

The Statistical Distribution

Statistical distribution refers to the representation of the possible values that a variable can take and the relative probabilities associated with each of these values. It is a fundamental concept in statistics and data analysis, as it provides crucial information about the behavior and nature of random variables.

Statistical distributions can be of different types, but the main categories are **discrete distributions** and **continuous distributions**.

Discrete distributions

Discrete distributions refer to random variables that take on discrete or distinct values. These values are usually integers. A common example is the probability distribution associated with rolling a die, in which the possible outcomes are integers from 1 to 6. Other discrete distributions include the binomial, Poisson, and geometric distributions.

Discrete distributions are characterized by various properties that define their behavior and the statistical representation of the data.

Discrete Values:

Discrete distributions are associated with random variables that can take on specific, separate values. These values are often non-negative integers.

Probability function:

For discrete distributions, we use the discrete probability function that assigns to each possible value of the random variable the probability of obtaining that value.

Sum of probabilities:

The sum of all the probabilities associated with the individual values of the random variable is equal to 1.

Mean:

The mean of a discrete distribution is calculated as the sum of the products between the values of the variable and their corresponding probabilities.

Variance:

Variance is a measure of the dispersion of data around the mean. For discrete distributions, the variance is calculated as the sum of the squares of the differences between the variable's values and the mean, each multiplied by the probability of that value.

Mode:

The mode of a discrete distribution represents the value that occurs with the greatest frequency or probability.

Cumulative distribution function:

Discrete distributions can be described through the cumulative distribution function, which indicates the probability that the random variable is less than or equal to some specific value.

Continuous distributions

They concern random variables that can take on an infinite number of values within a continuous interval. For example, elevations, temperatures, or time taken to complete a task are continuous variables. The normal (or Gaussian) distribution is an example of a continuous distribution and is widely used in many disciplines due to its statistical properties.

Continuous distributions are characterized by the following properties:

Continuous values:

Continuous distributions are associated with random variables that can take on an infinite number of values within a continuous interval. There are no jumps or discontinuities in the possible values of the random variable.

Probability density function:

For continuous distributions, the probability density function (**PDF**) is used, which represents the probability density for different values of the variable within the interval. The probability of getting an exact value is zero, but the area under the PDF curve within a specific range represents the probability of getting values in that range.

Integral of the PDF:

The integral of the probability density function over the range of possible values of the random variable is equal to 1.

Mean:

The mean of a continuous distribution is calculated as the integral of the product between the values of the variable and their corresponding probability density.

Variance:

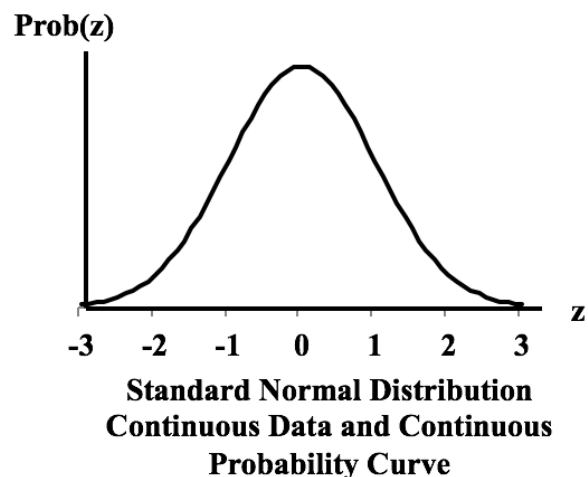
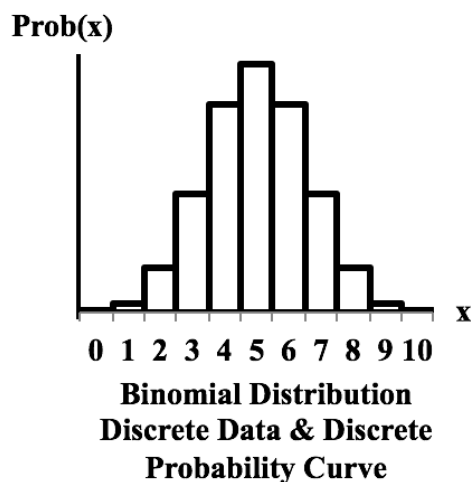
The variance of a continuous distribution is a measure of the dispersion of the data around the mean. It is calculated through the integral of the squares of the differences between the values of the variable and the mean, weighted by their probability density.

Median:

The median of a continuous distribution is the central value that divides the distribution into two equal parts in terms of the area under the curve of the PDF.

Statistical distributions are used to model and understand data behavior, as well as make predictions and inferences in a wide range of disciplines.

Understanding the type of distribution associated with a given data set is critical to correctly analyzing and interpreting statistical results.



References:

Probability Distribution - *Wikipedia*

https://en.wikipedia.org/wiki/Probability_distribution

Discrete Distribution vs Continuous Distribution - *Statistics from A to Z*

<https://www.statisticsfromatoz.com/blog/statistics-tip-of-the-week-different-distributions-can-have-discrete-or-continuous-probability-graphs-for-discrete-or-continuous-data>