

Pearson Correlation Coefficient

The Pearson correlation coefficient (PCC, pronounced /'piərsən/), also referred to as Pearson's r , the Pearson product-moment correlation coefficient (PPMCC), or the bivariate correlation, is a measure of the linear correlation between two variables X and Y whose coordinates are x and y . The coefficient r is defined variously but one definition is as follows:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Granted, this is a pretty messy relationship. Fortunately, calculators and computer programs (e.g. Excel, online apps, etc.) will perform these calculations for you.

For a situation where the relationship is expected to be linear...



When the Pearson correlation coefficient r has a value of:

- +1, it is a perfect positive correlation.
- 0, it is no correlation (the values aren't linked at all).
- -1, it is a perfect negative correlation.

The r value shows how good the correlation is (not how steep the slope is), and if it is positive or negative. The correlation coefficient is:

- positive when the values increase together, and
- negative when one value decreases as the other increases

Keep in mind that correlation does not always imply cause and effect. While it usually does in physics, it's not always so in other real-world situations. For instance, geese tend to fly in a northerly direction in the spring and in a southerly direction in the autumn. The flight of geese, however, does not produce the seasons.