Glossary

1. Bivariate

Bivariate data is data that has two variables. The quantities from these two variables are often represented using a scatter plot. This is done so that the relationship (if any) between the variables is easily seen

2. Coefficient of determination

Coefficient of determination R^2 is used in the context of statistical models whose main purpose is the prediction of future outcomes on the basis of other related information. It is the proportion of variability in a data set that is accounted for by the statistical model. It provides a measure of how well future outcomes are likely to be predicted by the model.

3. Constant

Constant is something that is unchanging or invariable, or that is continually occurring.

4. Correlation

Correlation is a statistical technique that can show whether and how strongly pairs of variables are related.

5. Correlation coefficient

Correlation coefficient is a measure of the interdependence of two random variables that ranges in value from -1 to +1, indicating perfect negative correlation at -1, absence of correlation at zero, and perfect positive correlation at +1. This is also called as coefficient of correlation.

6. Covariance

Covariance is a measure of how much two random variables change together. If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the smaller values, i.e., the variables tend to show similar behavior, the covariance is a positive number. In the opposite case, when the greater values of one variable mainly correspond to the smaller values of the other, i.e., the variables tend to show opposite behavior, the covariance is negative.

7. Density

Density is the quantity of something per unit measure, especially per unit length, area, or volume or the number of individuals, such as inhabitants or housing units, per unit of area.

8. Probable error

The probable error of a quantity is a value describing the probability distribution of that quantity. It defines the half-range of an interval about a central point for the distribution, such that half of the values from the distribution will lie within the interval and half outside.

9. Product moment correlation coefficient

The Pearson product-moment correlation coefficient (sometimes referred to as the PPMCC or PCC or Pearson's r, and is typically denoted by r) is a measure of the correlation (linear dependence) between two variables X and Y, giving a value between +1 and -1 inclusive. It is widely used in the sciences as a measure of the strength of linear dependence between two variables. It was developed by Karl Pearson from a similar but slightly different idea introduced by Francis Galton in the 1880s.

10. Raw data

Raw data is a term for data collected from a source. Raw data have not been subjected to processing or any other manipulation, and are also referred to as primary data.

11. Standard deviation

Standard deviation (represented by the symbol sigma, σ) shows how much variation or dispersion exists from the average (mean, or expected value). A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data points are spread out over a large range of values.

12. Substitution

The replacement of a term of an equation by another that is known to have the same value in order to simplify the equation. Substitution of variables (also called variable substitution or coordinate transformation) refers to the substitution of certain variables with other variables.

13. Summation

Summation is the operation of adding a sequence of numbers; the result is their sum or total.

14. Tabulated data

Tabulated data is a term for placing classified data into tabular form. A table is a symmetric arrangement of statistical data in rows and columns. Rows are horizontal arrangements whereas columns are vertical arrangements.

15. Variance

Variance is a measure of how far a set of numbers is spread out. It is one of several descriptors of a probability distribution, describing how far the numbers lie from the mean (expected value).