Rocco Vulpis

December 14, 2023

Normal, Gamma, and Beta Probability Distributions

In the field of statistics there are various distributions that are used to evaluate different outcomes. A common mistake made when studying statistics is the purpose of distributions, and how to properly use them. A distribution is defined as a mathematical equation that is used to show the spread of a dataset and how often values occur. According to IBM, “A statistical distribution, or probability distribution, describes how values are distributed for a field” (IBM.com, 1). Specifically, throughout statistics there are three commonly used distributions known as the “Normal Probability Distribution”, “Gamma Probability Distribution”, and “Beta Probability Distribution” that each have their own unique way to calculate probability.

The most widely known/used continuous probability distribution is the normal distribution. This distribution is also called the Gaussian distribution. Its purpose is to describe the distribution of values for many natural occurrences such as heights, weights, births, reading ability, SAT scores, etc. It’s symmetric about the mean showing that the data around the mean occurs more frequently than the data further from the mean. When graphed, the distribution appears as a bell curve with most of the values being clustered around the central region. The data is distributed symmetrically with no skew. When calculating the distribution, there are two parameters necessary which are the mean and the standard deviation. The mean is the location parameter. The peak of the curve will be centered based on the mean. If the mean increases, the curve will move right. If the mean decreases, the curve moves left. The standard deviation is the scale parameter. A lower standard deviation results in a curve appearing narrow. When it’s higher, the curve appears wider.

A normal curve can be fit to a dataset once the mean and standard deviation have been calculated. This is done by using a probability density function. The area under the curve shows the probability in a probability density function. The total area under the curve is always 1 (100%) in a probability distribution. This is because the sum of all outcomes is always one, meaning that the occurrence is always certain. Areas under the normal density function require the evaluation of the integral

An additional widely used distribution is the gamma distribution that is used to describe random variables that are always nonnegative. It can be used to represent time modeling and reliability modeling for events such as predicted rainfall or the reliability of machines. Since the applications are often unbalanced, the distribution takes on its skewed shape. The distribution consists of three parameters: The shape parameter, location parameter, and the scale parameter. The shape parameter is represented by α (Alpha). It allows the distribution to take on a variety of shapes. Whereas β (Beta) is known as the scale parameter and it affects the overall scale of the distribution. Similarly, to normal distribution, the gamma distribution is described using a probability density function (PDF), but this is not exclusive. Generally, the gamma distribution is presented by the scale and shape parameters that together determine its characteristics. The variable Γ ​(α) is used to represent the gamma distribution. Areas under the gamma density function require the evaluation of the integral

The final distribution being discussed is known as the beta probability distribution. It is a two-parameter function that is used to model proportions such as the proportion of students who receive a certain grade and the proportion of product defect rate. The graphs assume widely differing shapes for various values of its two parameters which are α and β. For this distribution, the parameters are both shape parameters, and both values must be positive. Beta distribution can be related to the binomial probability distribution because the parameters are positive. A random variable Y is said to have a beta probability distribution with parameters α > 0 and β > 0 if and only if the density function of Y is

where

The cumulative distribution function for the beta random variable is commonly called the incomplete beta function and is denoted by

In conclusion, the Normal, Gamma, and Beta Probability Distributions each offer unique approaches for modeling different types of data and outcomes in the field of statistics. These distributions, each with their distinct applications, are essential tools in the field of statistics, enabling a deeper understanding of data behavior and assisting in making informed predictions and decisions.

Sources

1.3.6.3. families of distributions. (n.d.). https://www.itl.nist.gov/div898/handbook/eda/section3/eda363.htm

Bhandari, P. (2023, June 21). *Normal distribution: Examples, formulas, & uses*. Scribbr. https://www.scribbr.com/statistics/normal-distribution/

Chen, J. (n.d.). *Normal distribution: What it is, properties, uses, and Formula*. Investopedia. https://www.investopedia.com/terms/n/normaldistribution.asp

Frost, J. (2023, October 6). *Normal distribution in statistics*. Statistics By Jim. https://statisticsbyjim.com/basics/normal-distribution/

Scheaffer, R. L., Mendenhall, W., & Wackerly, D. D. (n.d.). *Mathematical Statistics with Applications, Seventh Edition*.

*Statistical distribution*. IBM. (n.d.). https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=terms-statistical-distribution

*What is the Gamma Distribution?*. Built In. (n.d.). https://builtin.com/data-science/gamma-distribution