Kyle Antczak

Normal, Gamma, Beta Distributions

Normal distribution is the most common continuous probability distribution. The shape of a normal distribution follows a bell shape curve, which is what is used for the empirical rule. The empirical rule states that almost all of the data will fall within three standard deviations() of the mean(). Specifically, it is that 68% of the values fall within the first standard deviation, 95% within the first two standard deviations and 99.7% within the first three. Because the normal distribution is bell shaped, that means if you are looking at just above or below the mean within one of those standard deviations its half those percent’s. For example, 34% of the values less than the mean fall within one standard deviation. The normal density function is:

Where

The expected for a normally distributed random variable is just the mean and the variance is the standard deviation squared. This like the book says implies that the mean locates the center of the distribution, and the standard deviation measures the spread. The area under the normal density function is the integral of y of the normal density function from a to b, but because there is no closed-form expression for this integral the most common approach to getting areas is by using a z table. The z table calculates the area from the mean to point z, where z is the distance from the mean measured in standard deviations. The normal distribution is symmetrical, meaning the area from z standard deviations below the mean to the mean is the same as z. For example P(Z > 2) is equal to P(-2 > Z), which is both .0228. The easiest way to find any value for Z is to find the area for it in comparison to the right side of the mean and then if necessary to either add it or subtract from one to find the area of the corresponding distance on the left side.

Lesson 4.6 covers the gamma probability distribution. Unlike a normal distribution which is symmetrical, the gamma distribution is skewed and nonsymmetric. An example of this being the length of time to complete a maintenance checkup for an automobile. The two parameters for a gamma distribution are α > 0 and β > 0. Its density function is:

α is sometimes called the shape parameter of the gamma distribution because of the different shapes produced with different α and the same β, while β is called the scale parameter because multiplying a gamma distributed variable by a positive constant produces a gamma distribution with the same α but different β. Unless α is an integer, there is no closed form expression for the integration of the area. When α is an integer, the gamma distribution random variable can be expressed as a sum of certain Poisson probabilities. The easiest way to compute probabilities is using statistical software. There are some tables where certain values are integrated. The expected mean for gamma distribution is αβ and the variance is α.

There are two special cases for gamma distributed random variables. The first is where a random variable Y is said to have a chi-square distribution with v degrees of freedom if and only if Y is a gamma-distributed random variable with parameters α = v/2 and β = 2. A variable with a chi-square distribution is called a chi-square random variable and these variables occur often in statistical theory, and tables of probabilities are common. The expected mean for chi-squared random variables is v and the variance is 2v. The next case is where α = 1, which is called the exponential density function. This has its own density function with the only parameter being β > 0. This density function is useful in modeling the length of life for electronics. The expected mean for an exponential distribution is β with the variance being

The last distribution is the beta probability distribution. This density function is defined over the closed interval 0 ≤ y ≤ 1. Its distribution is often used to model proportions like proportion of time a machine is under repair. The beta density function is:

Where the parameters are the same as the gamma density function, α > 0 and β > 0

The density functions have widely different shapes for different α and β. The beta density function can be applied to random variable defined on the interval c ≤ y ≤ d because (y – c) / (d – c) = y’ makes the variable fit on the scale 0 ≤ y’ ≤ 1 by translation and a change of scale. The cumulative distribution function can be written as , where n = . Like the other distributions statistical software is the easiest way to compute the binomial probabilities and there are also certain tables of the beta function. The beta density function can be integrated directly when are both integers. The expected for beta distributed random variables is and the variance is .