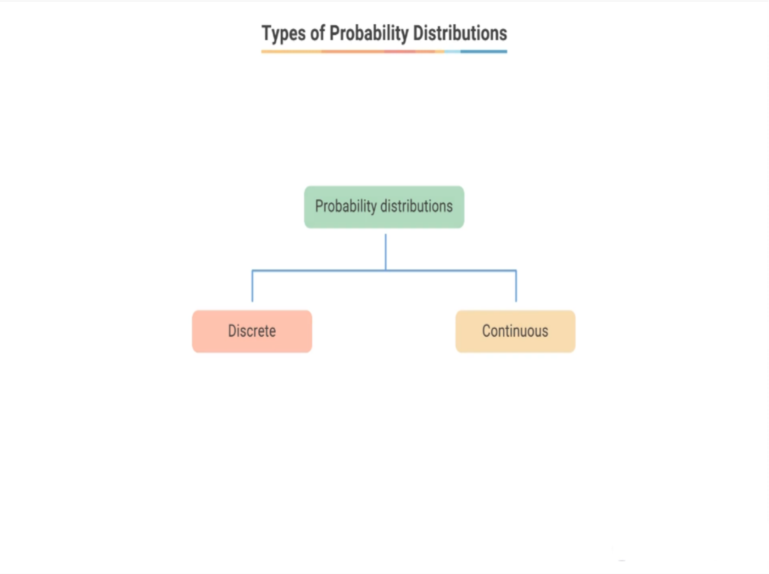
Artificial Intelligence (AI) for Investments
Prof. Abhinava Tripathi
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

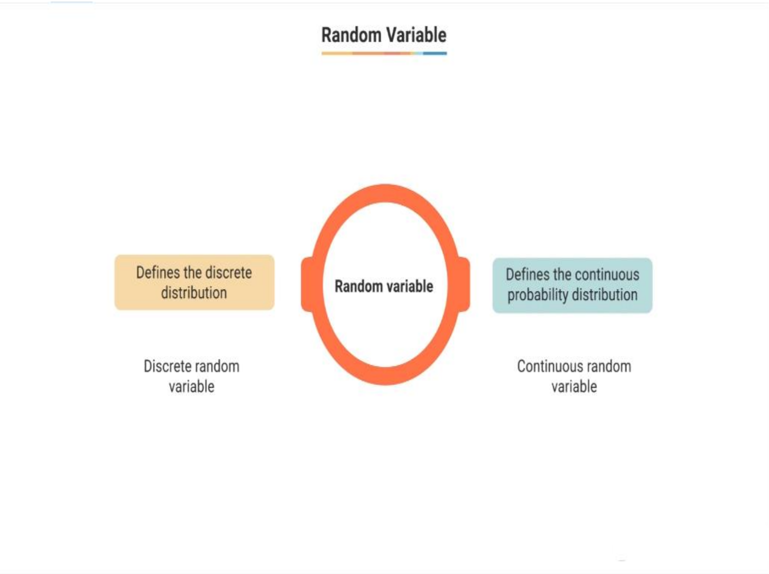
Lecture - 14

(Refer Slide Time: 00:16)



Continuous random variables. Probability distributions are of two types, one is discrete and other is continuous probability distribution.

(Refer Slide Time: 00:22)

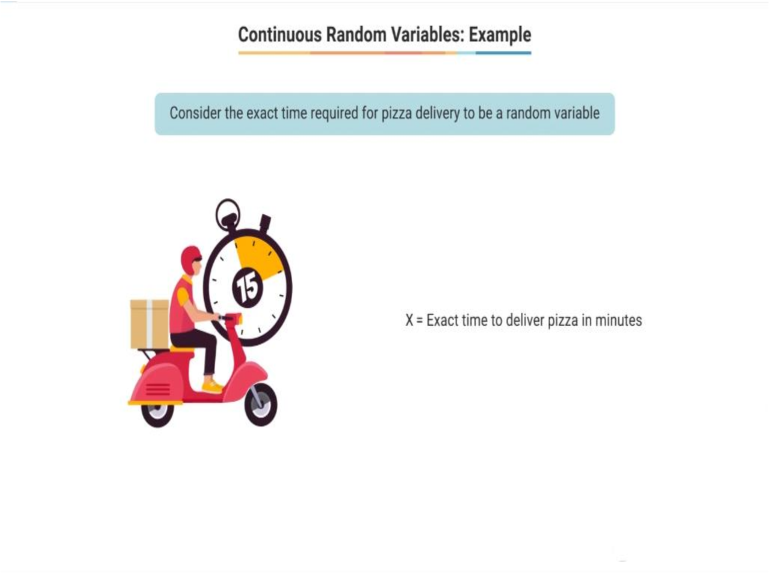


The random discrete variable used to define the discrete distribution is called discrete random variable and for a continuous probability distribution we use the continuous random variable. So, let us try to understand what a continuous random variable is.

(Refer Slide Time: 00:37)

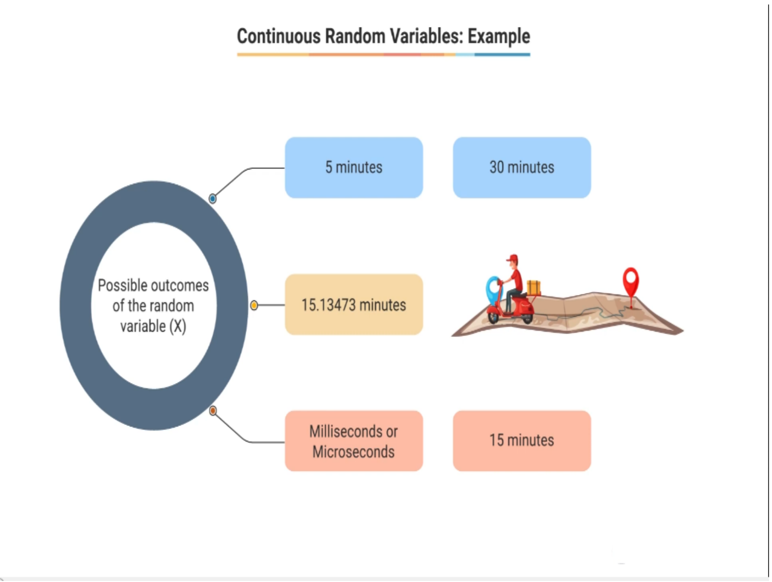
Let us consider the example, all of us have ordered a pizza at some times in our life. So, consider yourself as a manager at one of the pizza delivery outlets. Now, as a delivery manager you are concerned with the average time it takes for a pizza delivery to reach a customer.

(Refer Slide Time: 00:52)



So, let us consider a random variable which is the exact amount of time required for a pizza delivery to reach a customer. Let us name this capital X and let us say that this X is calculated in minutes.

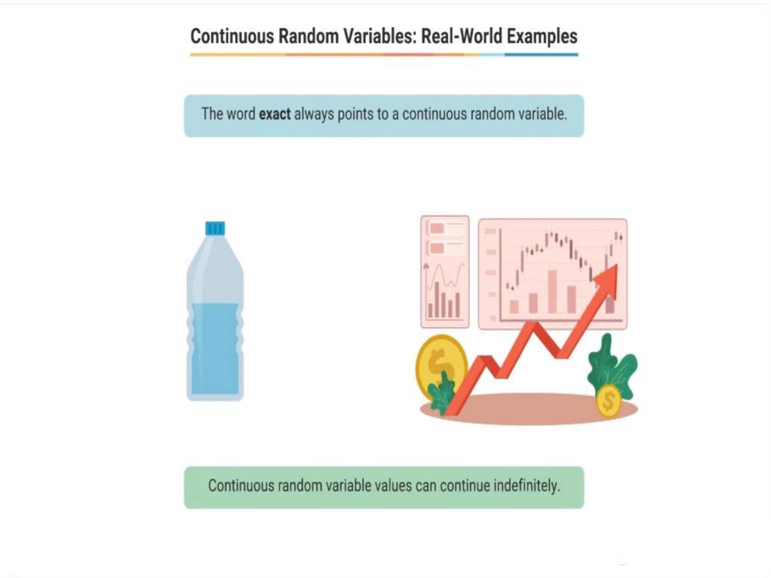
(Refer Slide Time: 01:04)



Now, let us list down all the possible outcomes of this random variable, it could be 5 minutes which means that the customer was probably in the neighbouring building or it could be 30 minutes it could also be something like 15.13473 minutes. You can even think up to the last millisecond or up to the last micro second level even 15.13473 might not represent the exact amount of time because we could go up to 100 of decimal points rate.

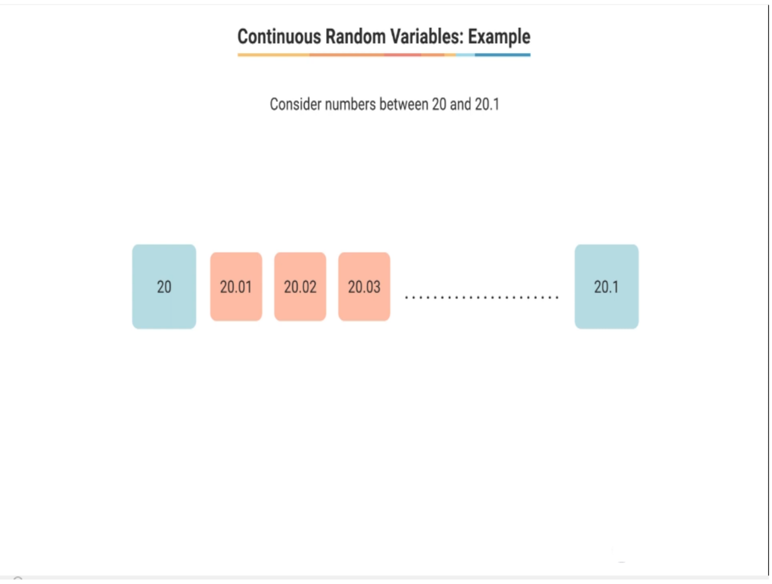
So, technically it could be a whole number like 15 but it could also be a decimal and could go on till infinite decimal points.

(Refer Slide Time: 01:38)



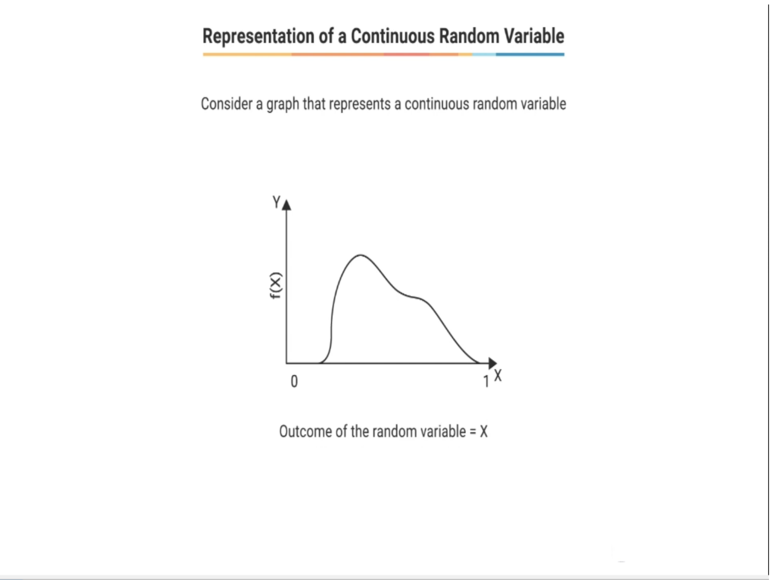
So, when you use words like exact amount of time or say the amount of water present in a bottle or exact stock price at the end of a trading session or anything that is generally exact this is always going to be a continuous random variable. And the reason is that the moment you say exact the values that the random variable can take is going to be infinite.

(Refer Slide Time: 01:57)



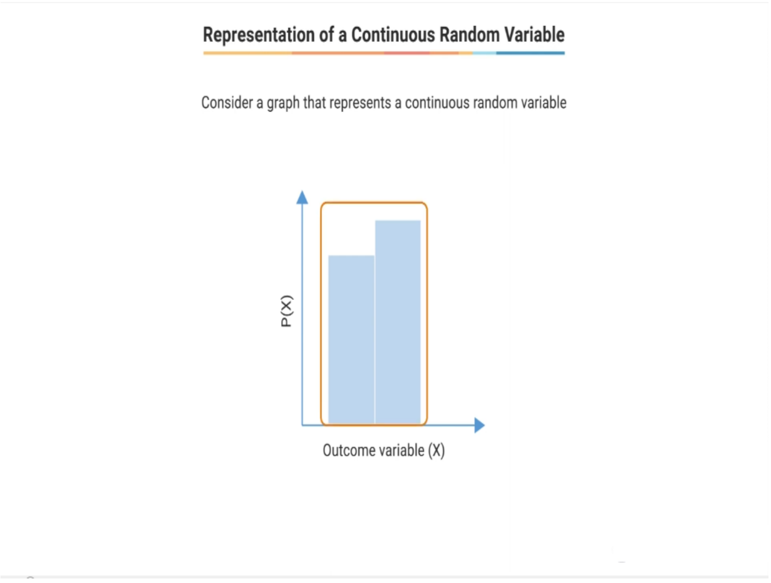
Even just between 20 and 20.1 there could be a million values, whenever you have such kind of random variables, they are called continuous random variables. So, to understand how we actually represent a continuous random variable on a graph.

(Refer Slide Time: 02:06)



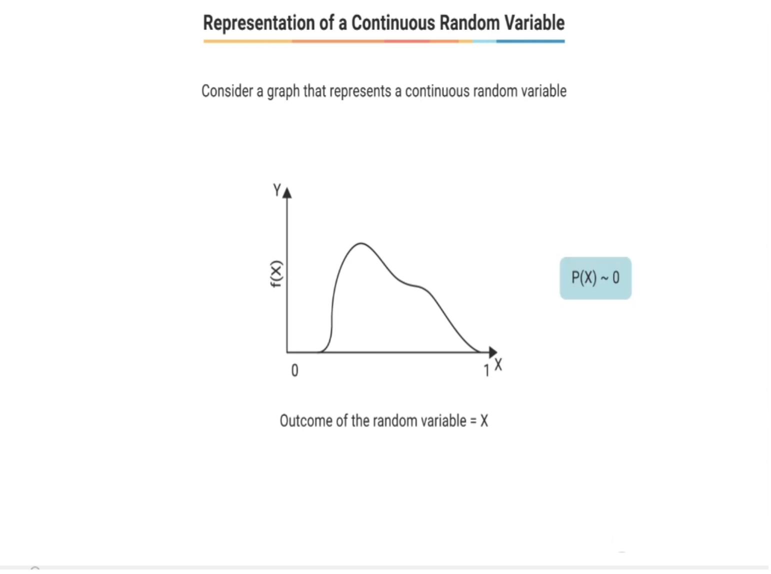
So, let us create a plot with X axis as the outcomes of the random variable.

(Refer Slide Time: 02:15)



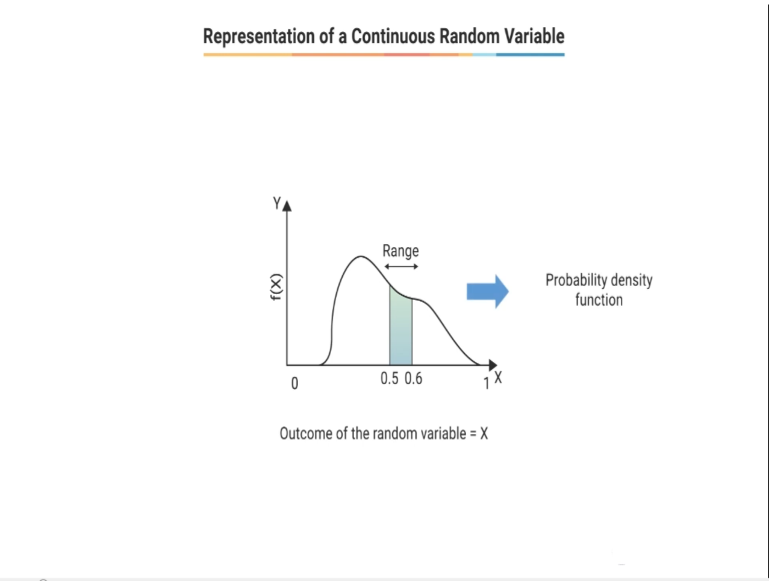
Now, in case of a discrete random variable remember that the y-axis used to represent the probability value for that specific value of x and the plot would look very similar to a histogram plot.

(Refer Slide Time: 02:27)



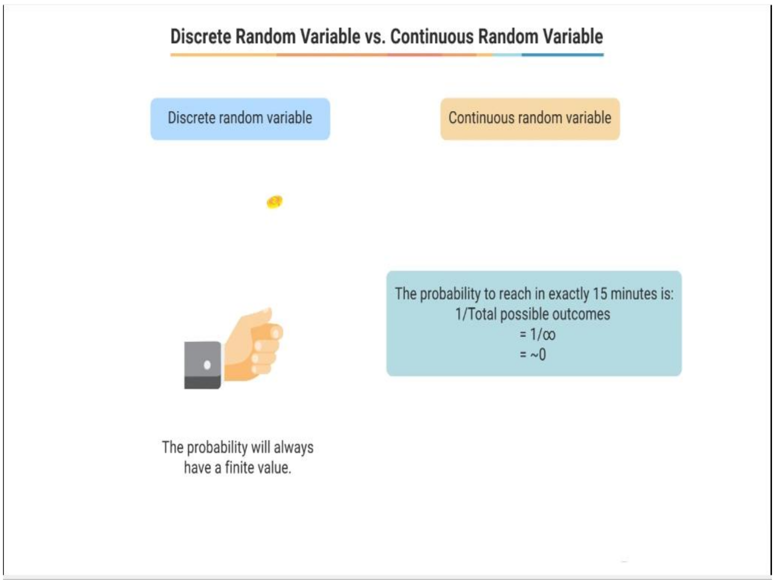
But in the case of a continuous random variable that is not possible, this is because you cannot define the probability for any specific value of x and to be too small almost close to 0.

(Refer Slide Time: 02:36)



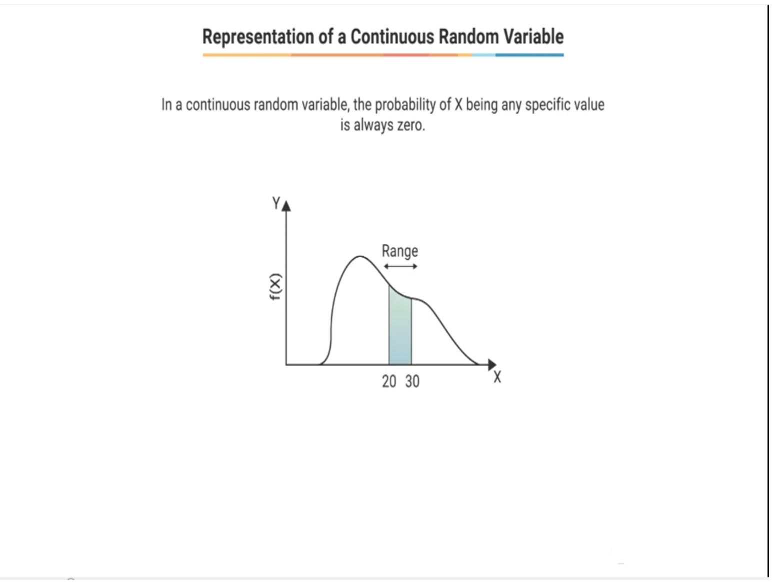
So, typically how you represent a continuous random variable on a graph as the name suggests it will have a continuous line and this graph is known as the probability density function.

(Refer Slide Time: 02:46)



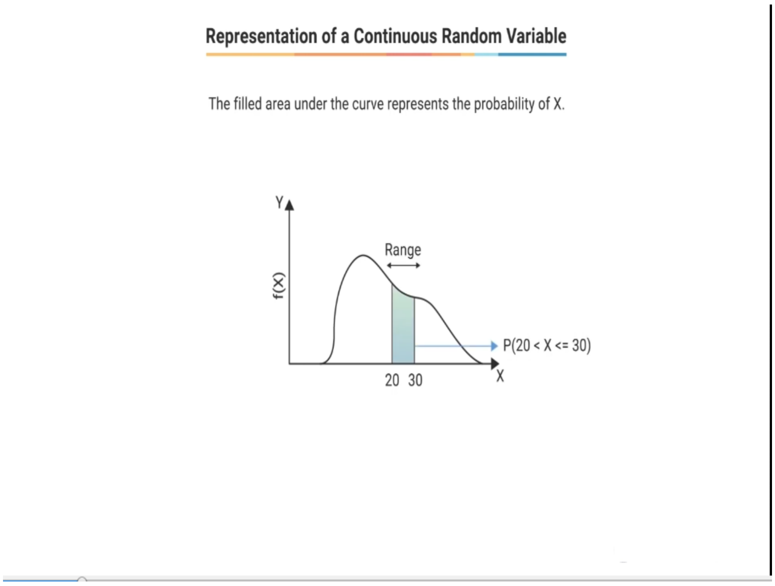
Now, in the case of a discrete random variable we used to calculate the probability by saying probability of x = 0 or probability of x = 1 or we would simply plot these values on the plot. Now, think in the case of continuous random variables. If I go back to the example of commute time to a customer and ask you what is the probability that X will be exactly 15 minutes since the possible outcomes that X can take is almost infinite. So, the probability that X will be exactly 15 will be almost 0 and this is true for any value of x.

(Refer Slide Time: 03:15)



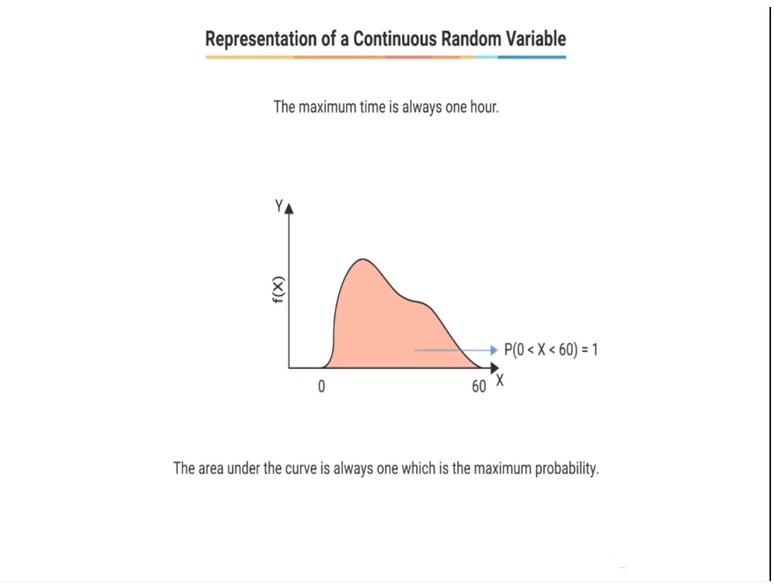
Therefore, we can say that the probability of X being any specific value is always 0. So, instead what we do we measure the probability of X lying in a certain interval let us take an interval from 20 to 30 minutes.

(Refer Slide Time: 03:30)



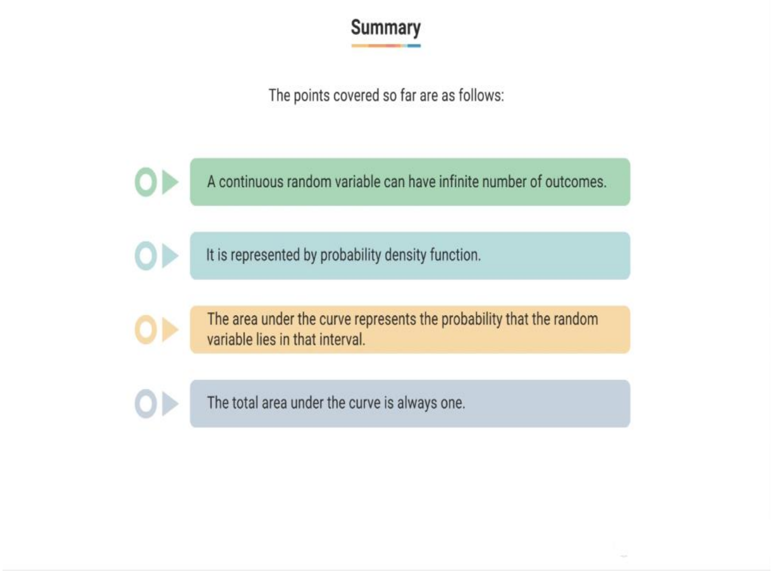
The way we will find this probability from the graph is that it will be the area under the curve between 20 and 30 minutes. So, the area that has been coloured in the plot is the probability that X lies between 20 to 30.

(Refer Slide Time: 03:40)



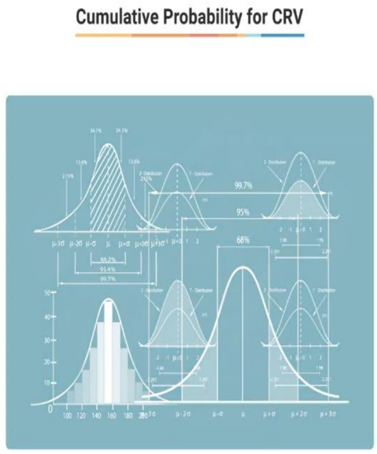
Now, let us assume that the maximum time is always 1 hour and any time above that is not considered in our data as that scenario is extremely rare then the commute time can range anywhere from 0 to 60 minute. Thus, a random variable X will also be a line that will range from 0 to 60 minute and total area under the curve from 0 to 60 will be 1 which is the maximum probability of x.

(Refer Slide Time: 04:00)



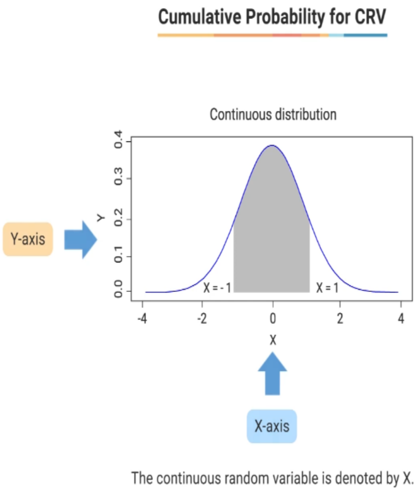
So, we will conclude this discussion by quickly going over the points that we have learned first a continuous random variable can have infinite number of outcomes. Second, we represent a continuous random variable using what we call a probability density function which is essentially a continuous line drawn for all the range of values that X can take. Thirdly the area under the curve represents the probability that the random variable lies in that interval and finally the total area under the curve will always be equal to 1.

(**Refer Slide Time: 04:32)**



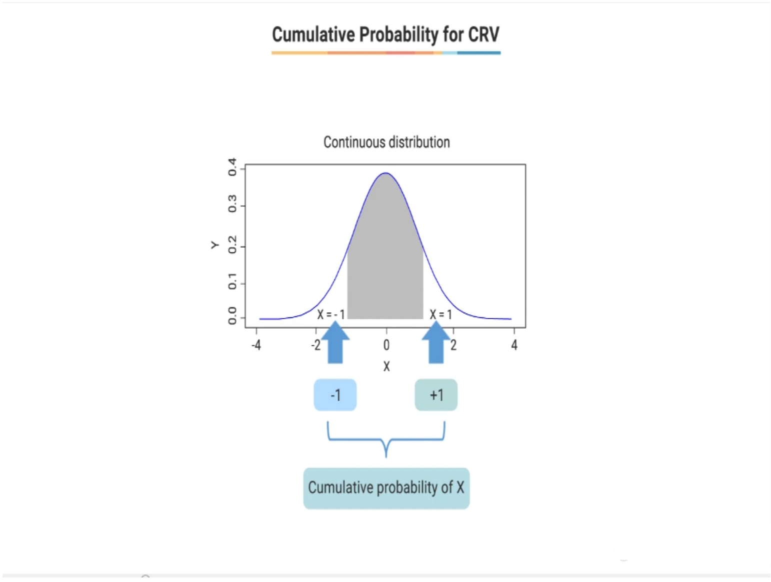
Continuous probability distributions cumulative probability for continuous random variables. Let us understand the concept of cumulative probability for a continuous random variable. So, let us take an example where we have considered some probability density function.

(Refer Slide Time: 04:39)



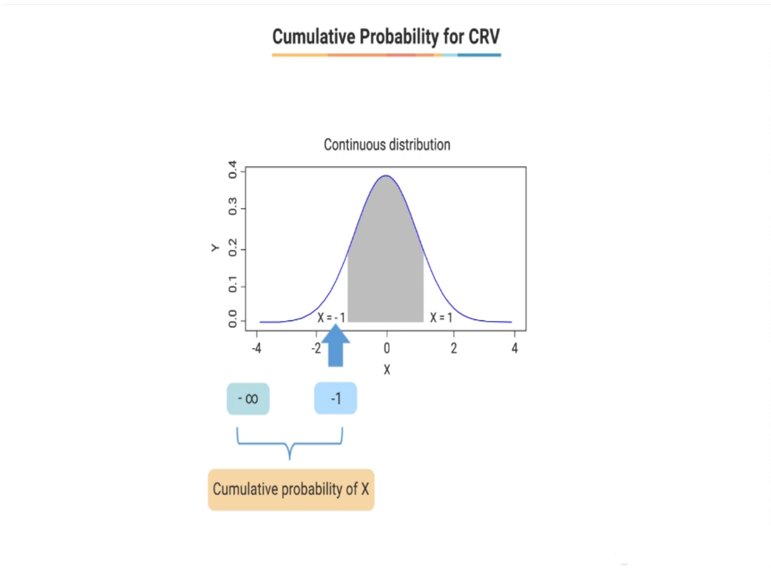
So, we have X and Y axis, so this is our continuous random variable let it be defined by capital X.

(Refer Slide Time: 04:48)



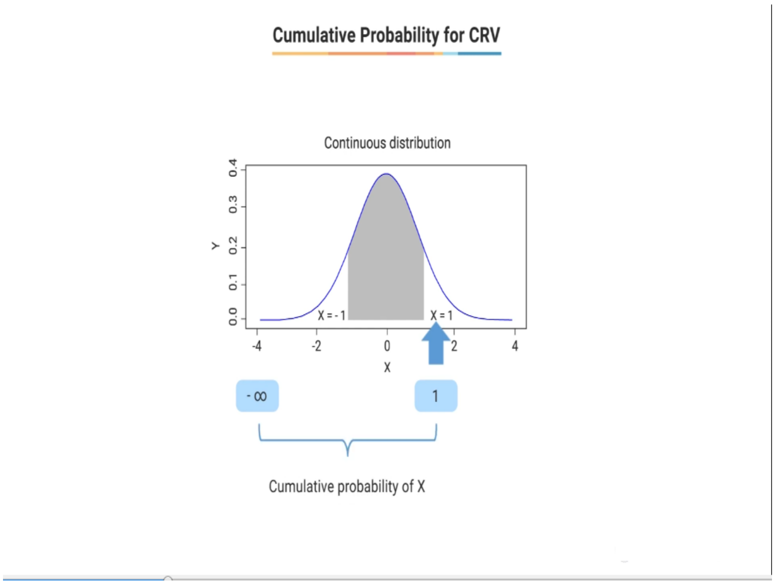
The X axis goes from plus infinity to minus infinity. On this chart we have also marked - 1 and + 1 for X random variable. And if we have learned that we want to calculate probability of between X between 1 to - 1 then that will be the area under the curve from minus 1 to plus one. Let us understand what will be the cumulative probability for this graph.

(Refer Slide Time: 05:08)



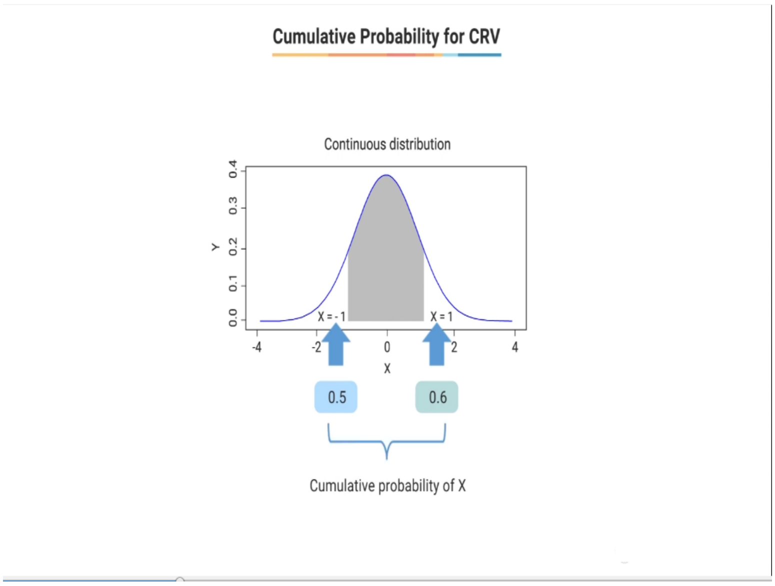
Suppose, we want to find the cumulative probability of x = - 1 what it signifies is that is going to be the area from - infinity to -1 though in this case it does not go till - infinity but basically till whatever point it goes, so this region is the cumulative probability of x = - 1 that is from - infinity to - 1.

(Refer Slide Time: 05:30)



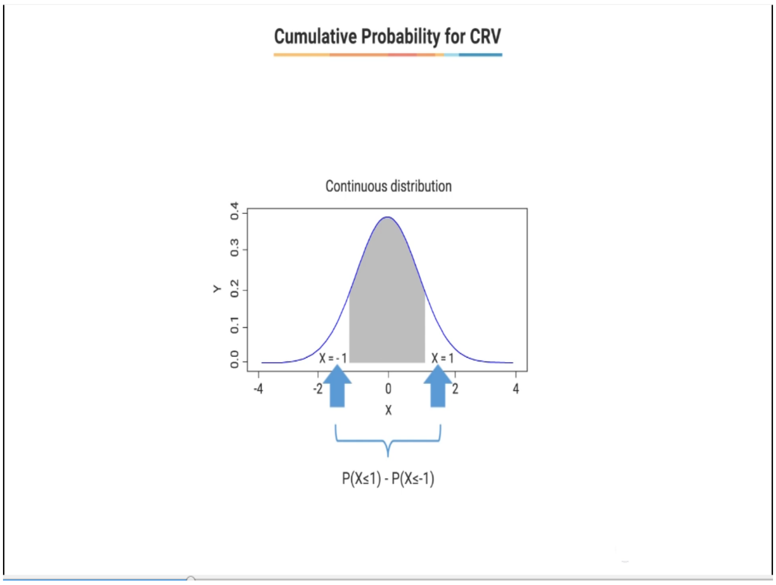
The same concept can be applied to find the cumulative probability at x = 1 and that would be denoted by the shaded region from - infinity to + 1 basically, all the values from - infinity to + 1. Now, the cumulative probability becomes very important when we talk about continuous random variables because in such cases we are always dealing with ranges.

(Refer Slide Time: 05:49)



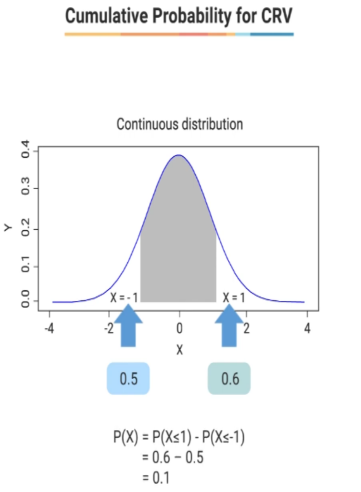
So, let us say you are given the cumulative probability for X = 1 and X = - 1 as 0.6 and 0.5 respectively. So, using these values can we find the probability of X lie from -1 to + 1.

(Refer Slide Time: 06:02)



We can say that the probability of X between - 1 to + 1 is nothing but $P(X \leq 1) - P(X \leq -1)$ and this value can give us the area under the curve from - 1 to + 1.

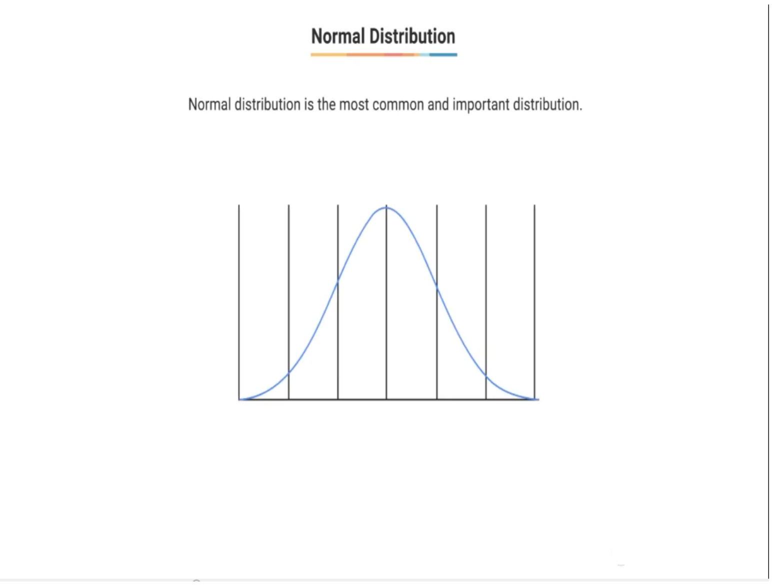
(Refer Slide Time: 06:16)



So, in this case probability of X less than 1 is 0.6 and probability of X less than - 1 is 0.5. So, the probability of X lies between + 1 to - 1 is 0.1.

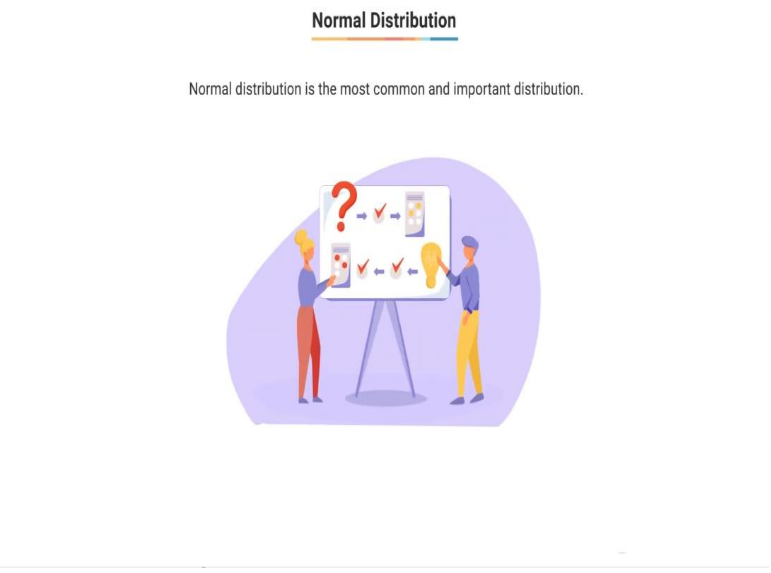
Continuous probability distributions, normal distribution.

(Refer Slide Time: 06:31)



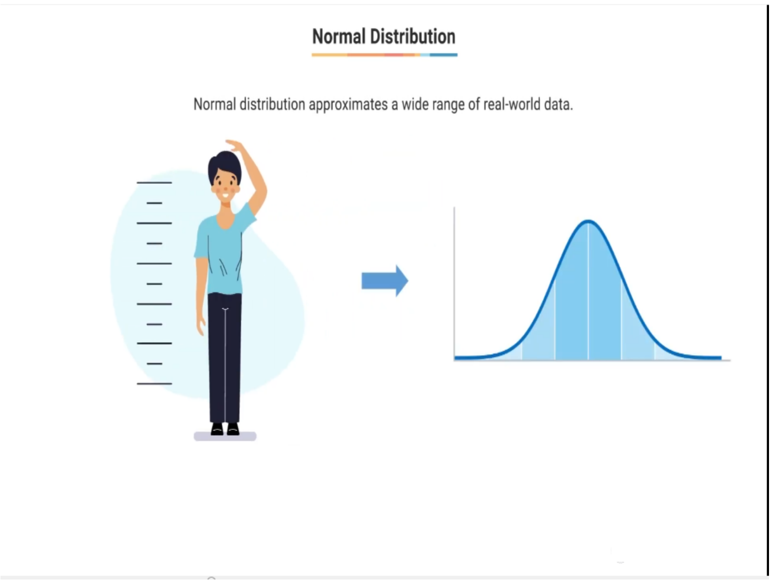
The normal distribution is perhaps the most widely used and most important distribution when talking about distributions of continuous random variables.

(Refer Slide Time: 06:40)



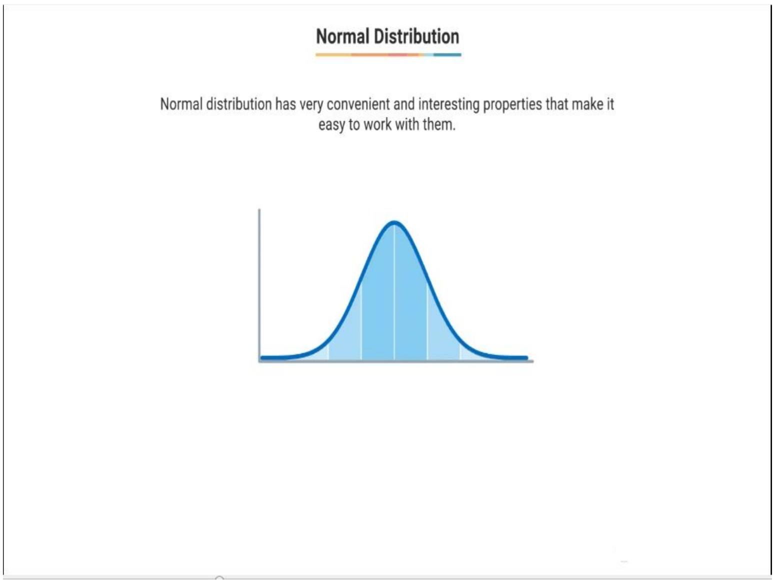
And there are a few valid reasons why this is the case there are many places where the normal distribution appears naturally.

(Refer Slide Time: 06:46)



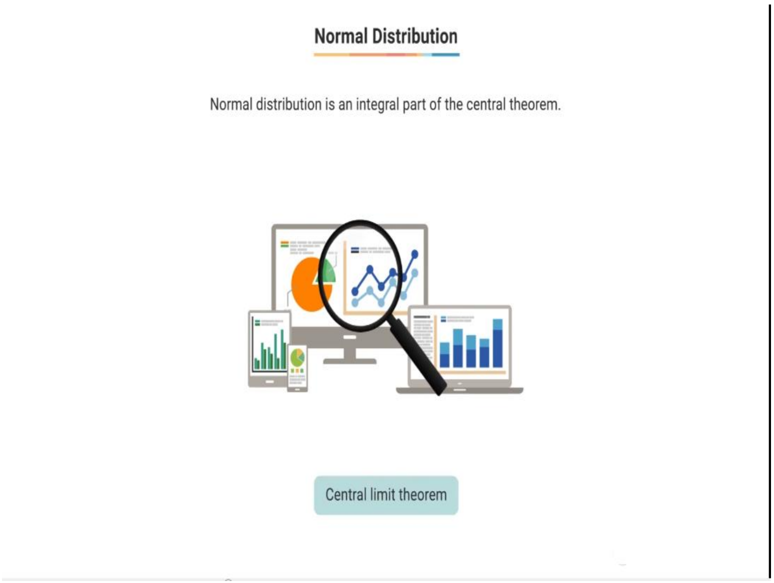
For example, if you take the height of an individual in a country it will most likely follow a normal distribution therefore the normal distribution approximates a lot of real life data.

(Refer Slide Time: 06:56)



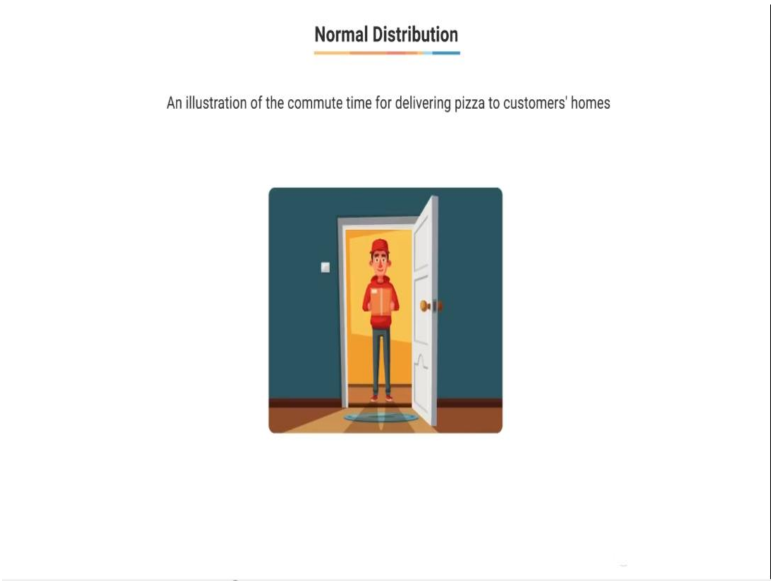
The second reason is that once we know that something follows a normal distribution and knowing that normal distributions have something some interesting properties it makes it very convenient to work with them.

(Refer Slide Time: 07:07)



And the third reason is that normal distribution appears in a very important theorem which will form the base of most of the inferential statistics problems that we will solve. This theorem is called the central limit theorem. For these reasons the normal distribution is an extremely important distribution.

(Refer Slide Time: 07:22)

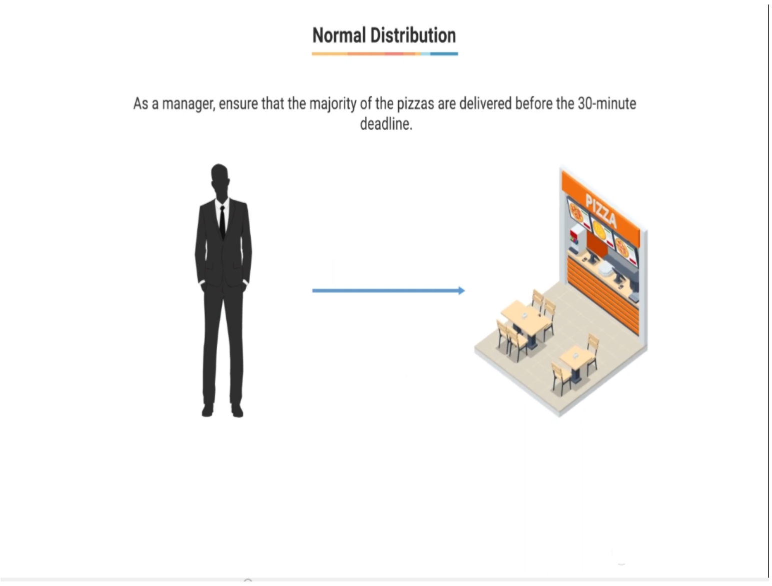


So, to understand normal distribution, let us go back to our earlier example of commute time for pizza delivery to the customer's houses.

(Refer Slide Time: 07:31)

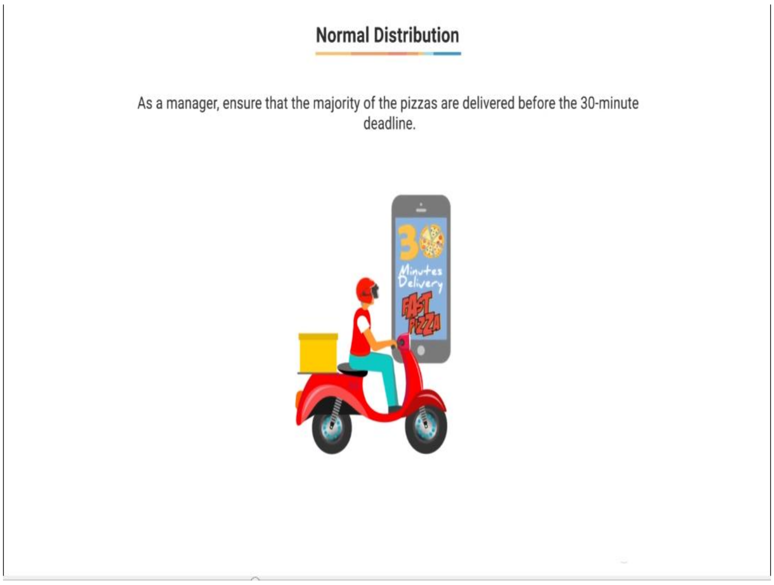
We are aware that pizza delivery outlets have come up with this 30 minute guarantee of delivering pizza from the time an order is placed. If the time taken to deliver the pizza is more than 30 minutes then depending on the purchase value, the pizza comes free or with a huge discount on total purchase.

(Refer Slide Time: 07:48)



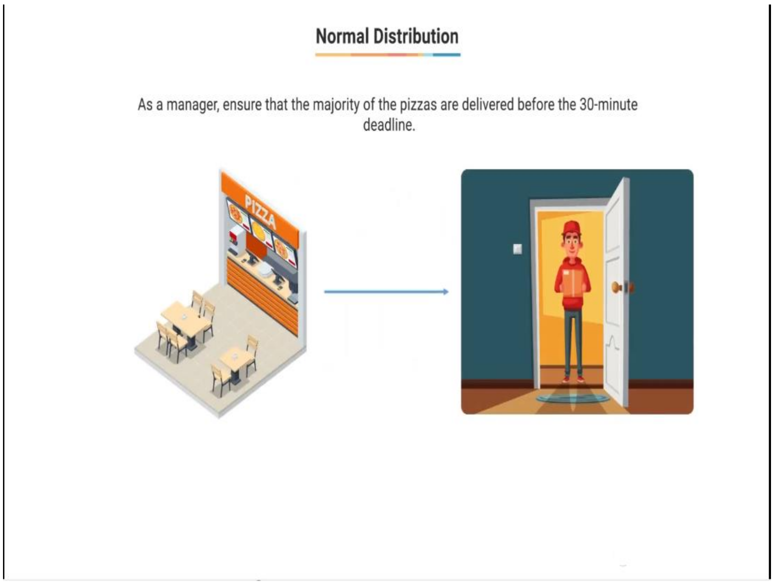
As a manager at one of the pizza outlets and given this 30-minute guarantee.

(Refer Slide Time: 07:55)



You want to ensure that most of the pizzas are delivered well before these 30 minutes.

(Refer Slide Time: 08:01)

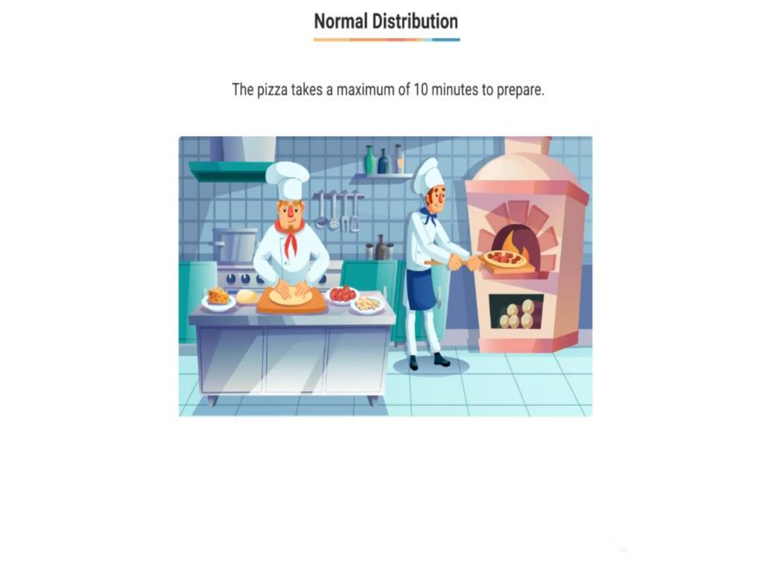


For now, let us consider only those commutes from the pizza outlet to the customer's location and not include the deliveries,

(Refer Slide Time: 08:04)

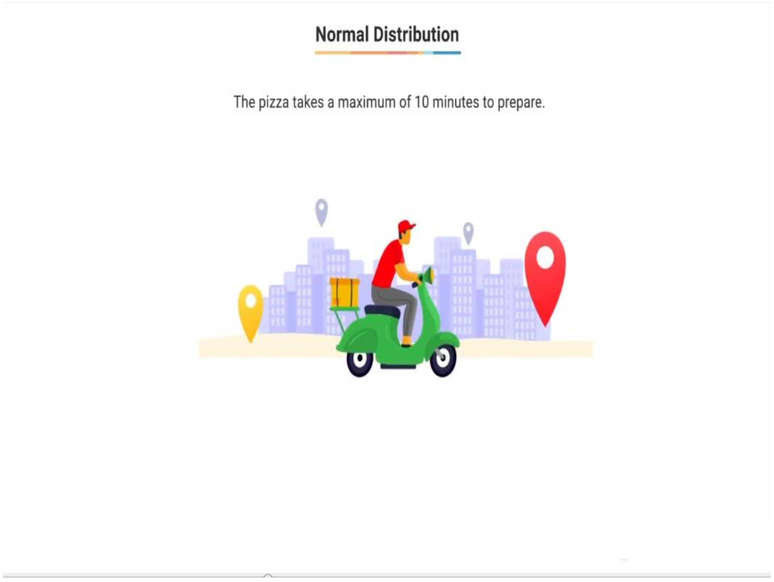
where the delivery boy has to visit multiple locations as he is delivering multiple orders.

(Refer Slide Time: 08:10)



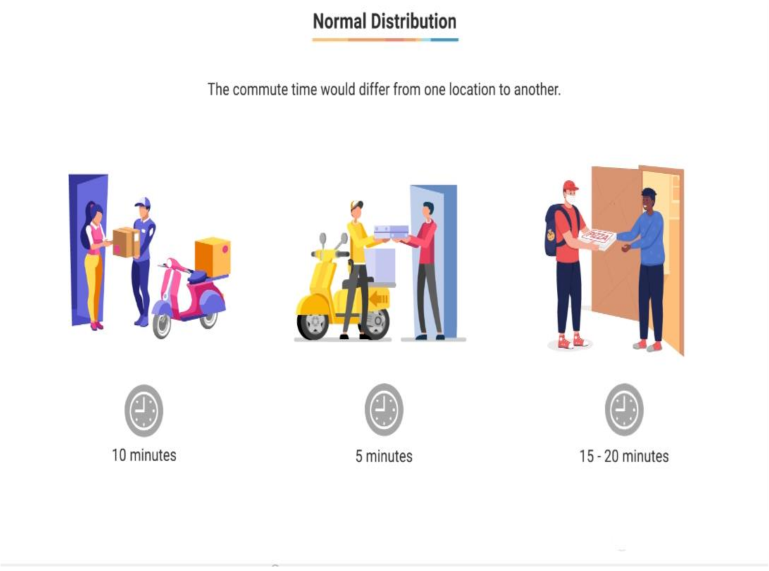
So, assume that it takes a maximum of 10 minutes to make the pizza.

(Refer Slide Time: 08:15)



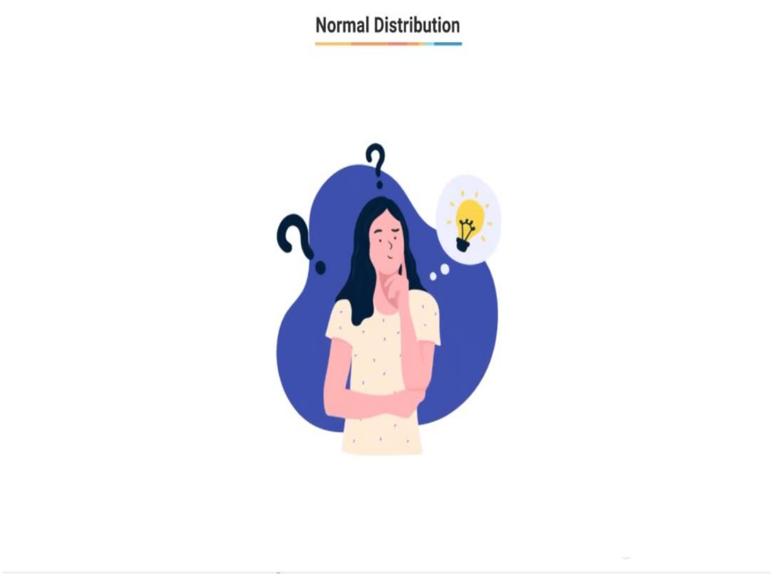
That leaves us with only 20 minutes to deliver the pizza from the outlet to the customer's location.

(Refer Slide Time: 08:19)



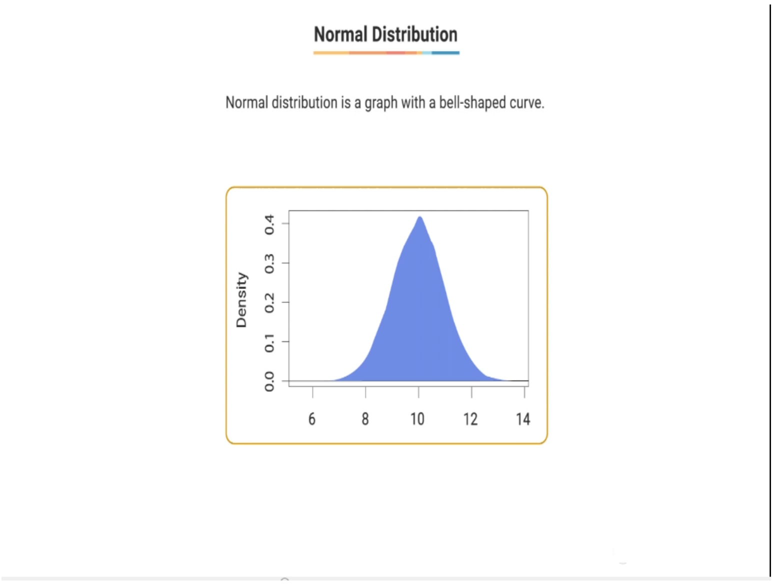
The commute time would obviously vary from one location to another for one customer the commute time could be 10 minutes while, for another customer the commute time could be just 5 minutes while, for yet another customer the commute time could go up to 15 to 20 minutes.

(Refer Slide Time: 08:35)



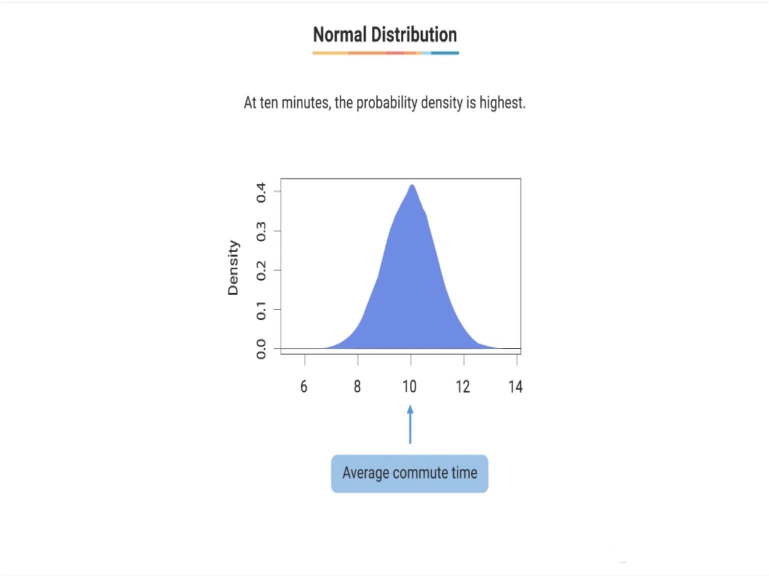
How do you think the probability density function for such a scenario might turn out.

(Refer Slide Time: 08:39)



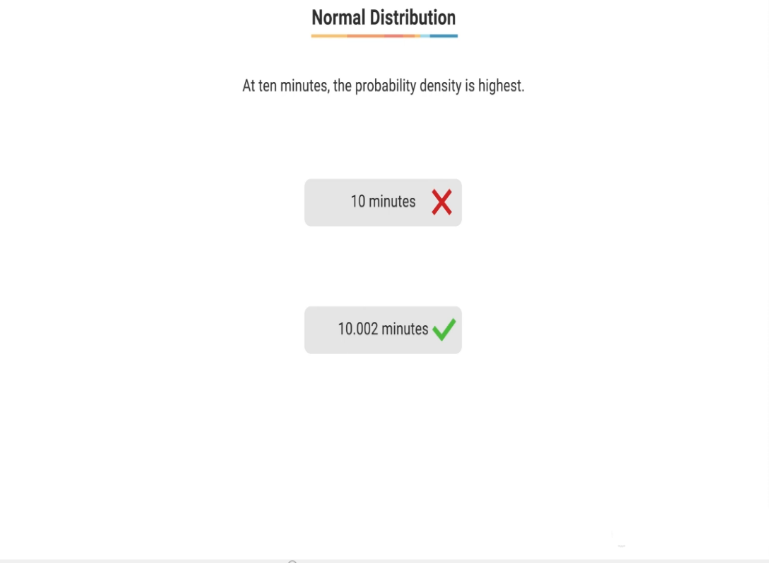
When we plotted the graph, we got a bell-shaped curve that looks like this as shown here. This chart is known as normal distribution. Let us try to understand this chart.

(Refer Slide Time: 08:48)



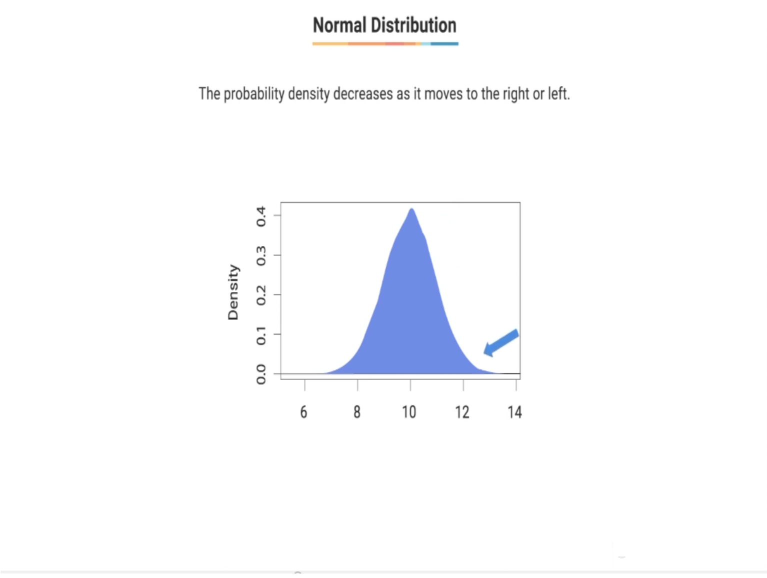
You can see that the probability density is highest at 10 minutes this shows that the average commute time is 10 minutes and in most cases commute time will be close to 10 minutes.

(Refer Slide Time: 08:58)



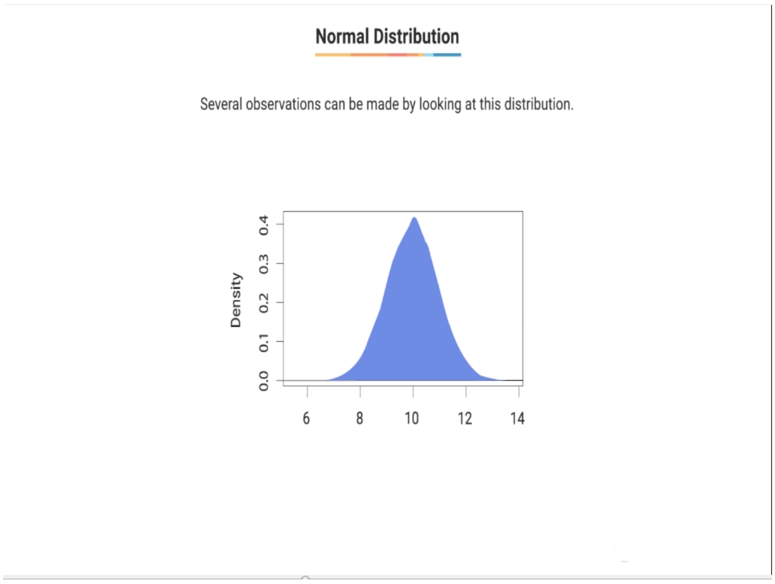
Obviously, it may not be exactly 10 minutes it might be a value like 10.002 minutes but, we shall ignore this difference for now.

(Refer Slide Time: 09:06)



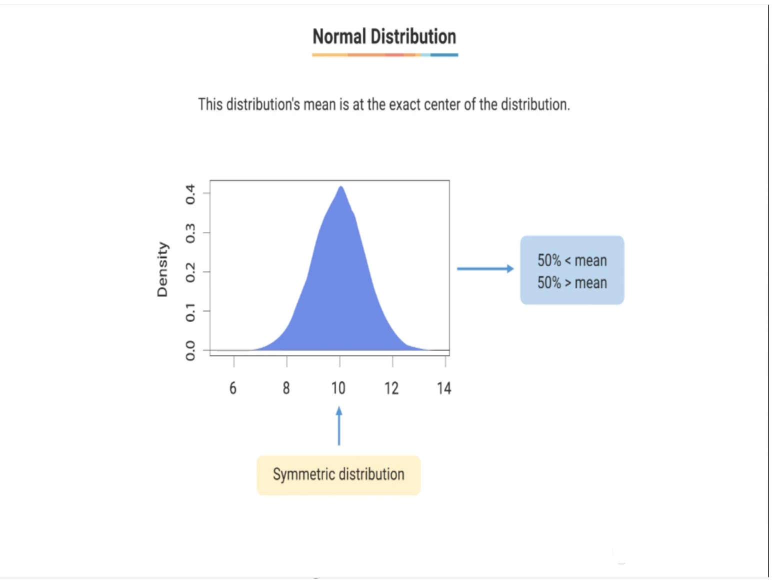
We can also see that the probability density starts decreasing as we move towards right or left such a plot is known as normal distribution.

(Refer Slide Time: 09:17)



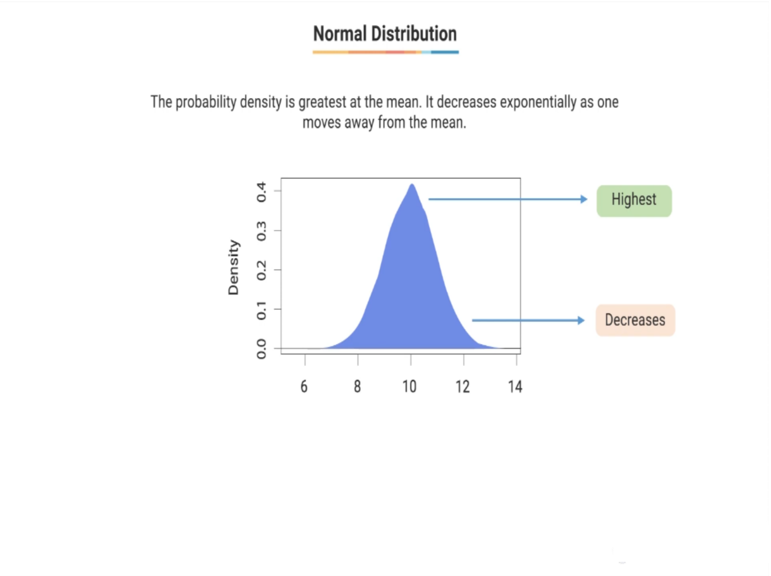
Just by looking at this distribution we can take several observations about the characteristics of this distribution.

(Refer Slide Time: 09:21)



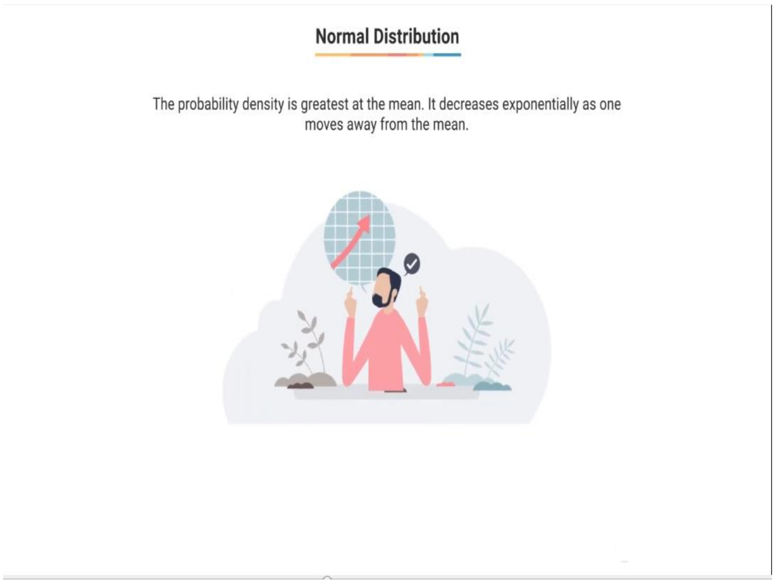
First, the mean of this distribution which is in our case 10 minutes lies in the exact centre of this distribution and this distribution is symmetric around its mean. Second, since the distribution is symmetric around the mean which is 10 minutes it means that 50 percent of the values are less than the mean and 50 are greater than the mean.

(Refer Slide Time: 09:38)



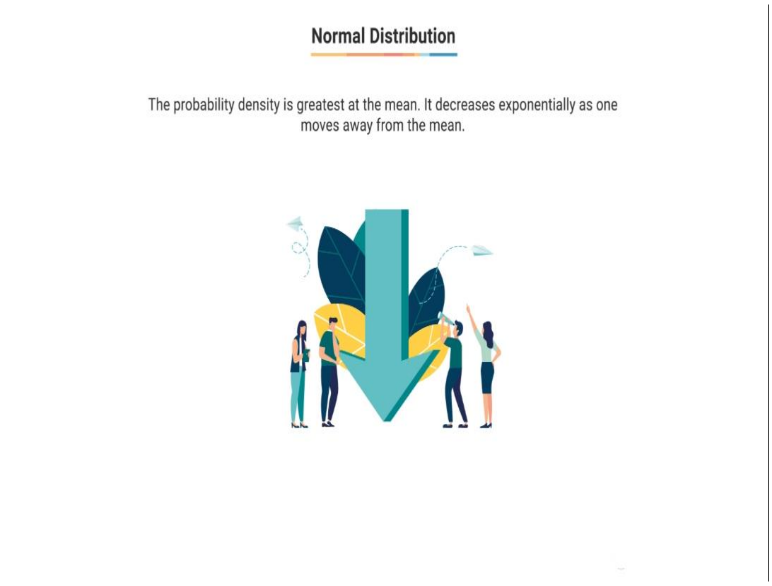
This kind of a shape is called a bell curve as it looks like the shape of a bell. Third we can see that the probability density is highest at the mean and decreases exponentially as we move further away from the mean.

(Refer Slide Time: 09:51)



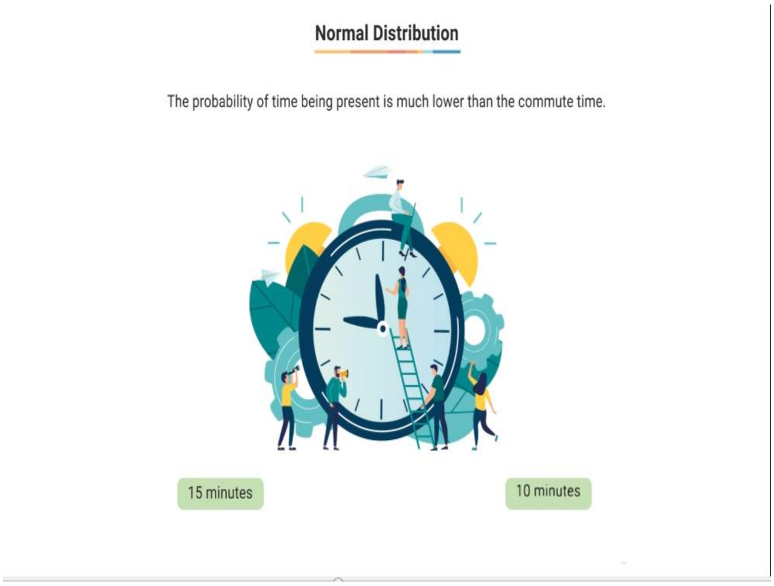
In simple language it means that there is a high probability that the value of the random variable is close to the mean.

(Refer Slide Time: 09:57)



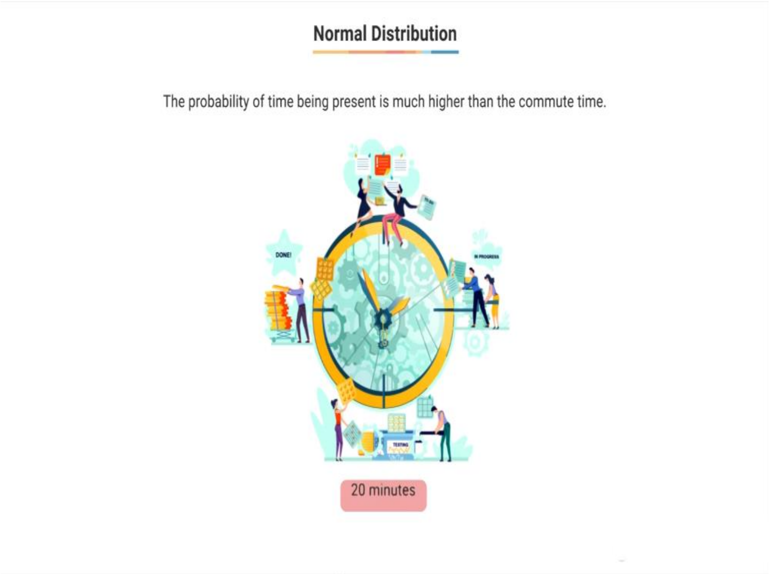
As we move further away from the mean, the probability of the occurrence of such values decreases.

(Refer Slide Time: 10:00)



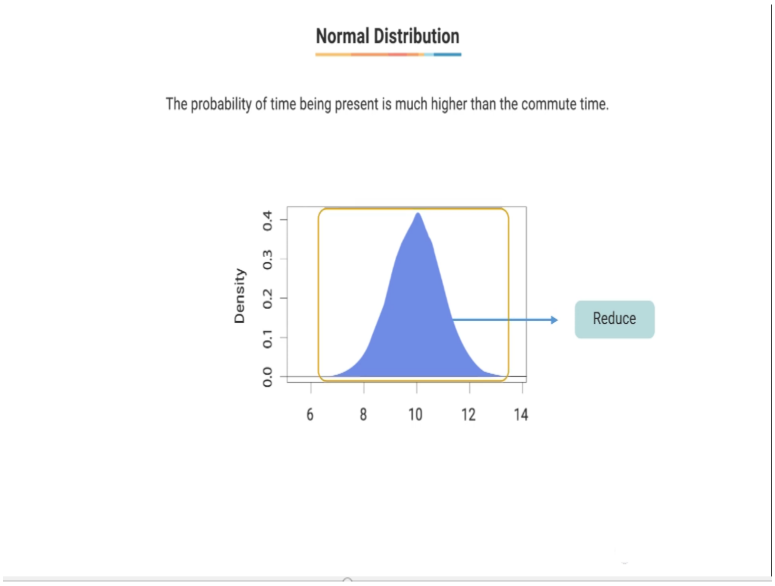
In our commute example this would mean that the probability of the time being around 15 minutes is much lower as compared to the probability of the commute time being around 10 minutes.

(Refer Slide Time: 10:10)



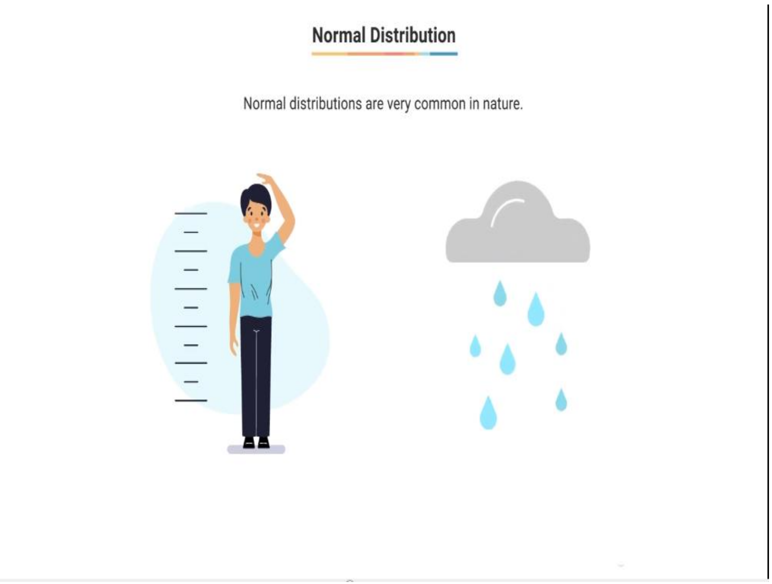
But it would still be higher compared to the probability of the commute time being around 20 minutes.

(Refer Slide Time: 10:16)



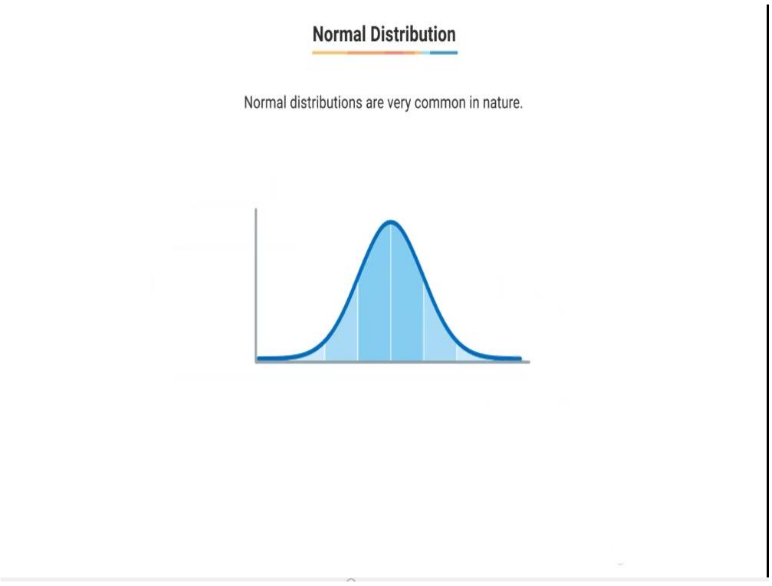
Similarly, if we go below 10 minutes then the probability density again starts to reduce and we would get a plot similar to the one shown here. So, this is how you identify normal distribution.

(Refer Slide Time: 10:26)



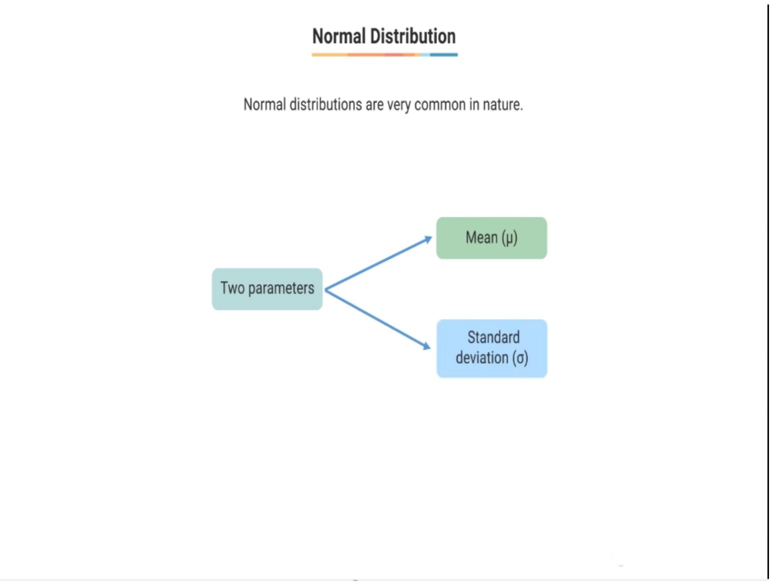
Such distributions are very common in nature, be it heights and weights of people amounts of rainfall and many other places.

(Refer Slide Time: 10:33)



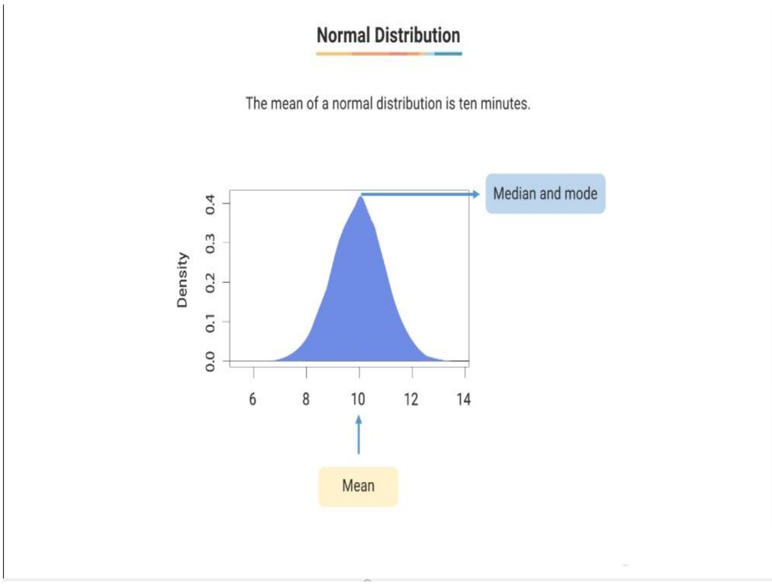
Let us not generalize any normal distribution using some parameters.

(Refer Slide Time: 10:38)



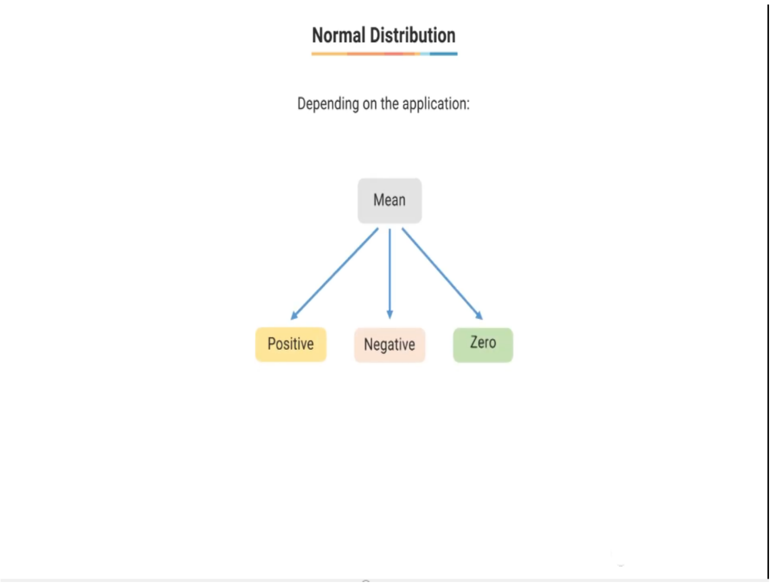
Any normal distribution can be defined using only 2 parameters mean μ and standard deviation σ.

(Refer Slide Time: 10:42)



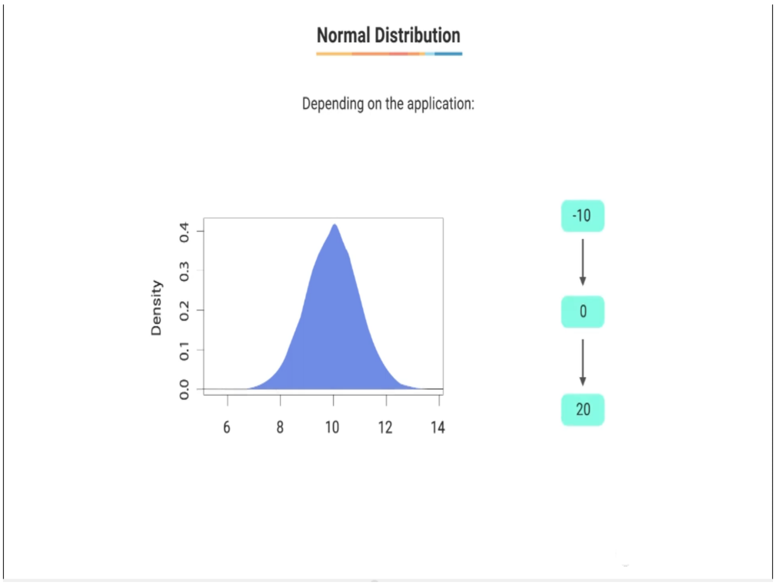
The mean which is μ is located at the centre of normal distribution and that point also denotes the median and the mode of distribution. In our case the mean was 10 minutes but in general it can be any value.

(Refer Slide Time: 10:54)



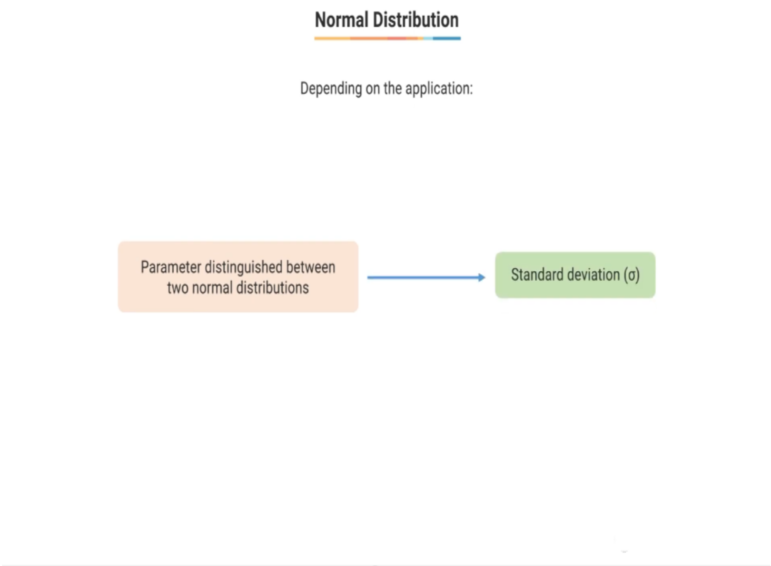
Depending upon the application, the mean could be positive negative or 0.

(Refer Slide Time: 10:59)



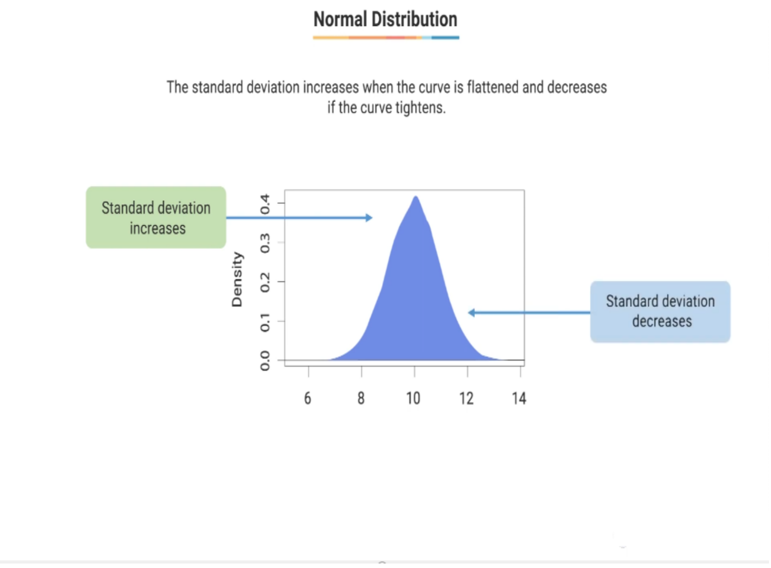
Here, you can see how the distribution changes as we are changing the mean from - 10 to 0 to 20.

(Refer Slide Time: 11:05)



The second important parameter to distinguish between two normal distributions is the standard deviation which we have denoted by σ

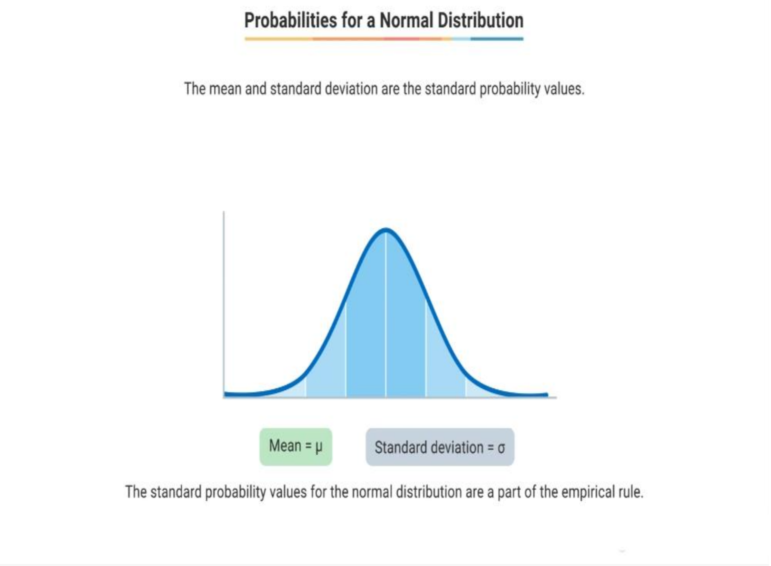
(Refer Slide Time: 11:11)



For example, if I flatten the curve you can see in the figure and they show that the standard deviation has increased. Similarly, if you try to narrow the same curve it shows that the standard deviation is decreasing.

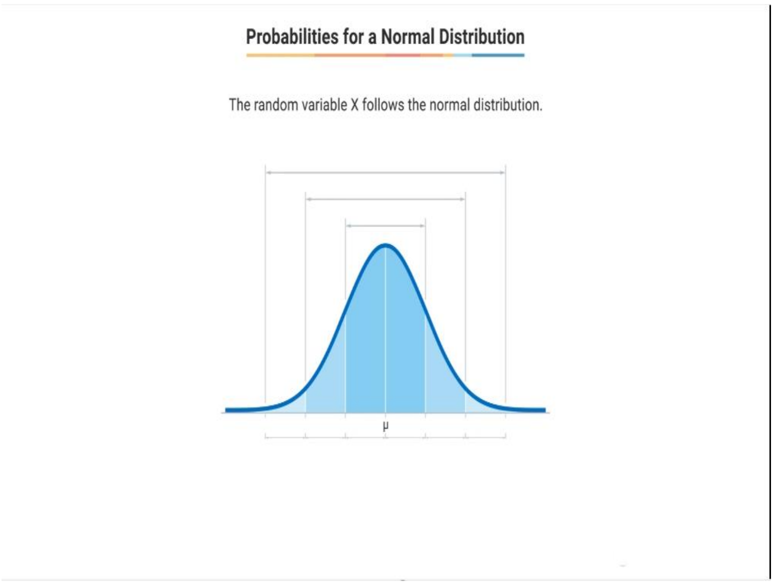
Probabilities for a normal distribution.

(Refer Slide Time: 11:25)



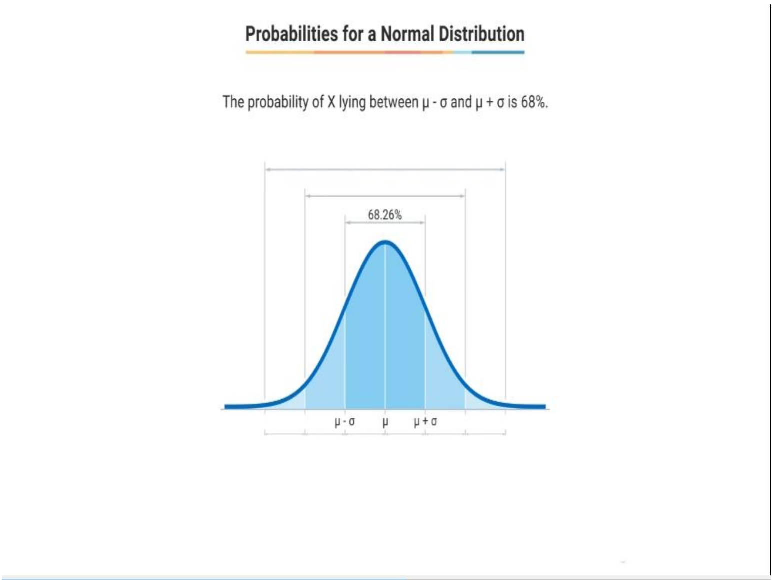
Suppose we have a normal distribution graph here and let us say that the mean of this distribution is μ we will assume that it is standard deviation is given by σ. Now, very important rule that will keep coming as we will see the normal distribution is the empirical rule. This is nothing but some standard probability values that we can easily calculate for any normal distribution.

(Refer Slide Time: 11:46)



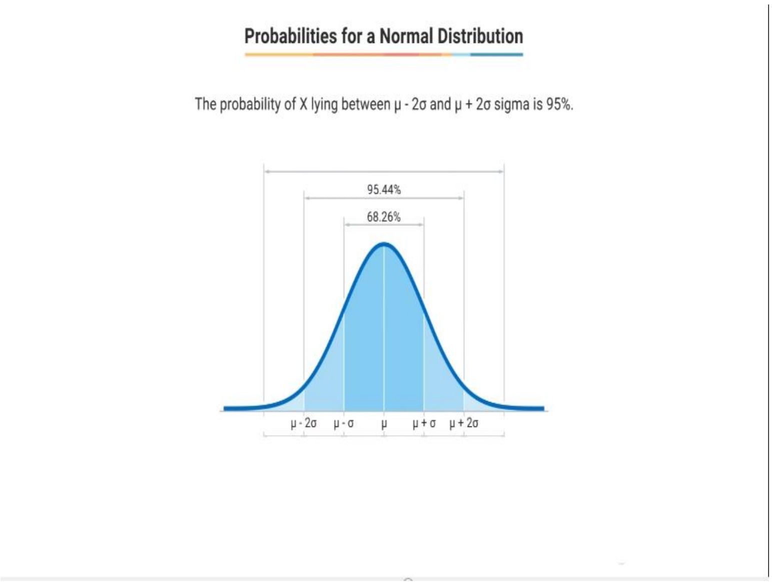
So, again we have a random variable X and we know that X follows a normal distribution with mean μ and standard deviation σ.

(Refer Slide Time: 11:55)



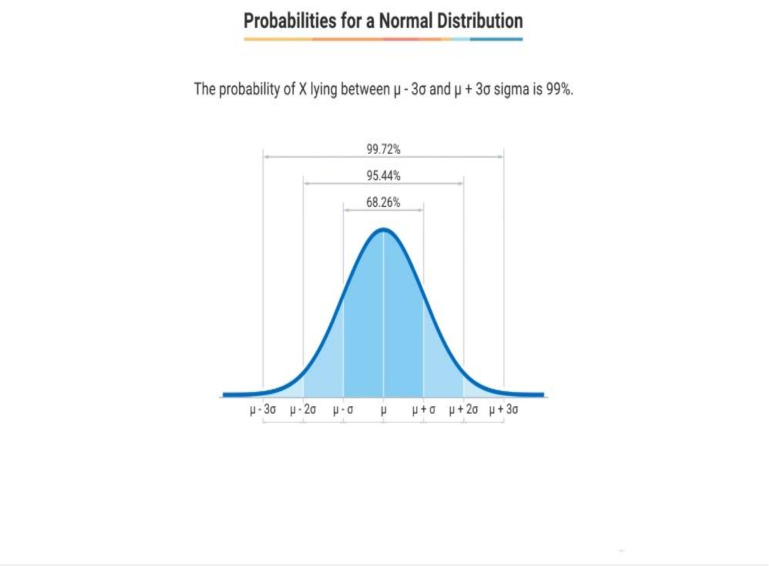
The empirical rule states that the probability of X lying between $\mu - \sigma$ and $\mu + \sigma$ is around 68 percent or 0.68.

(Refer Slide Time: 12:01)



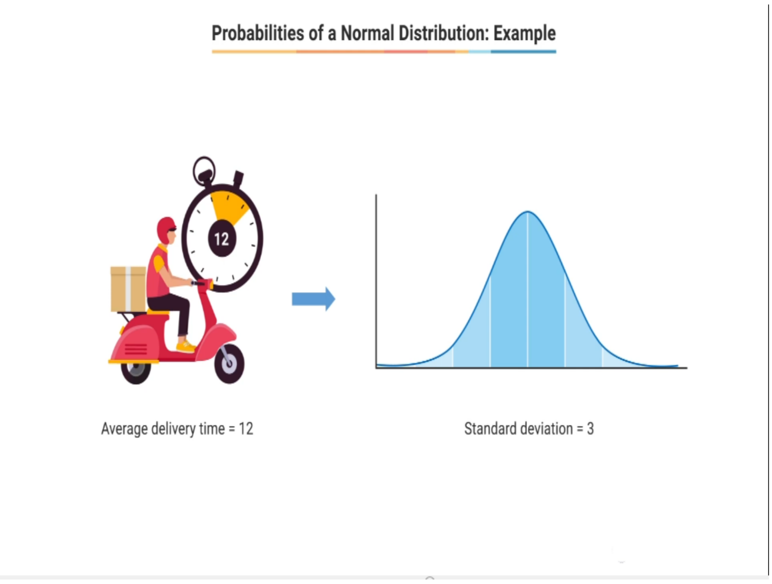
The probability of X lying between $\mu - 2\sigma$ and $\mu + 2\sigma$ is around 95 percent or 0.95.

(Refer Slide Time: 12:08)



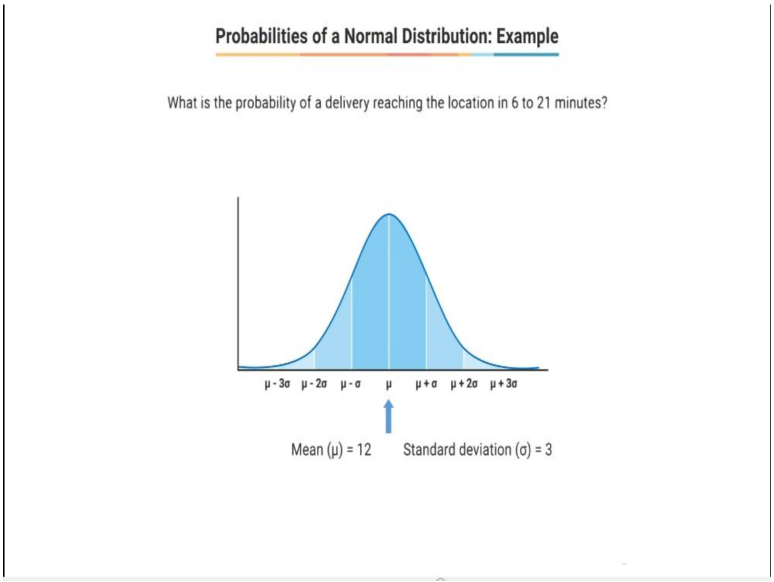
And the probability of X lying between $\mu - 3\sigma$ and $\mu + 3\sigma$ is around 99.7 percent or 0.997, this is a very important concept in a normal distribution. Let us go back to the pizza delivery example where we need to deliver pizzas to various customers who live in different parts of the city and hence the delivery boys have different commute times.

(Refer Slide Time: 12:29)



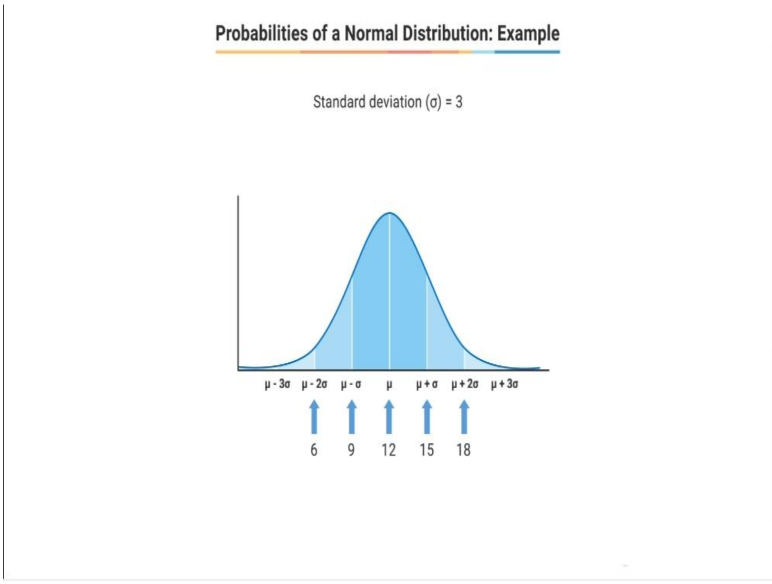
Now, the commute time of each of these deliveries were measured and was found that the commute time follows the normal distribution with the average time of 12 minutes and the standard deviation of 3 minutes. So, what information do we have about our distribution.

(Refer Slide Time: 12:43)



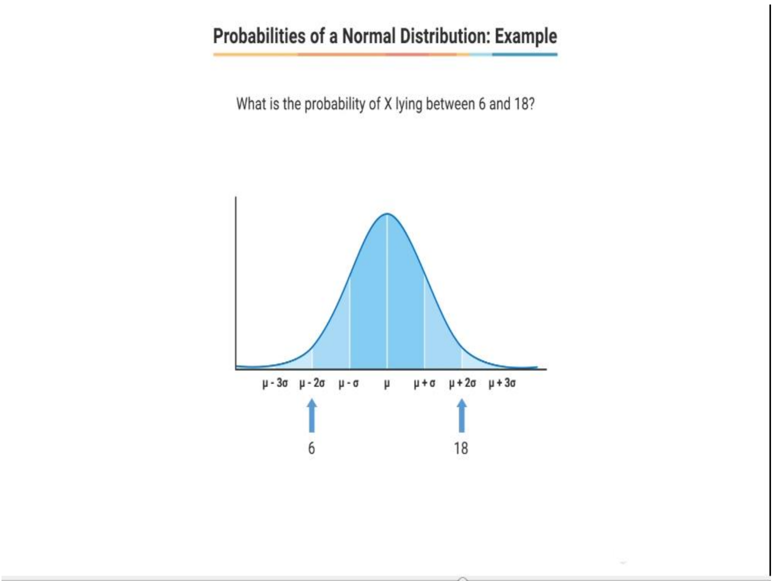
We know that the mean of the normal distribution is 12 minutes which is our μ and the standard deviation is 3 minutes which is our sigma. Suppose we want to know the probability of any delivery taking about 6 to 18 minutes to reach a customer location.

(Refer Slide Time: 12:56)



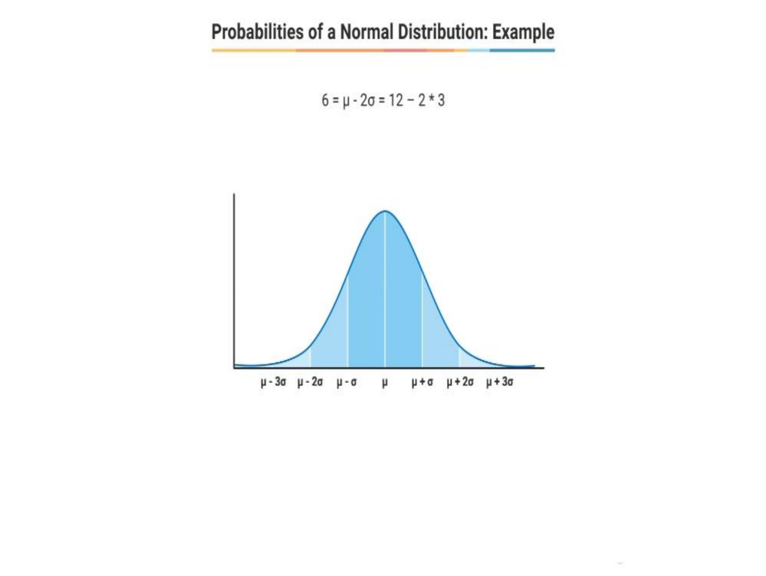
If we plot this on normal distribution, we know the value mu is given to us as 12 minutes and if we see on the left symmetry and right symmetry, we know that the standard deviation is 3 minutes. So, $\mu + \sigma$ =15 and $\mu + 2\sigma$ = 18 and so on. Similarly, on the left hand side $\mu - \sigma$= 9 and $\mu - 2\sigma =$ 6 again, because our μ is 12 and standard deviation is 3.

(Refer Slide Time: 13:24)



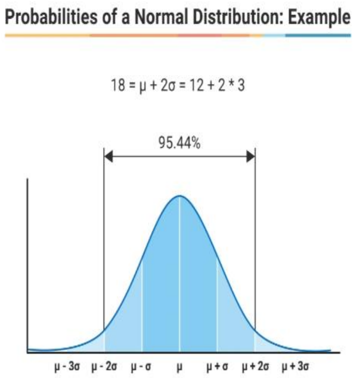
Now, as we have our 1, 2, 3 rule we want to know the probability of the commute time x line between 6 and 18 minutes.

(Refer Slide Time: 13:31)



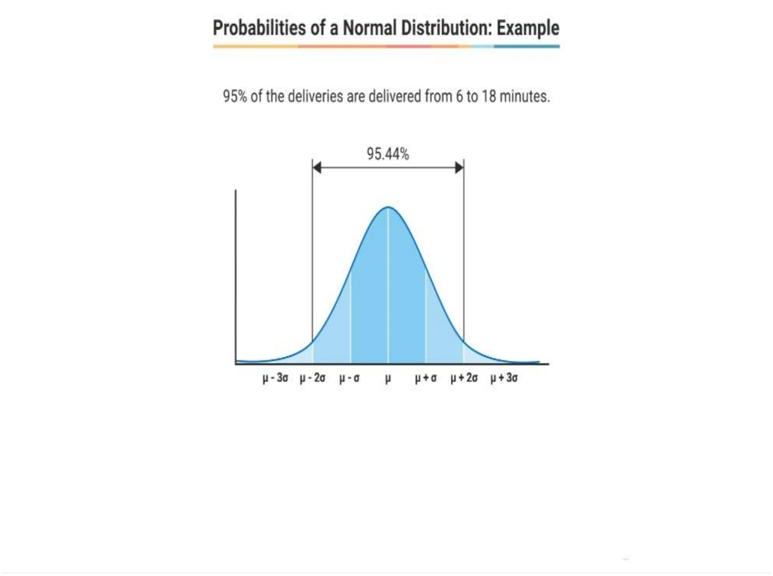
Since we know μ and σ, we can write $6 = \mu - 2\sigma = 12 - 2 * 3$.

(Refer Slide Time: 13:36)



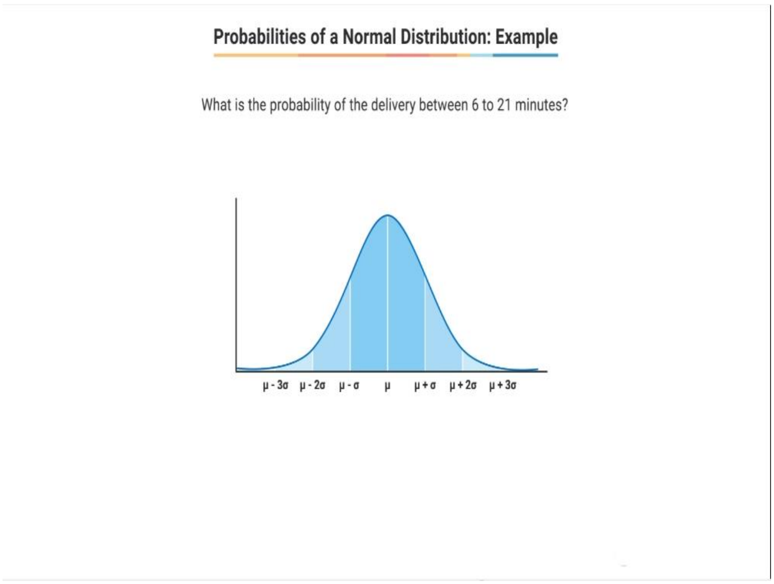
And we can write $12 = \mu + 2\sigma = 12 + 2 * 3$ which and using the empirical rule we know that 95 percent of the population lies between $\mu - 2\sigma$ and $\mu + 2\sigma$.

(Refer Slide Time: 13:48)



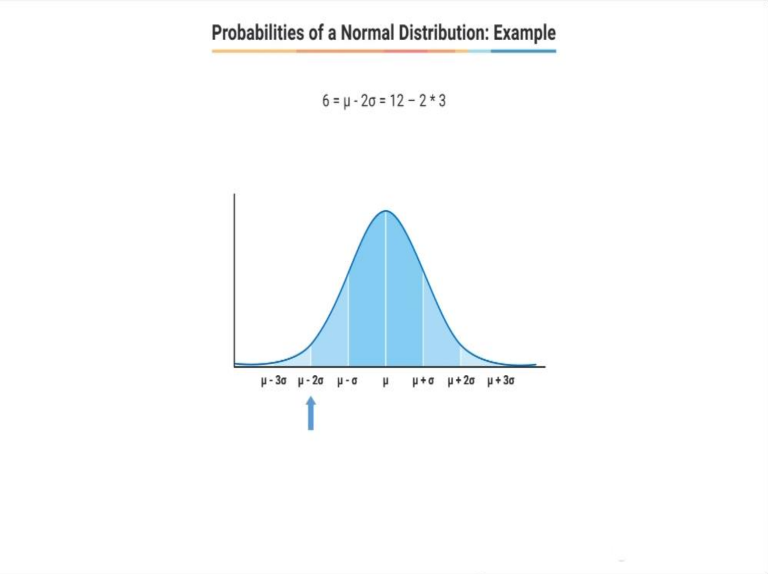
Or in other words 95 percent of the deliveries are actually taking 6 to 18 minutes to reach the customer location.

(Refer Slide Time: 13:53)



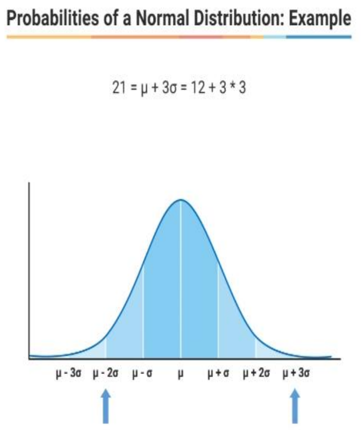
Now, let us say we want to know the probability that the delivery time is between 6 and 21 minutes.

(Refer Slide Time: 14:00)



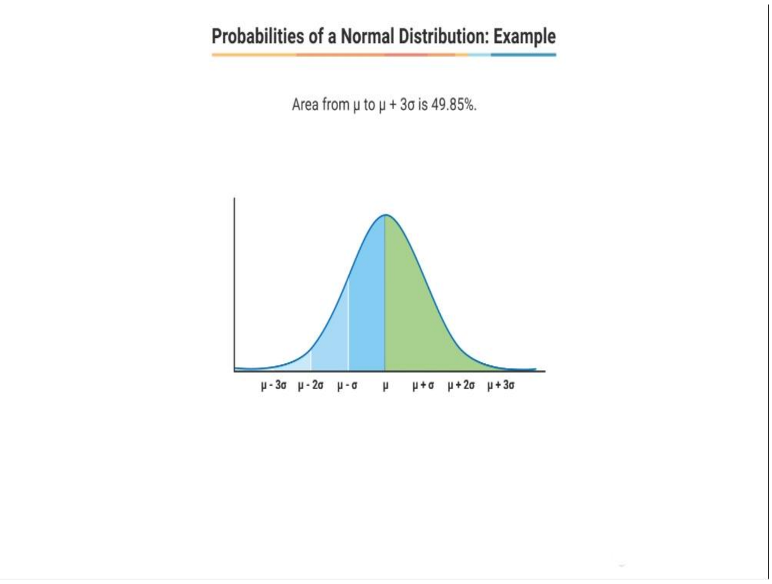
So, number6 $= \mu - 2\sigma$ because mu is 12 and sigma is 3.

(Refer Slide Time: 14:05)



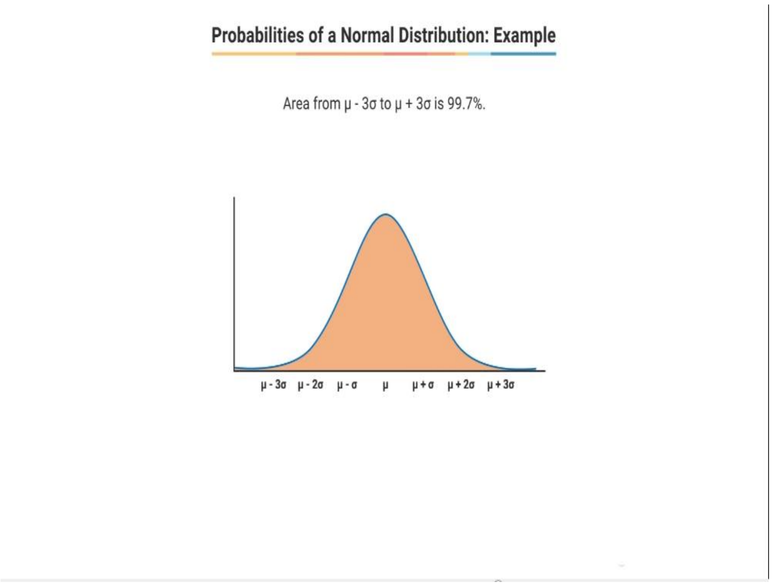
And now we can write $21 = \mu + 3\sigma$.

(Refer Slide Time: 14:08)



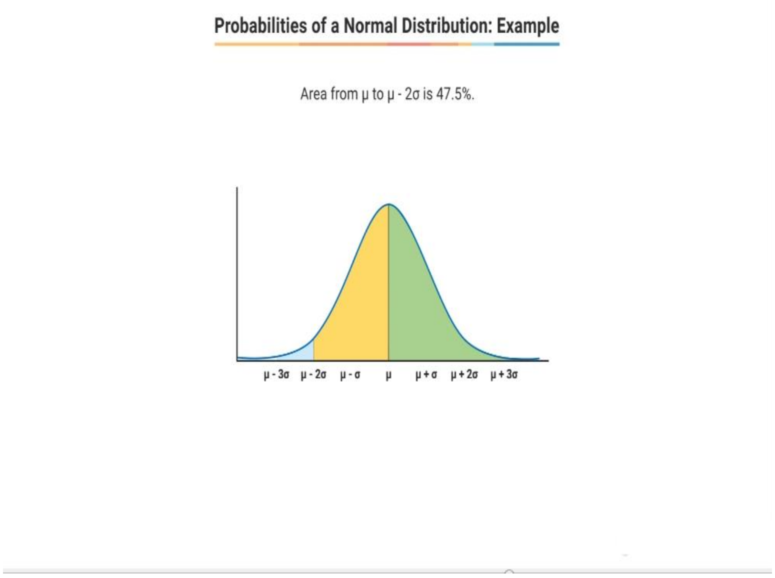
Similarly, if you see the right symmetry of the normal distribution plot you can see that the area under μ to $\mu + 3\sigma$ is like 49.85 percent of the population, why 49.85 percent?

(Refer Slide Time: 14:21)



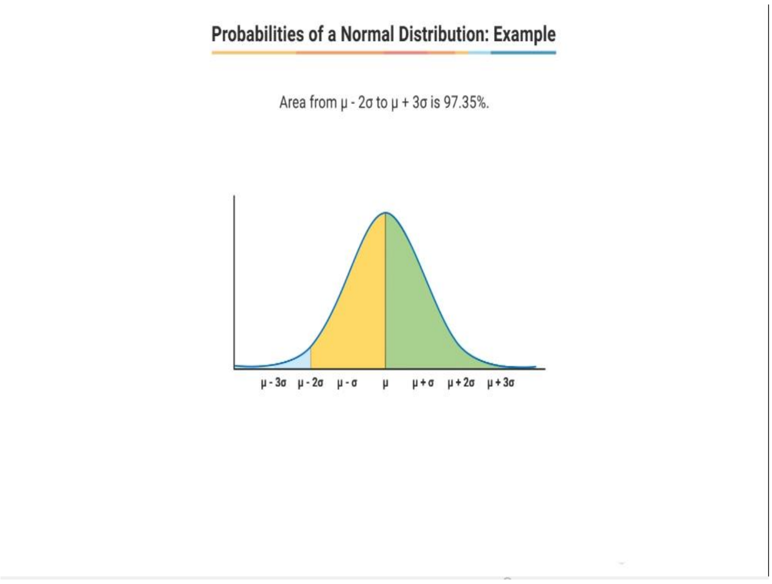
Because, we know already that the area from my $\mu - 3\sigma \; to \; \mu + 3\sigma$ is 99.7 percent and therefore we are just simply taking half of that. So, on the right side of the curve we will say that the area under the curve is 49.85 percent.

(Refer Slide Time: 14:34)



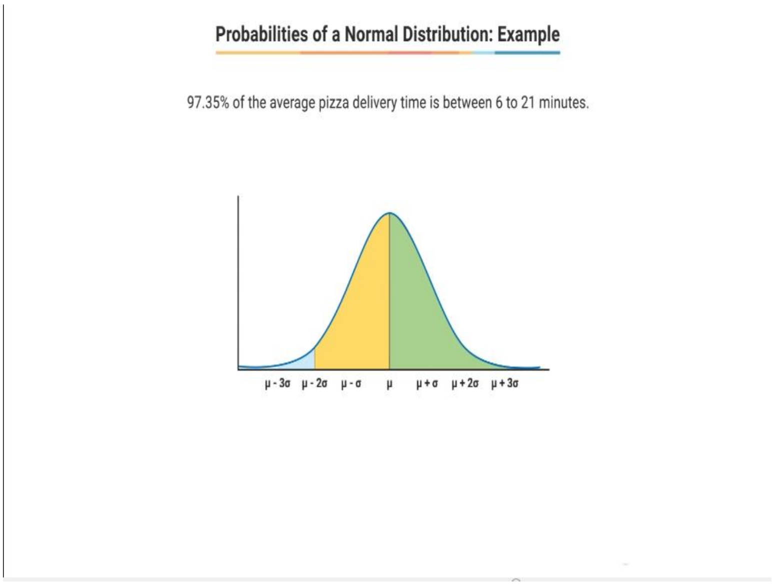
And on the left side of the curve, we see that the area will be half of 95 percent which is 47.5 percent of the population.

(Refer Slide Time: 14:40)



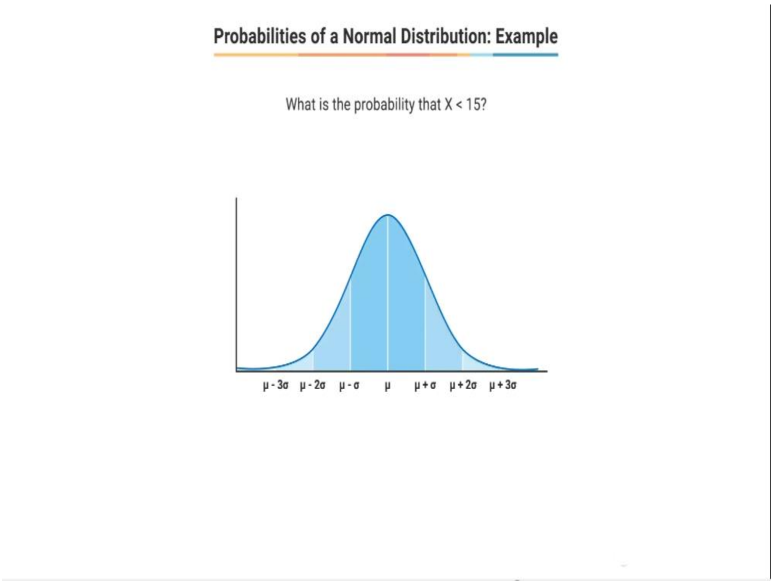
So, upon adding these two we get 97.35 percent of people.

(Refer Slide Time: 14:45)



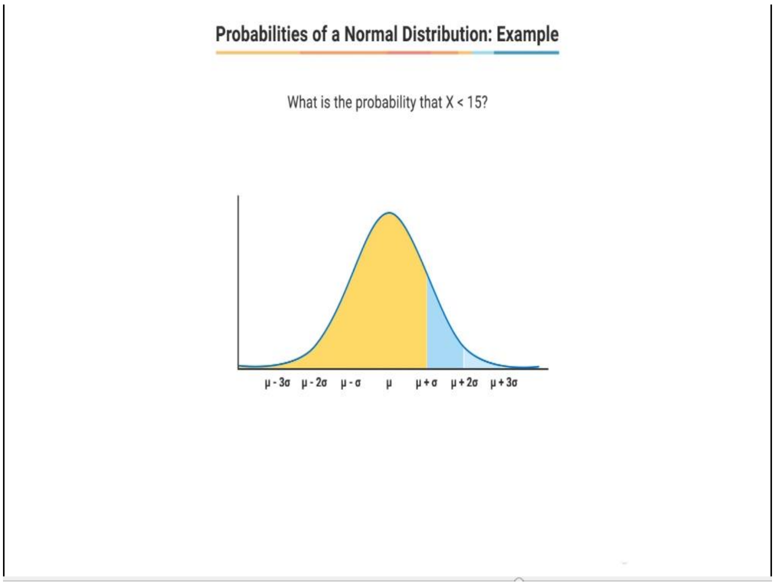
This means that 97.35 percent of the pizza delivery actually take 6 to 21 minutes as commute time to reach the customer locations.

(Refer Slide Time: 14:53)



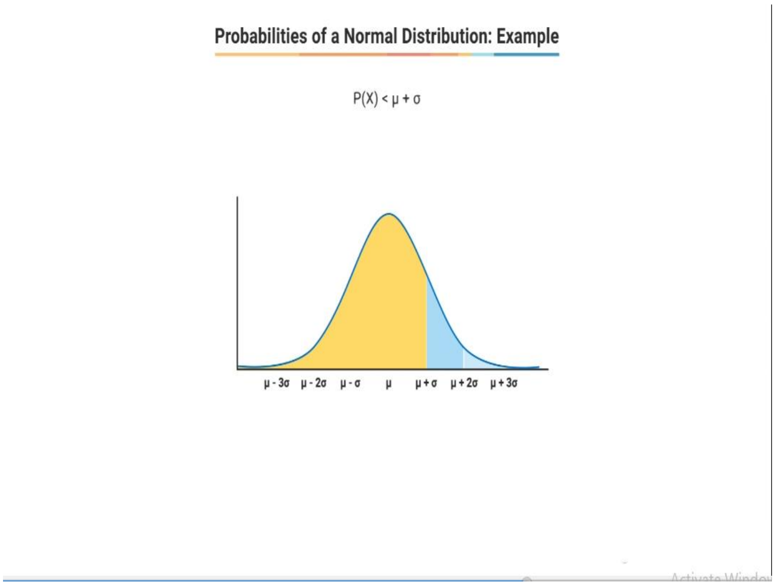
On, similar lines suppose we ask what is probability that X is less than 15,

(Refer Slide Time: 14:56)



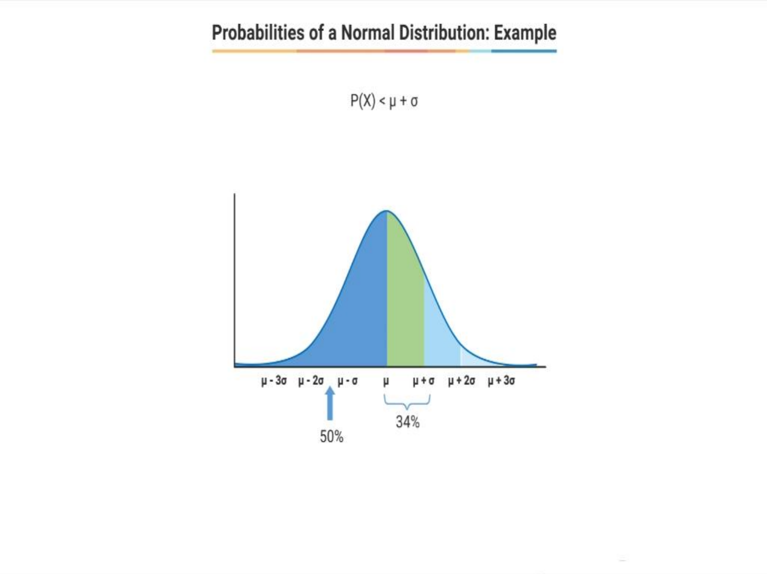
or in other words the percentage of pizza delivery is there the delivery boy is taking less than 15 minutes to commute to the customer location. Then it will be probability of X less than $\mu + 1\,\sigma$.

(Refer Slide Time: 15:04)



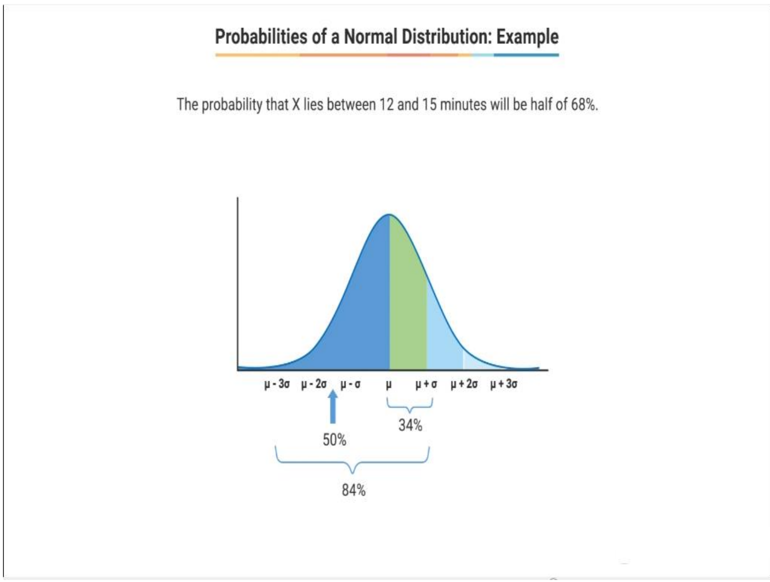
Now, let us try to calculate this value.

(Refer Slide Time: 15:09)



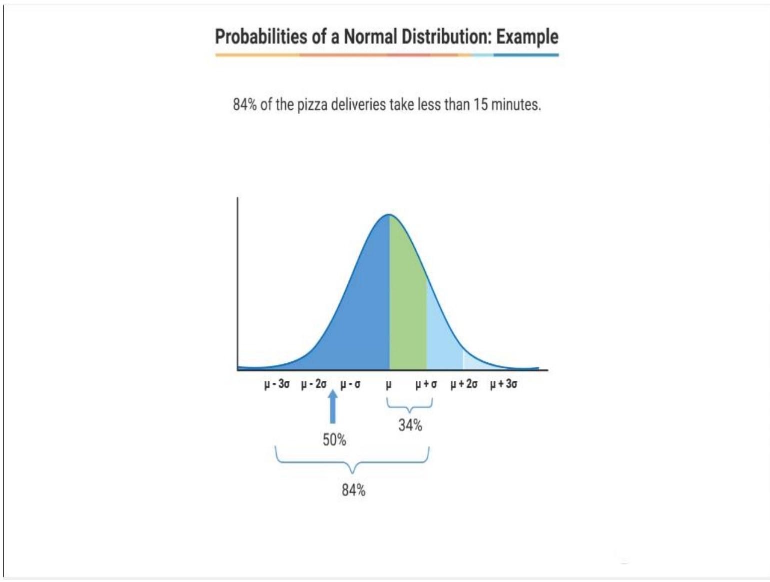
We can divide this area into 2 parts one below 12 minutes which is going to be 50 percent because 12 is the mean of the distribution and it is on the centre, and the second region will be the probability that X lies from μ to $\mu + 1\,\sigma$ and this value will be half of X lying within one standard deviation of the mean which was 68 percent.

(Refer Slide Time: 15:26)



Hence, the probability that X lies between 12 and 15 minutes will be half of 68 percent which is 34 percent. Now, when we add these two areas we get 50% + 34% which comes out to 84 percent.

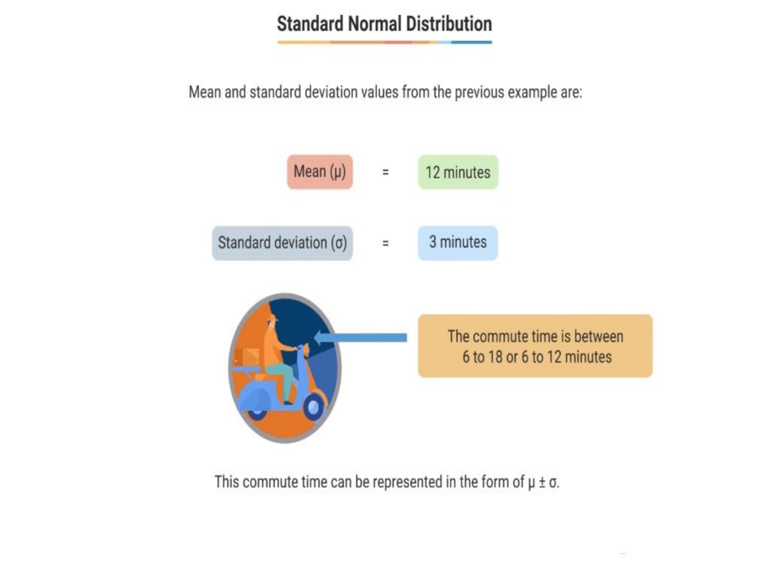
(Refer Slide Time: 15:40)



Thus, we can say that 84 percent of the pizza deliveries are actually taking less than 15 minutes of commute time.

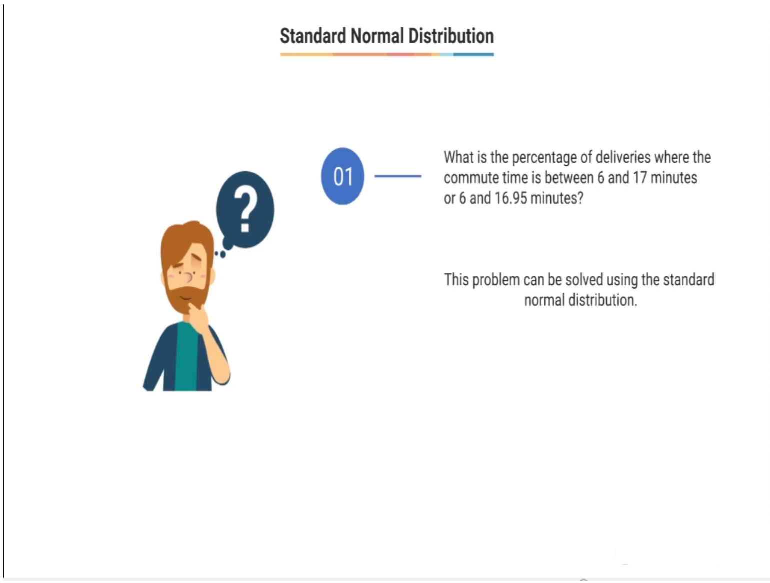
Standard normal distribution.

(Refer Slide Time: 15:47)



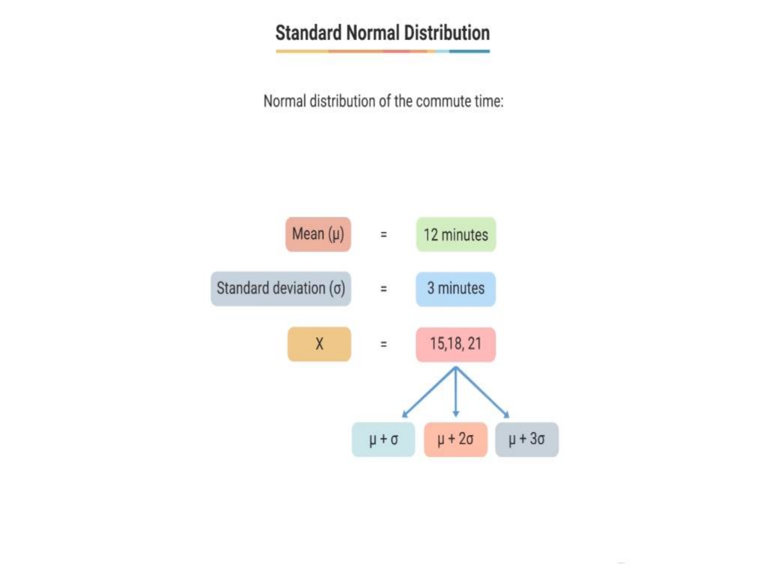
In previous example we had seen that our mean commute time was 12 minutes and our standard deviation was 3 minutes and we tried to find the percentage of pizza deliveries where the commute time is between 6 and 18 minutes or 6 to 12 minutes. In those cases, we were trying to find the commute time which was exactly falling within $\mu +$ some fixed number of standard deviations and it could be 1 standard deviation or - 1 standard deviation.

(Refer Slide Time: 16:13)



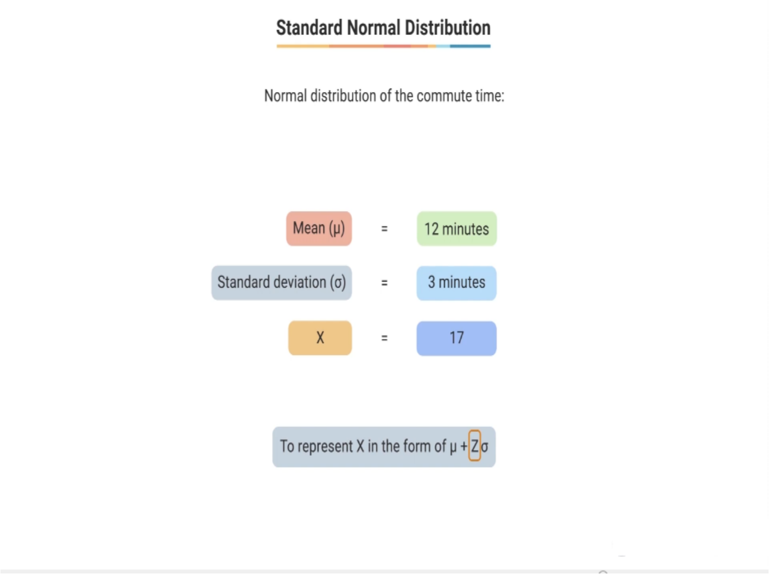
But suppose you want to find what is the percentage of deliveries where the commute time is between 6 and 17 minutes or between 6 and 16.95 minutes. In those cases, we will need some other techniques for example, the standard normal distribution.

(Refer Slide Time: 16:27)



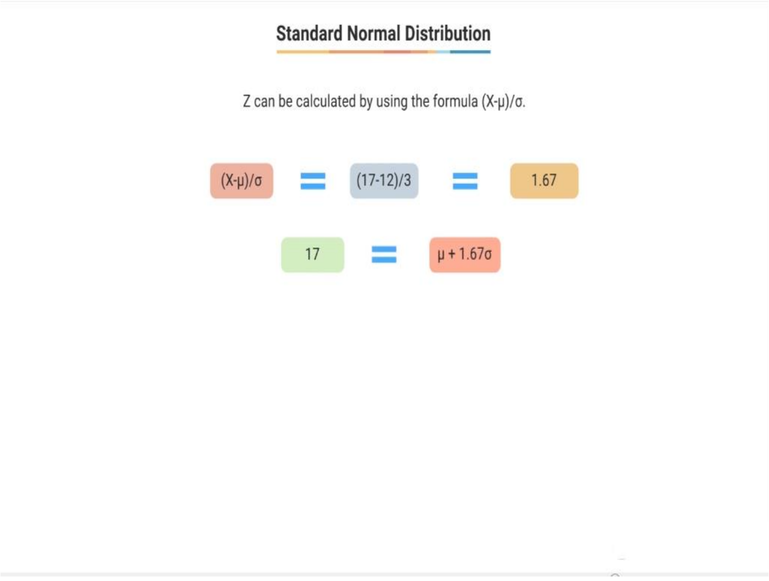
Let us go back to our normal distribution of the commute time. Here you can see that we have our mean as 12 when standard deviation is 3. Earlier we were considering X to be values like 15 18 and 21 which were multiples of 3. So, that we could represent them as mu + 1 sigma mu + 2 sigma and mu + 3 sigma respectively, but what if I consider value like 17?

(Refer Slide Time: 16:46)



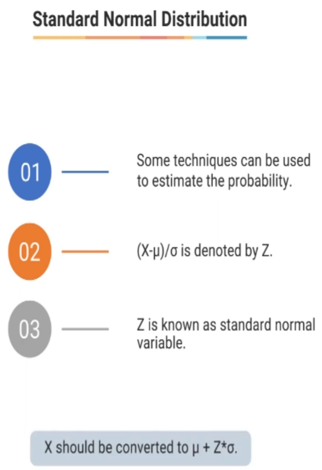
Here, we can see that μ is 12 and X is 17 and the difference between these two values should be 5. So, if we have to represent X in the form of $\mu +$ some other multiples of σ then we can find this multiplication factor by using $\frac{x-\mu}{\sigma}$.

(Refer Slide Time: 17:00)



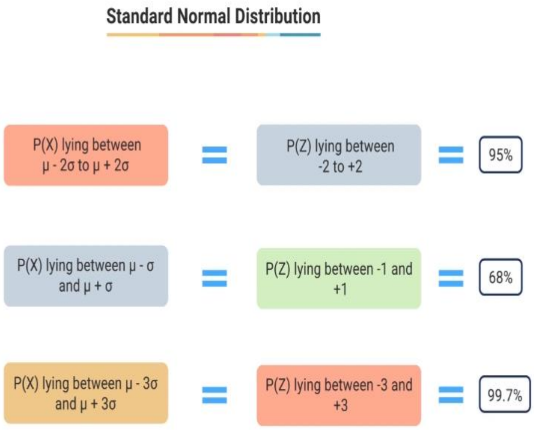
For instance, in our case $\frac{x-\mu}{\sigma} = \frac{17 - 12}{3}$ which equates to $\frac{5}{3}$ which = around 1.67. We can now represent 17 as $\mu + 1.67\sigma$.

(Refer Slide Time: 17:17)



And we can use some technique to estimate the probability. The value $\frac{x-\mu}{\sigma}$ is denoted by Z and Z is called the standard normal variable. So, essentially, we always convert are X into some form of $\mu + Z * \sigma$.

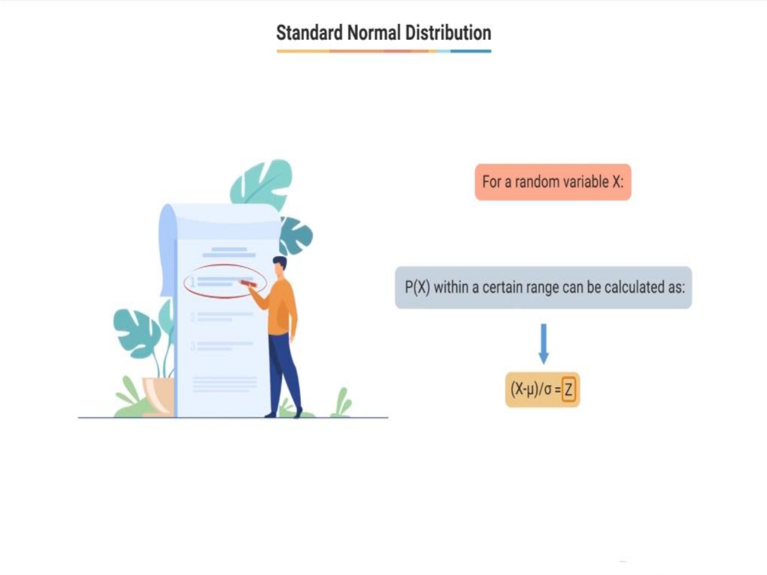
(Refer Slide Time: 17:30)



If you want to calculate the probability of X lying between $\mu - 2\sigma \ to \ \mu + 2\sigma$ it is the same as finding the probability that a new random variable Z which we mentioned earlier was called the standard normal variable and lies between - 2 to + 2. Thus, the probability of Z lying between - 2 to + 2 will be 95 percent as per the empirical rule. Similarly, the probability of X lying between $\mu - 1\sigma \ to \ \mu + 1\sigma$ is the same as the probability of Z line between - 1 and + 1 which is 68 percent.

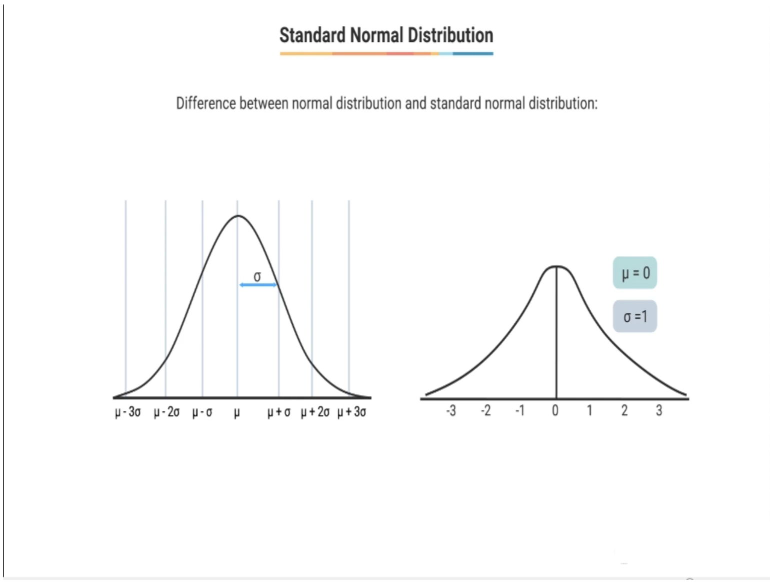
And finally, the probability of X lying between $\mu - 3\sigma$ *and* $\mu + 3\sigma$ is the same as the probability of Z lying between - 3 + 3 which is 99.7 percent.

(Refer Slide Time: 18:09)



Whenever we have some random variable X and we want to find the probability of X within a certain range we can calculate a new random variable by finding out $\frac{(X-\mu)}{\sigma}$ which is z and we can now say that Z will follow the standard normal distribution.

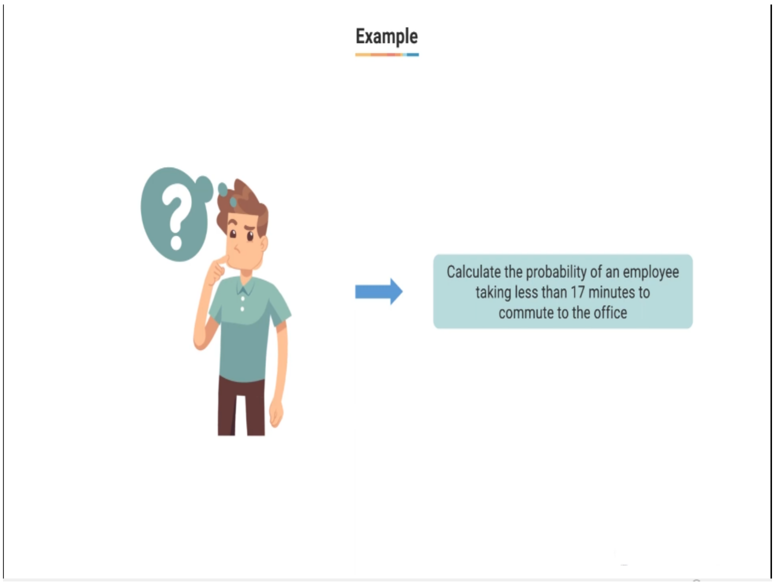
(Refer Slide Time: 18:24)



Let us see the difference between normal distribution and the standard normal distribution. If, you see the normal distribution you have a μ like in our case μ was 12 minutes and on the right side you can see that we have marked $\mu + 1\sigma$, $\mu + 2\sigma$, and $\mu + 3\sigma$. Similarly, on the left side you have $\mu - 1\sigma$, $\mu - 2\sigma$, $\mu - 3\sigma$ and this we have seen with the empirical rule in the previous discussion.

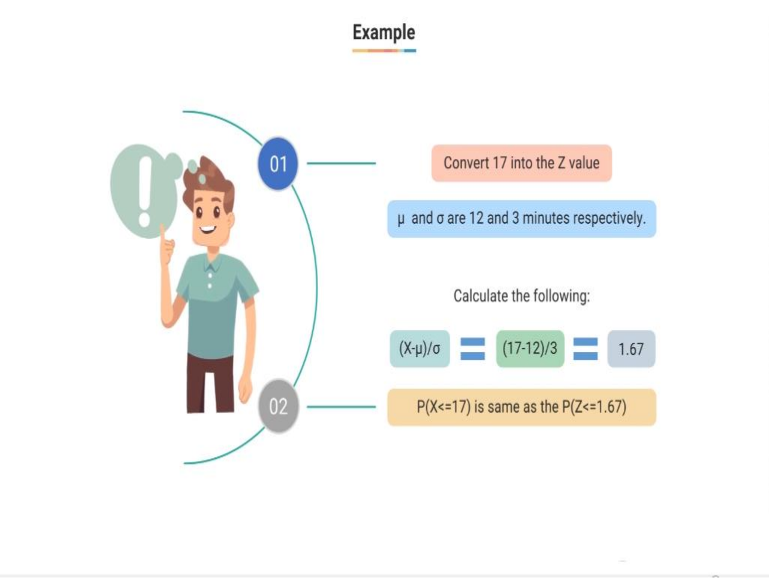
If you see the standard normal distribution then our mean is always 0 and the standard deviation is always 1, and $\mu + 1\sigma$ will be 1, $\mu + 2\sigma$ will be 2, and on the left side you will have $\mu - 1\sigma$ marked as minus 1 and so on. So, that is the difference between our normal distribution which we marked by X and the standard normal distribution between marked by Z. We will now look at an example problem on standard normal distribution.

(Refer Slide Time: 19:14)



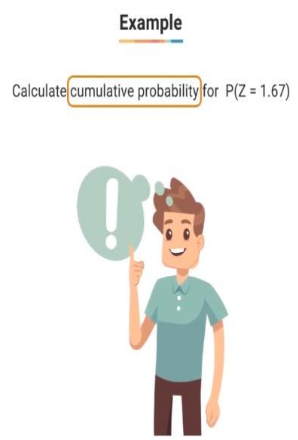
How do we proceed to solve the problem of calculating the probability that an employee will take less than 17 minutes to compute to the office. So, let us first convert this to the standard normal distribution.

(Refer Slide Time: 19:24)



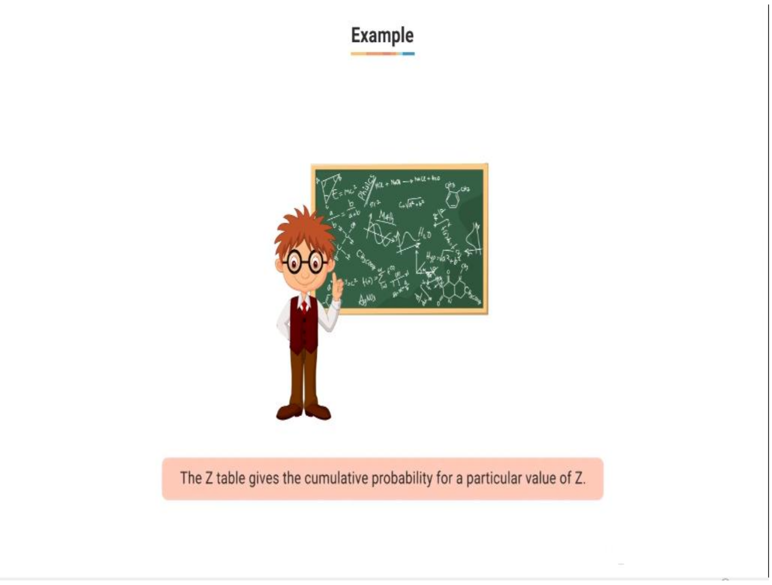
And then see how we can calculate the probability from here. First let us convert 17 into the Z value. Remember that the μ and \σ for this distribution are 12 and 3 minutes respectively thus calculating $\frac{(X-\mu)}{\sigma}$ we get $\frac{(17-12)}{3}$which equates $\frac{5}{3}$ which is 1.67 thus probability that X is less than 17 is the same as the probability t z less than 1.67.

(Refer Slide Time: 19:47)



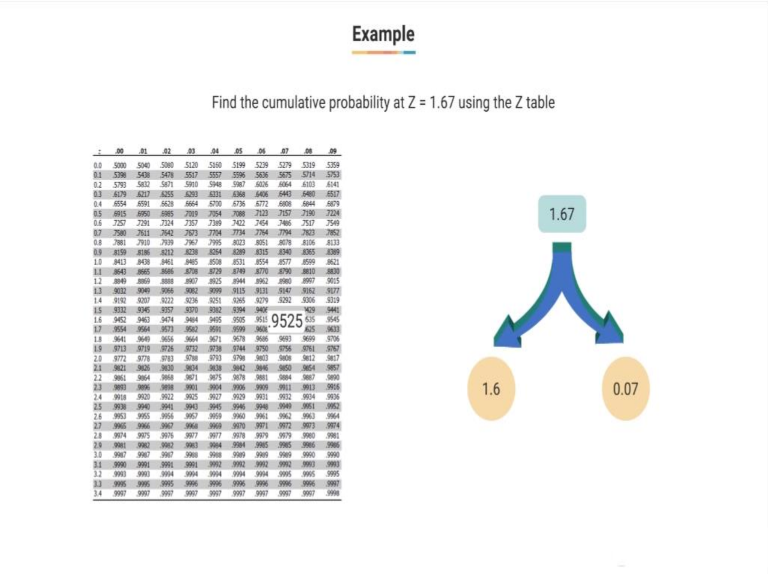
So, how can you find out the probability of Z = 1.67? If you observe carefully, we are essentially calculating the cumulative probability for Z = 1.67.

(Refer Slide Time: 19:58)



Fortunately, some mathematicians have calculated all the probabilities and put them in a table which is called Z table, this table gives the cumulative probability for a value of Z.

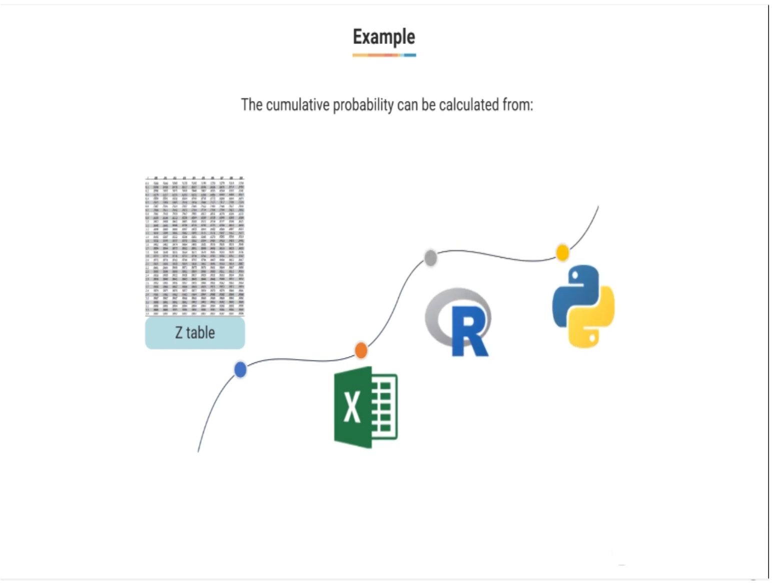
(Refer Slide Time: 20:08)



And let us see how you can find this table to find the cumulative probability at Z = 1.67 How doing that we will see the row based on before and after the decimal. So, this value 1.67 can be broken into 1.6 and 0.07. Next, we will search for the row corresponding to 1.6 as you can see that we are there and we can see the second decimal place in the column. So, we will select the column marked as 0.07.

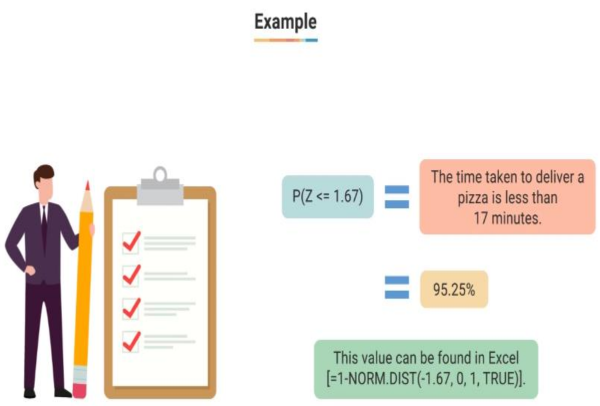
So, at the intersection of this row and column you can find the cumulative probability of Z being 1.67 which is 0.9525.

(Refer Slide Time: 20:40)



This kind of calculation can be easily done from Z table Excel and various other programs like R and python.

(Refer Slide Time: 20:46)



Thus, we now have the probability of Z less than 1.67 or the pizza delivery commute time which is less than 17 minutes which is approximately 95.25 percent. In Excel this can be found using the function [=1 – NORM.DIST (- 1.67,0,1, TRUE)] or we can also use that table for here.

(Video Starts: 21:09)

In this lesson we understood the concept of probabilistic events and sample space. We also examined the probabilities of intersection and union of events when 2 or more events are combined. We also discussed the computation of conditional probabilities with the help of Bayes' theorem. We examined the computation of joint and marginal probabilities. Next, we introduce random variables and probability distributions.

We also computed the expectations and variance of a random variable with a given probability distribution. Two key distributions which we focused in detail included binomial and normal probability distributions. For these distributions we examined key properties such as expected value standard deviation and cumulative probabilities. For more advanced discussions our focus remained on continuous random variables and normal distributions.

Given their integral role in applications such as central limit theorem, confidence interval estimation and hypothesis testing. We also examine a very important version of normal distribution that is standard normal distribution. These discussions will act as building blocks to more advanced applications in statistical inference and predictive Analytics. In the next lesson we will discuss descriptive statistics.

(Video Ends: 22:17)