

Step 2: Calculate the Difference Between Means

Next, we'll calculate the difference between each of the individual X and Y values and their respective means:

	x	y	$x - x_{mean}$	$y - y_{mean}$
	6	45	-15.3	-19.1
	12	47	-9.3	-17.1
	13	39	-8.3	-25.1
	17	58	-4.3	-6.1
	22	68	0.7	3.9
	25	76	3.7	11.9
	27	75	5.7	10.9
	29	74	7.7	9.9
	30	78	8.7	13.9
	32	81	10.7	16.9
Mean	**21.3**	**64.1**		

Step 3: Calculate the Remaining Values

Next, we'll calculate the remaining values needed to complete the Pearson Correlation Coefficient formula:

	x	y	$x - x_{mean}$	$y - y_{mean}$	$(x - x_{mean})*(y - y_{mean})$	$(x - x_{mean})^2$	$(y - y_{mean})^2$
	6	45	-15.3	-19.1	292.23	234.09	364.81
	12	47	-9.3	-17.1	159.03	86.49	292.41
	13	39	-8.3	-25.1	208.33	68.89	630.01
	17	58	-4.3	-6.1	26.23	18.49	37.21
	22	68	0.7	3.9	2.73	0.49	15.21
	25	76	3.7	11.9	44.03	13.69	141.61
	27	75	5.7	10.9	62.13	32.49	118.81
	29	74	7.7	9.9	76.23	59.29	98.01
	30	78	8.7	13.9	120.93	75.69	193.21
	32	81	10.7	16.9	180.83	114.49	285.61
Mean	**21.3**	**64.1**					

Step 4: Calculate the Sums

	x	y	$x - x_{mean}$	$y - y_{mean}$	$(x - x_{mean})*(y - y_{mean})$	$(x - x_{mean})^2$	$(y - y_{mean})^2$
	6	45	-15.3	-19.1	292.23	234.09	364.81
	12	47	-9.3	-17.1	159.03	86.49	292.41
	13	39	-8.3	-25.1	208.33	68.89	630.01
	17	58	-4.3	-6.1	26.23	18.49	37.21
	22	68	0.7	3.9	2.73	0.49	15.21
	25	76	3.7	11.9	44.03	13.69	141.61
	27	75	5.7	10.9	62.13	32.49	118.81
	29	74	7.7	9.9	76.23	59.29	98.01
	30	78	8.7	13.9	120.93	75.69	193.21
	32	81	10.7	16.9	180.83	114.49	285.61
Mean	**21.3**	**64.1**		**Sum**	1172.7	704.1	2176.9

Step 5: Calculate the Pearson Correlation Coefficient

Now we'll simply plug in the sums from the previous step into the formula for the Pearson Correlation Coefficient:

	x	y	$x - x_{mean}$	$y - y_{mean}$	$(x - x_{mean})*(y - y_{mean})$	$(x - x_{mean})^2$	$(y - y_{mean})^2$
	6	45	-15.3	-19.1	292.23	234.09	364.81
	12	47	-9.3	-17.1	159.03	86.49	292.41
	13	39	-8.3	-25.1	208.33	68.89	630.01
	17	58	-4.3	-6.1	26.23	18.49	37.21
	22	68	0.7	3.9	2.73	0.49	15.21
	25	76	3.7	11.9	44.03	13.69	141.61
	27	75	5.7	10.9	62.13	32.49	118.81
	29	74	7.7	9.9	76.23	59.29	98.01
	30	78	8.7	13.9	120.93	75.69	193.21
	32	81	10.7	16.9	180.83	114.49	285.61
Mean	**21.3**	**64.1**		**Sum**	1172.7	704.1	2176.9

$r = 1172.7 / \sqrt{(704.1)*(2176.9)} = \mathbf{0.947}$

The Pearson Correlation Coefficient turns out to be 0.947.

Since this value is close to 1, this is an indication that X and Y are strongly positively correlated.

In other words, as the value for X increases the value for Y also increases in a highly predictable fashion.

An Introduction to the Pearson Correlation Coefficient

How to Find a Confidence Interval for a Correlation Coefficient