Darwyn Abreu

Beta, Gamma, and Normal Distributions

When looking at beta, gamma, and normal distribution one thing that is important to note is that Normal distribution is a distribution with a bell curve and is said to have a normal probability distribution if the density function is as follows provided that with being Variance/V and = (E)xpectation/Mean/E. In the book Theorem 4.7 states that E is where the center of the distribution lies while V is the measure of spread. When addressing the area of the region of the normal density function the integral of f(y) should be evaluated from the start and end points of the area given. When solving problems an important thing to remember is that the values are symmetric around E so only one side of the mean needs to be tabulated. An example problem is given in the book where a table of probabilities are given and you are given a probability of p(Z>2) provided E=0 and = 1 in order to solve the problem the distance between E and E+Z must be 2 or greater in order to fulfill the probability which means that Z must equal two. After looking at the chart and seeing that the first Z digit is 2 followed by .00 the table tells you the answer is .0228 which also represents the area under the curve. Given that information that we know the range for Z=2 is .0228 then given P() the Probability can be figured out to be 1-2(.0228) with 2(.0228) being the two ranges where z=2. Similarly the problem can be inversely solved if we have a range 0<y<1.73 given that the positive range accounts for half the data you have .5-(z=1.73). Z can also be expressed as units of standard deviation through the equation which can then be used to read the table and plugged into problems again. Overall it seems like the most important concept of normal distribution is that it is symmetrical so there is some data inherently given without being written out.

When looking at gamma probability distribution an important point to keep in mind is that it is going to be used for non-negative skewed/non symmetric data. The book describes the graph and most of the area being towards the origin and dropping off as y increases. Some examples of this type of data might be the amount of time between checking people out at a store or amount of hits landed in every round of a mma match. The function for gamma distribution is as follows : where . The function is known as the gamma function and as well as for any value of and . In general different values of will change the shape of the function so is referred to as the shape parameter while is referred to as the scale parameter. When the shape parameter is an integer the function can be expressed as a sum of Poisson probabilities. When the shape parameter is not an integer but is 0<c<d<infinity the result can be expressed as the integral of f(y) from c to d except when the shape parameter is equal to 1. When the shape parameter is equal to 1 the integral is impossible to take so therefore in order to find the value tabulated areas can be used. The mean of Gamma distribution is E(y)= and the variance is V(y)=. The book states the easiest way to deal with Gamma Function is by using pre-existing software and gives some use cases such as finding a variable with a chi-square value or when looking at the exponential density function which is when =1.

The last distribution to cover is Beta distribution. The beta density function is another two parameter density function however this one is only defined over the interval 0<=y<=1. The purpose of the function is usually for modeling proportions like the proportion of chemical impurities in a product The function for a beta distribution defined over the closed y range of (0,1) is as follows: where B(. The shape of the beta function’s graph will change significantly when either or are changed in any way. It is also important to not that if y is not on the interval (0,1) and is on interval (c,d) a y\* can be formed in order to fit into the interval of (0,1), this y\* is given as shown . There is also a cumulative distribution function for the beta variable which is called the incomplete beta function and is denoted as such F(y)=. When both variables are positive integers then the incomplete beta function is related to the binomial probability function. When both variables are integers and 0<y<1 F(y)= where n=. The mean for beta distribution is E(y)= and the variance is V(y)=

Overall it seems like the key takeaways is that when there is a bell curve of data available then normal distribution can be applied, when speaking about non-symmetric rates gamma distribution can be applied, and when speaking about proportions beta distribution can be applied. Each distribution has its own unique shape to its graph and its own unique ways in which it will or will not shift when variables are changed.