Types of distributions

Types of Data Distribution in Statistics

There are two main types of data distribution in statistics: continuous and discrete.

Continuous data:

Continuous data is a type of information that can range from one extreme to another, usually measured on a scale such as temperature or weight. It can also be presented in the form of a histogram which allows for easier comparison and understanding between different sets of data. With Continuous Data, you are able to gain insights into trends and relationships that might not ordinarily be seen with other types of datasets.

Discrete data:

Discrete data has a limited set of values and ranges, such as countable elements like the student population in a classroom or cars passing through an intersection. Representing this kind of information with bar graphs allows for quick understanding at-a-glance!

Prior to utilising statistical tools, it is essential that you comprehend the kind of data at your disposal as the tools accessible are contingent on your type of data. Knowing this detail first will enable you to make an informed decision and utilise the most appropriate tool for successful results

Discrete Distribution Types:

- Binomial distribution
- Poisson distribution
- Hypergeometric distribution
- Geometric distribution

Binomial Distribution

The binomial distribution describes the probability of a certain number of successes (or failures) in a given number of trials or events. This type of distribution is used when there are only two possible outcomes for each trial, such as success or failure, heads or tails, yes or no etc., with equal probabilities for each product. The binomial distribution can be used to calculate the likelihood of achieving a specific result from multiple independent trials that have only two possible effects.

Poisson Distribution

The Poisson distribution describes the probability that an event will occur within a fixed time period when its rate is known but its exact timing cannot be predicted accurately enough to measure it directly. This type of distribution is useful for modelling random occurrences such as customer arrivals at stores, phone calls received by call centers etc., where the average rate of occurrence is known but the exact timing cannot be measured.

Hypergeometric Distribution

The hypergeometric distribution describes the probability of a certain number of successes (or failures) in a given number of draws from an urn or population when the draws are made without replacement. This type of data distribution is used in situations where an urn contains different items, such as colored balls, and you want to evaluate the probability that a certain number of those items will be drawn out with each draw being made without replacing the item that was just taken out.

Geometric Distribution

The geometric distribution describes the probability of a success occurring on any given trial in a series of independent trials when the probability of success for each trial is known. This type of data distribution can be used to model the number of failures that occur before a success in situations such as manufacturing processes, where there are multiple attempts at creating a product and each attempt has a given probability of success.

Continuous Distribution Types

- Normal Distribution
- Lognormal distribution
- F distribution
- Chi-Square distribution
- Exponential distribution
- T Student distribution
- Weibull Distribution
- Non-normal distributions

Data distributions are a way of describing data sets by plotting individual data points on a graph. This graphical representation can help us to understand the data and make predictions about it. There are different types of data distributions, each with its characteristics and uses. In this article, we will discuss eight types of data distributions: Continuous Distribution, Normal

Distribution, Lognormal distribution, F distribution, Chi-Square distribution, Exponential distribution, T Student distribution and Weibull Distribution.

Continuous data distributions measure data points over a range instead of as individual data points. Continuous data distributions are typically used for data that is normally distributed (data that does not have any outliers). Examples of continuous data distributions include Normal Distribution, Lognormal distribution and F distribution.

Normal Distribution

Normal Distributions are one of the most commonly used data distributions. This distribution measures data points in a bell-shaped curve, with an equal number of data points to the left and right of the mean value. Normal Distributions can be used to predict future outcomes based on past trends.

Lognormal Distribution

Lognormal distributions measure data points in a curve shaped like a sigmoid function – a curved line beginning at zero and then increasing sharply to a peak and slowly decreasing. This data distribution is often used in financial data, allowing us to extrapolate potential future stock prices based on past data.

F Distribution

F Distributions measure data points spread out over a broader range than normal distributions. This data distribution is often used when measuring data with higher variability, such as performance data or customer satisfaction surveys.

Chi-Square Distribution

Chi-Square Distributions measure the difference between observed data and expected results. This data distribution can be used to identify significant differences between two data sets and help us understand which factors may be influencing our results.

Exponential Distribution

Exponential distributions measure data points with an exponential curve – a curve beginning at zero and gradually increasing in value. This data distribution is often used when data points are expected to increase over time, such as population data or customer data in a given market.

T-Student Distribution

T Student Distributions measure data points spread out more than normal distributions. This data distribution can be used for data sets with higher variability and outliers, such as performance data.

Weibull Distribution

Weibull Distributions measure data in an exponential curve – a curve beginning at zero and gradually increasing in value. This data distribution is often used for reliability tests and can help us predict how long it will take for a system to fail.

Non-normal Distribution

Non-normal distributions include data distributions such as the Poisson Distribution, Gamma Distribution, Beta Distribution, Logistic Distribution and Cauchy Distribution. Non-normal data distributions are often used when data does not fit into the normal data distribution categories, such as highly non-linear or data with outliers.