Glossary

1. Algorithm

An algorithm is an effective method expressed as a finite list of well-defined instructions for calculating a function. Starting from an initial state and initial input the instructions describe a computation that, when executed, will proceed through a finite number of well-defined successive states, eventually producing an output.

2. Array

An array is a systematic arrangement of objects.

3. Correlation ratio

Correlation ratio is a measure of the relationship between the statistical dispersion within individual categories and the dispersion across the whole population or sample. The measure is defined as the *ratio* of two standard deviations representing these types of variation.

4. Covariance

Covariance is a measure of how much two random variables change together.

5. Intra-class correlation

Intra-class correlation is a descriptive statistic that can be used when quantitative measurements are made on units that are organized into groups. It describes how strongly units in the same group resemble each other.

6. Pearson's correlation

In statistics, the Pearson product-moment correlation coefficient or Pearson's r, (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.

7. Probability distribution

Probability distribution is a function that gives the probability of all elements in a given space.

8. Random variable

In probability and statistics, a random variable or stochastic variable is a variable whose value is subject to variations due to chance (i.e. randomness, in a mathematical sense). As opposed to other mathematical variables, a random variable conceptually does not have a single, fixed value rather; it can take on a set of possible different values, each with an associated probability

9. **Range of correlation**

The correlation ratio η takes values between 0 and 1. The limit $\eta = 0$ represents the special case of no dispersion among the means of the different categories, while $\eta = 1$ refers to no dispersion within the respective categories.

10. **Regression**

Regression analysis includes many techniques for modelling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables. More specifically, regression analysis helps one understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed.

11. **Skewness**

Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable. Skewness value can be positive or negative, or even undefined.

12. **Standard deviation**

This is the most commonly used measure of statistical dispersion. It is the square root of the variance, and is generally written as σ (sigma).

13. **Statistical dispersion**

Statistical dispersion (also called statistical variability or variation) is variability or spread in a variable or a probability distribution. Common examples of measures of statistical dispersion are the variance, standard deviation and interquartile range.

14. **Statistical parameter**

Statistical parameter is a parameter that indexes a family of probability distributions.

15. **Vector**

A vector is a one dimensional array. A multivariate random variable or random vector is a list of mathematical variables each of whose values is unknown, either because the value has not yet occurred or because there is imperfect knowledge of its value.