

Pearson Correlation Assumptions

<http://www.statisticssolutions.com/pearson-correlation-assumptions/>

The assumptions of the Pearson product moment correlation can be easily overlooked. The assumptions are as follows: level of measurement, related pairs, absence of outliers, normality of variables, linearity, and [homoscedasticity](#).

Level of measurement refers to each variable. For a [Pearson correlation](#), each variable should be continuous. If one or both of the variables are ordinal in measurement, then a [Spearman correlation](#) could be conducted instead.

Related pairs refers to the pairs of variables. Each participant or observation should have a pair of values. So if the correlation was between weight and height, then each observation used should have both a weight and a height value.

Absence of outliers refers to not having outliers in either variable. Having an outlier can skew the results of the correlation by pulling the line of best fit formed by the correlation too far in one direction or another. Typically, an outlier is defined as a value that is 3.29 standard deviations from the mean, or a standardized value of less than ± 3.29 .

Linearity and homoscedasticity refer to the shape of the values formed by the [scatterplot](#). For linearity, a “straight line” relationship between the variable should be formed. If a line were to be drawn between all the dots going from left to right, the line should be straight and not curved. Homoscedasticity refers to the distance between the points to that straight line. The shape of the scatterplot should be tube-like in shape. If the shape is cone-like, then homoskedasticity would not be met.