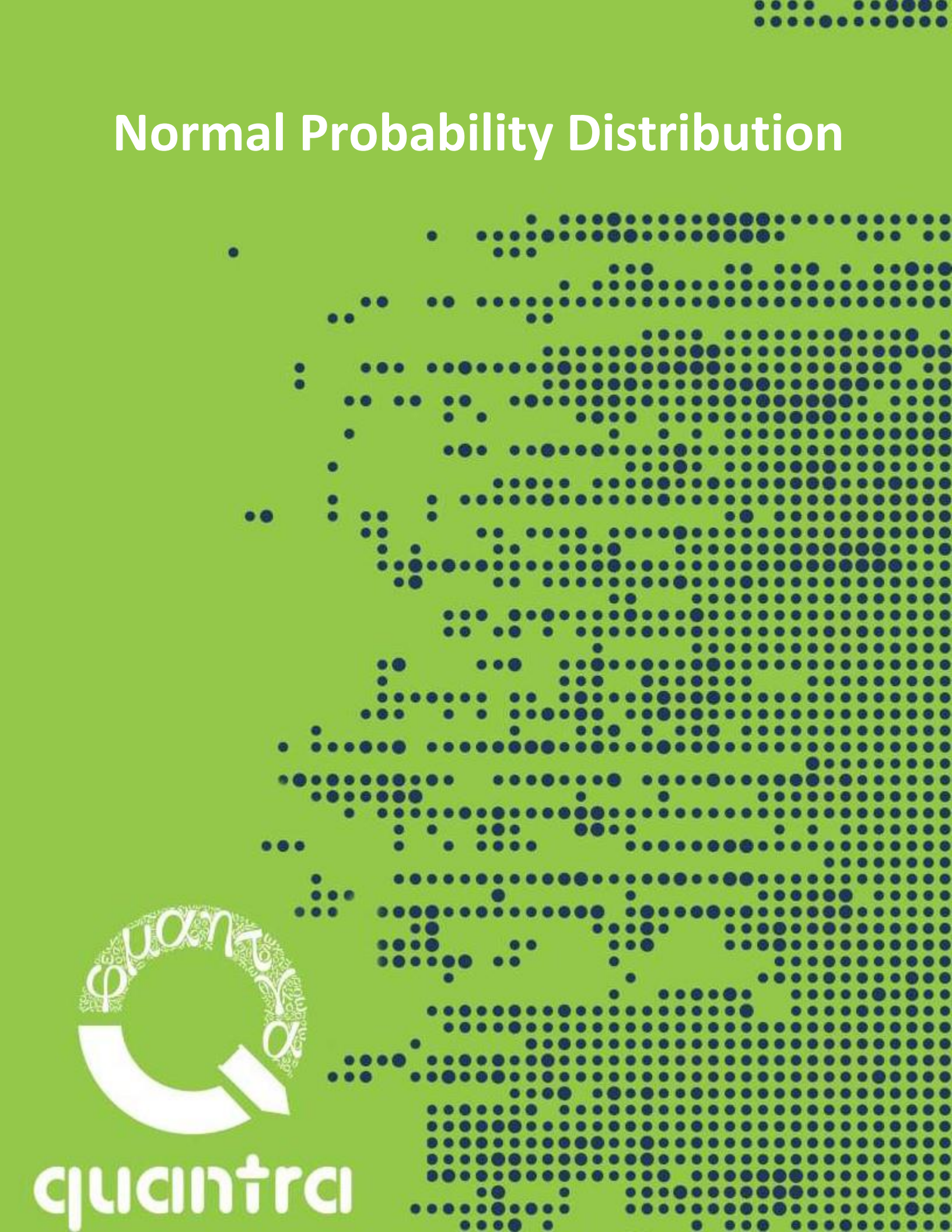
Normal Probability Distribution

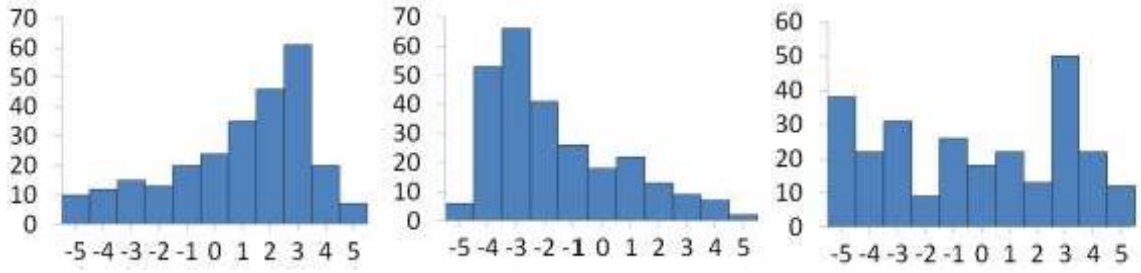


Contents

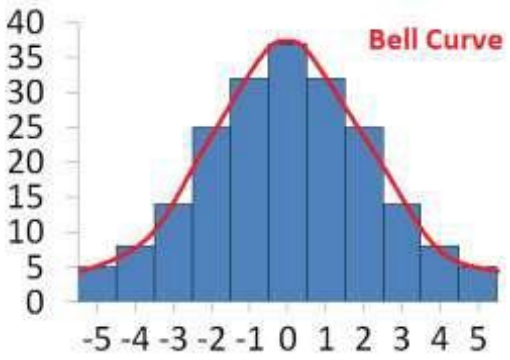
In this unit, you will learn about Normal Distribution. It is a very common continuous probability distribution function. Probability distribution is a mathematical function, which can be thought of as providing the probabilities of occurrence of different possible outcomes in an experiment. It is a description of a random phenomenon in terms of the probabilities of the events. Let us now look at Normal Distribution in more detail.

Normal Distribution

Normal distribution is the most fundamental way of distribution, where a central value has the maximum probability of occurrence. Values to the left and right of this central value have reduced probabilities of occurrence that are symmetric across the central value. Such kind of distribution can be observed in a lot of different places in nature. For example, the heights of humans are concentrated the most near 5'7", heights greater and lesser than this central value has decreased probabilities of occurrence as we go further away from the central value. Similarly, the distribution of weight, IQ, examination scores, blood pressure, etc. are all examples of normal distribution. It is an important distribution for representing real valued random variables, like the price of an asset, whose distribution is unknown. Data can be distributed or spread out in different ways. Look at the histograms below for a better understanding. This data is for daily returns of different assets. Each vertical column in the graph represents the total number of days for which the returns lied within a given range. The x-axis stores the values of daily returns and the y-axis indicates the number of days.



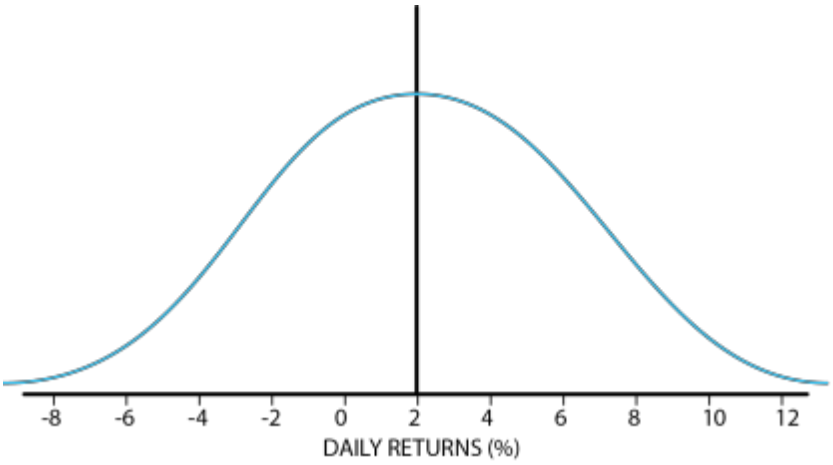
From these histograms, we get an understanding about which range of values occurs more frequently than others, or in other words, which values have a higher probability of occurring. We can see that for the three different assets, daily returns are spread out in different ways across the center of the graph, i.e. zero daily returns. The leftmost histogram represents data that is more spread out on the right or has more positive daily returns, which implies that the security is going through a bull run and is on the rise. The histogram in the center represents data that is more spread out to the left or has more negative daily returns, implying that the security is going through a bear run. The rightmost histogram represents data that is jumbled up and has no bias towards the positive or the negative side, which implies that the asset is moving sideways or is in consolidation. But there are several cases where the data tends to be around the central value with no bias towards the left or the right and is distributed evenly as shown in the plot below.



This type of distribution is called Normal Distribution and is also referred to as the bell curve because of its close resemblance to the shape of a bell. In the context of trading, it is assumed in many pricing models that the returns on an asset are normally distributed. For any trading strategy, the aim is to study the risk and returns of the trade. A trader wants to maximize his returns and minimize the risk of the trade. Both of these can be assessed by plotting the distribution function of the asset. The mean of the distribution represents the average return one could expect and the standard deviation of the distribution is the measure of risk. Let us now look at these in more detail.

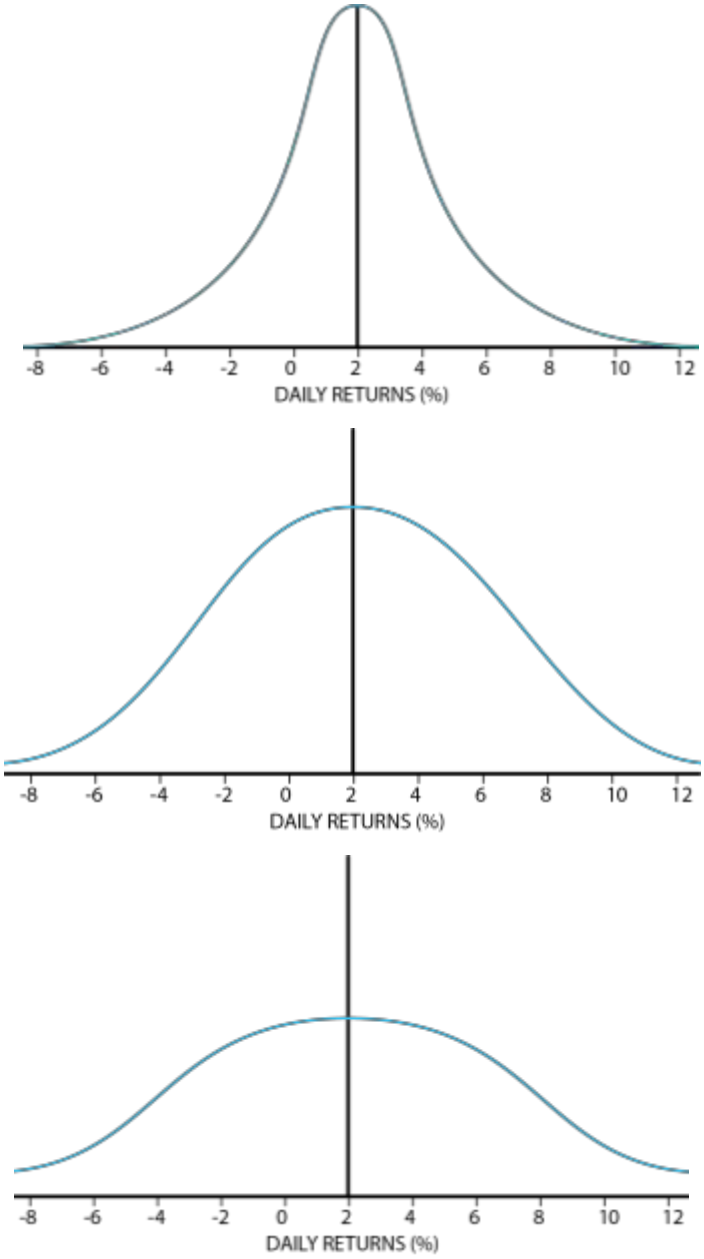
Mean or Expected value

The mean of a particular asset's return might say 2%, this implies that on an average over a certain time, the asset rises by 2% daily. This value can be computed by calculating the average on a large enough dataset containing historical data of asset prices. Look at the plot shown below; the mean of the Normal Distribution is at the center of the curve which is at 2% as shown. For the Normal Distribution function, the mean, median and mode all have the same values owing to the symmetry across the mean.



Standard deviation

The standard deviation indicates the amount by which the values deviate on an average from the mean. A higher standard deviation would imply that the returns on the asset vary over a broader set of range. Hence, the trade will be riskier as it leads to more uncertainty. Let us now look at the graphical representation of the same:



Therefore as seen above, the graphical representation of a Normal Distribution of daily returns through its mean and standard deviation enables understanding both returns and risk for an asset. The central value or the mean tells us the average returns that can be expected from the asset. The standard deviation depicts how much the returns curve stretches out taking extreme values, which gives us an insight into what the risk for investing the asset will be. If the graph is stretched out further, it would mean that we are at risk of losing more money than that for an asset with a lower standard deviation with a lesser probability of occurrence. An asset with a mean 1.5 and standard deviation of 2 is a riskier investment than an asset with a mean of 1.5 and standard deviation of 1.

It can be seen from the graphs above that the asset with standard deviation of 2 has higher probabilities of occurrence for more extreme values. These extreme values could be positive as well as negative; therefore it is a riskier bet to invest in such an asset.

The representation of Normal Distribution for these two assets will be N(1.5, 2) and N(1.5, 1). The Normal Distribution N(0, 1) having a mean value of zero and standard deviation of one is called the standard Normal Distribution. Knowing these values for each asset makes the investor aware of the expected returns and the risk of an investment.