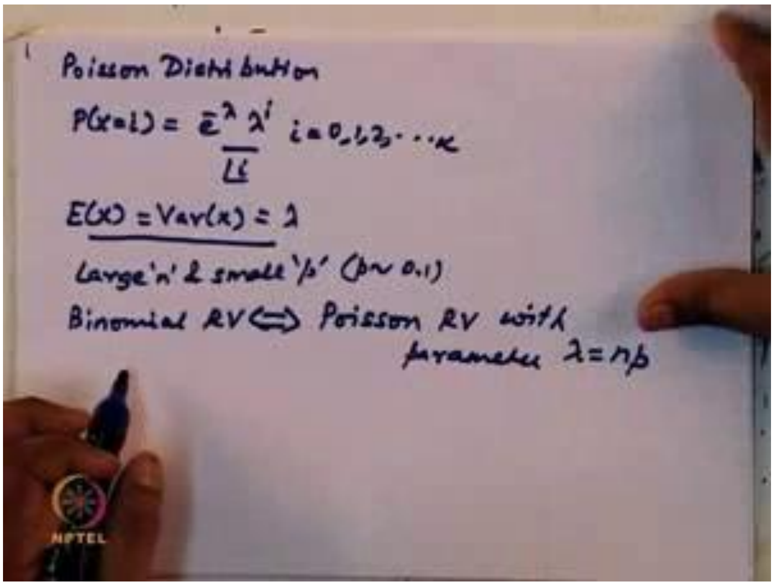
Introduction to Biostatistics
Prof. Shamik Sen
Department of Bioscience and Bioengineering
Indian Institute of Technology, Bombay

Lecture – 24
Uniform distribution Part-II and Normal distribution Part-I

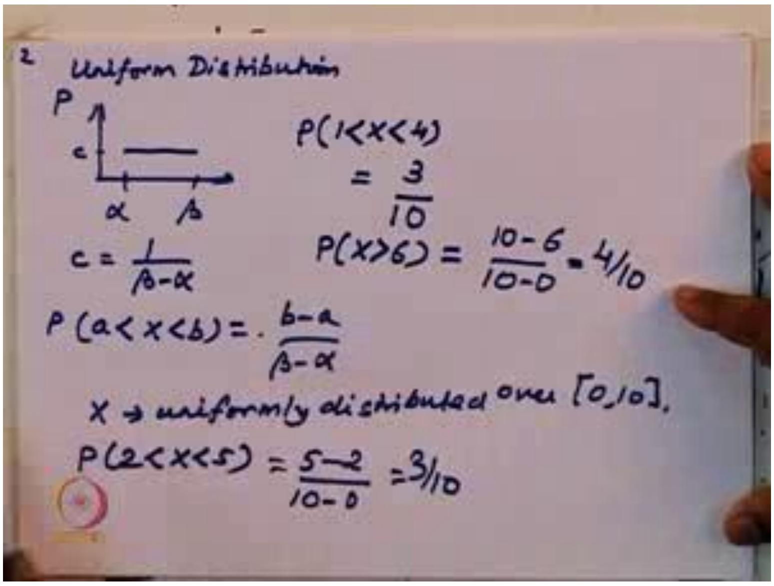
Hello and welcome to today's lecture. So, we would like to continue and discuss about 2 very important distributions to today, but before I start I would like to briefly review what was discussed in last class.

(Refer Slide Time: 00:35)



So, in last lecture we discussed about 2 distributions the Poisson distribution. So, for the Poisson distribution probability of x equal to i is given by e to the power minus lambda to the power i by factorial i, i is equal to 0, 1, 2 and it goes up to infinity. So, Poisson distribution the Poisson random variable is a discrete random variable, but it can go up to very high values and what we had demonstrated in last class was for the Poisson distribution your expectation and variance both return you a value of lambda. So, Poisson distribution also one more thing, so for large n and small p, so by small I mean let us say p is order 0.1. So, we can have the binomial random variable is almost equivalent to the Poisson random variable with parameter lambda is equal to n times.

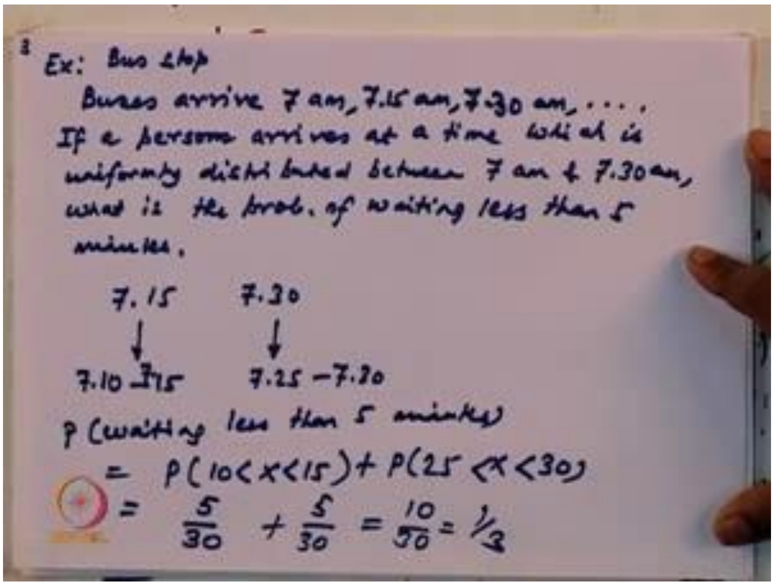
(Refer Slide Time: 01:58)



So, other than the Poisson distribution, the other distribution we discussed yesterday was the uniform distribution. And for a uniform distribution in a range alpha to beta your so your probability is uniform, it is a constant and this constant. So, this c is nothing, but c is given by 1 by beta minus alpha. So, probability of a less than x less than b is nothing, but b minus a by beta minus alpha; so let us discuss some few very simple cases. So, imagine you have a random variable x which is uniformly distributed, over 0 to 10. So, let us say you want to calculate probability of 2 less than x less than 5, what is this probability. So, this is nothing as you know from this example your b is equal to 5, a is equal to 2, so this probability is nothing, but 5 minus 2 and beta and alpha what is beta is 10 alpha is 0. So, 10 minus 0 is equal to three-tenth.

Similarly, probability of let us say 1 less than x less than 4 is nothing will return your same value. 2 less than x less than 5 and 1 less than x less than 4 both will return you a value of 3 by 10. What is probability of x greater than 6? X greater than 6 is nothing, but since you can go from 6 to 10; so x greater than 6 would be 10 minus 6 by 10 minus 0, is equal to 4 by 10. So, you see how you can make use of the uniform distribution to calculate each of the probabilities.

(Refer Slide Time: 03:56)

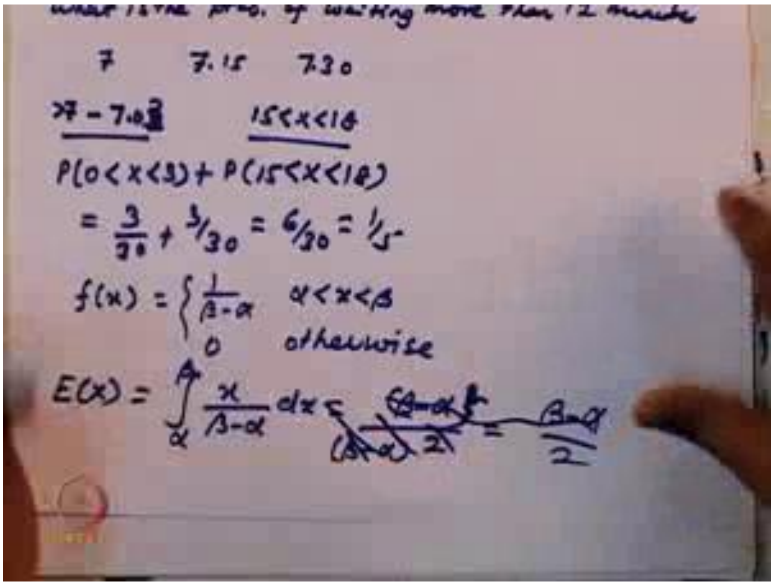


Let us take one more example. So, imagine at the bus stop, your buses arrive every 50 minutes. So, they arrive at 7 am, 7.15 am and so on and so forth. So, 7.30 am and so on and so forth. So, the question is: If a person arrives at a time which is uniformly distributed between 7 am and 7.30 am, what is the probability of waiting less than 5 minutes. So, let us read the problem. So, the buses arrive at every 15 minutes, so 7, 7.15, 7.30 so on and so forth. So, the person arrives at a time which is a random variable and which is uniformly distributed between 7 and 7.30, which means that person has equal chance of arriving at 7 7.01 so on and 7 .1 up to 7.30. So, what is the probability of waiting less than 15 minutes less than 5 minutes?

So, since your buses arrive at 7, 7.15 then 7.30 So, if he has to wait less than 5 minutes he can arrive anytime between 10 to 15 or 7 10 to 7.15 or anytime between 7.20 5 to 7 30. So, what is my probability? So, this probability waiting probability of waiting less than 5 minutes is probability of 10 less than x less than 15, plus the probability of 25 less than x less than 30. This is what this is 5 divided by 30 why because you have the entire duration is 30 minutes, and this is also 5 divided by 30. So, total is 10 by 30 equal to one-third.

Now, we can also ask the question what is the probability of waiting more than 12 minutes. So, the bus arrives at 7, 7.15, and 7.30.

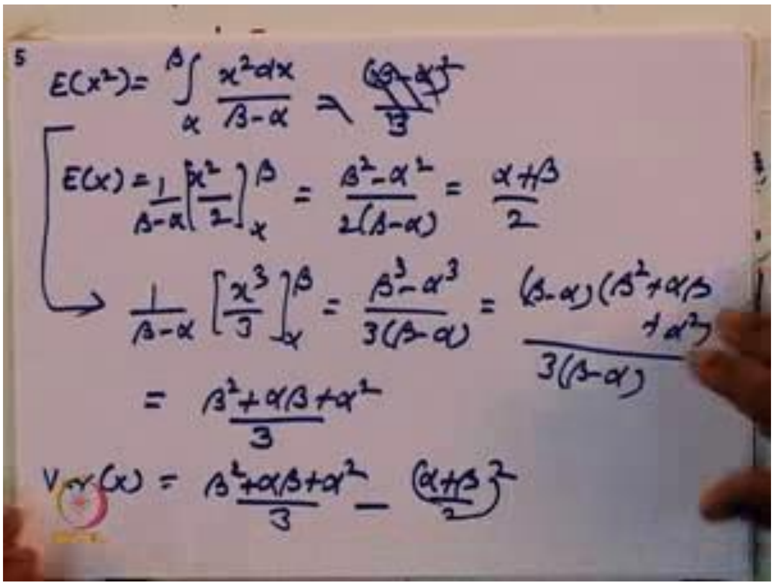
(Refer Slide Time: 07:03)



So, if the person has to wait more than12 minutes, then he has to arrive greater than 7 to 7. 03 right because if he arrives at 4 then he has to wait only for 11 minutes, since he has to wait at least more actually waiting at least. So, this if it is more than 12 minute then you have 7 to 7.02 or if also 7.03 also (Refer Time: 07:57) and similarly he can any anytime between 15 right after 15 to 18.

See if he arrives just after 15 then he would have to wait for 30 minutes, which is more than 12 minutes and latest he can arrive at 7.18 in that case you will have to wait for exactly 12 minutes. So, my probability is nothing, but probability of 0 less than x less than 3, plus probability of 15 less than x less than 18 equal to 3 by 30, plus 3 by 30 equal to 6 by 30 nothing but one-fifth. So, for a uniform distribution once again I right this is one by beta minus alpha for alpha less than x less than beta and 0 otherwise; you can calculate the expectation, expectation of this variable is nothing, but x this will return you a value of.

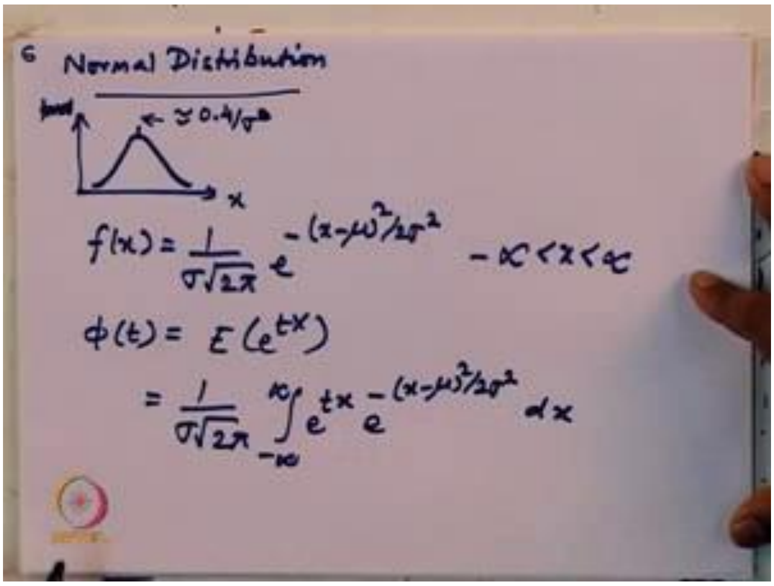
(Refer Slide Time: 09:23)



Expectation would return in a value of beta minus alpha by 2, similarly I can calculate expectation of x square. So, x cubed by 3 and x cubed is this. So, you will get a value of sorry sorry I think I have made a mistake. So, let me redo this. So, expectation of x is x square. So, 1 by beta minus alpha x square by 2 between alpha to beta, equal to beta square minus alpha square by 2 into beta minus alpha is equal to alpha plus beta by 2, that makes sense on an average you will get a value which is average of alpha and beta. This also I will redo. So, expectation of x square is going to be 1 minus beta minus alpha into x cube by 3, is beta cubed minus alpha cubed beta square plus alpha beta plus alpha square by 3 into beta minus alpha. So, this would give you a value of beta square plus alpha beta, plus Alpha Square into by 3.

So, your variance of x then becomes Alpha Square by 3 minus alpha plus beta by 2 whole square; you can simplify this and see what value you get. So, that brings our discussion of uniform distribution to a close, we will now discuss one of the most important distributions which is the normal distribution.

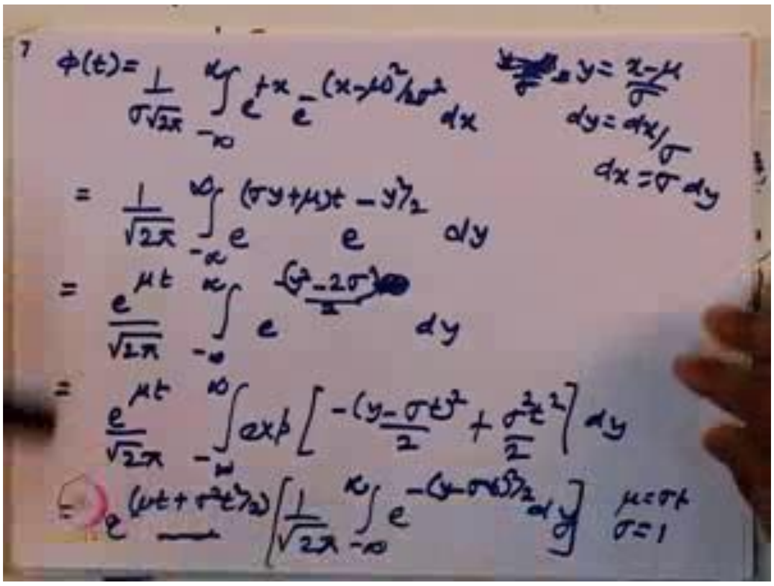
(Refer Slide Time: 11:33)



So, what is the normal distribution? So, normal distribution is the mount shaped distributions that you see at all times. This is your x value this is your probability. So, for normal distribution your f of x is given by. So, you see there are 2 parameters here. So, your probability density function is given by f x, and 1 by sigma root 2 pi to the power minus x minus mu whole square by 2 sigma square. So, typically this peak is roughly 0.4 by sigma square, this peak is roughly sorry 0.4 by sigma. So, we want to know what is this value of mu and sigma respectively.

So, for that as again we can use the moment generating function. So, for this is equal to expectation of e to the power t x, this will be one I can take out this is a constant. So phi of t then becomes I can let us say I put this mu minus y minus mu by sigma is equal to sorry.

(Refer Slide Time: 13:11)

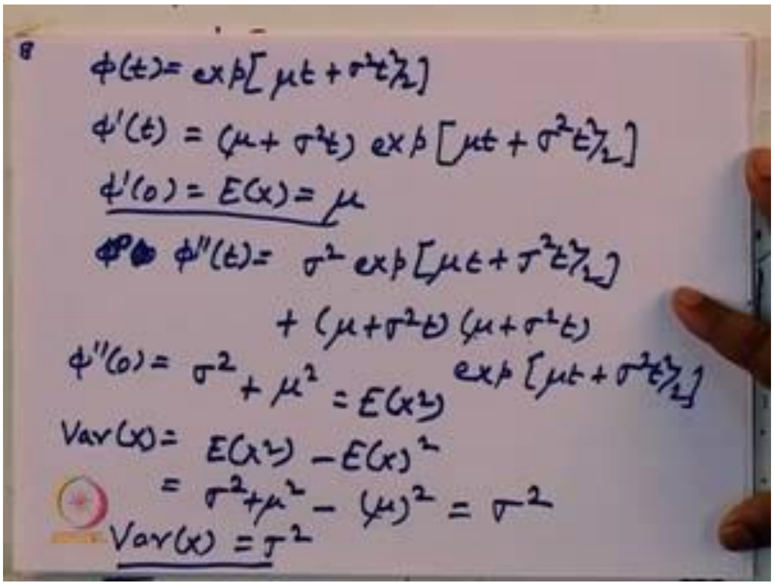


So, let me define y as x minus mu by sigma. So, phi of t becomes sigma root 2 pi. So, if I put this then my d y is equal to d x by sigma. So, I can actually convert it. So, I have minus 2 e to the power t x, e to the power minus x minus mu whole square 2 sigma square into d x right. So, in this would mean my d x is equal to sigma into d y. So, limits would not change, but if I put this values again. So, d x by sigma is equal to d y. So, this sigma goes the limits remain unchanged, power sigma y plus mu y square by 2 into d y.

So, I can take into this mu t term out, because the integral is over y and I can write it as minus infinity to infinity e to the power minus y square, minus 2 sigma, plus y by 2 into d y right. So, we have minus y square 2, minus y square by 2, minus 2 sigma. So, this you can. So minus 2 sigma is what you get. So, this is plus. So, this I can write as e to the power mu t by root 2 pi I can write minus infinity to infinity exponential. So, within the exponential I can write it as y minus sigma t whole square by 2, plus sigma square t square pi 2 into d y.

So, I can take out this term again. So, I can write it as e to the power mu t plus sigma square t square by 2 by root 2 pi can be separate. So, this is exponential together into1 by root 2 pi minus infinity to infinity, e to the power minus y minus sigma t whole square by 2 d y. So, this term this term is nothing, but the same distribution f of x with parameter of mu is equal to sigma t and sigma is equal to 1. So, this integral gives you nothing but 1.

(Refer Slide Time: 16:30)

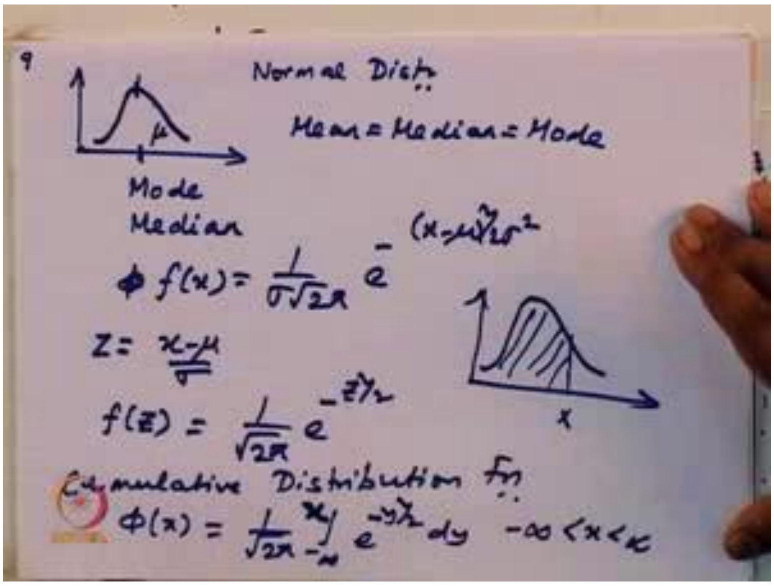


So, I can then write the entire phi of t is nothing but exponential mu t plus sigma square t square by 2. So, phi prime t then becomes simply mu plus (Refer Time: 16:50) sigma square t into exponential of the same thing.

So, phi prime 0 is equal to expectation of x, which will be giving me nothing, but mu if you put t equal to 0 here this will return you a value of 0, this will return you a value 0. So, exponential of 0 is 1, this will return you a value of 0. So, only you get this mu. So, expectation of x is equal to mu phi double prime of 0. So, first let us find out phi double prime of t is equal to phi double prime of t is equal to sigma square exponential of mu t plus sigma square t square by 2. So, basically I have taken a derivative of this term keeping this as constant, and plus into mu plus sigma square t exponential of mu t plus sigma square t square by 2.

So, this will return you a. So, phi double prime of 0 will give me a value of sigma square plus. So, t equal to 0 you only get a value of mu, here also you get a value of mu, and this whole thing is 1. So, you get a value of sigma square plus mu square. So, this is nothing, but E of x square. So, variance of x so this we have already opted is nothing, but E of x square, minus E x whole square is equal to sigma square plus mu square, sigma square plus mu square minus mu whole square nothing, but sigma square. So, you get variance of x is nothing but sigma.

(Refer Slide Time: 18:55)

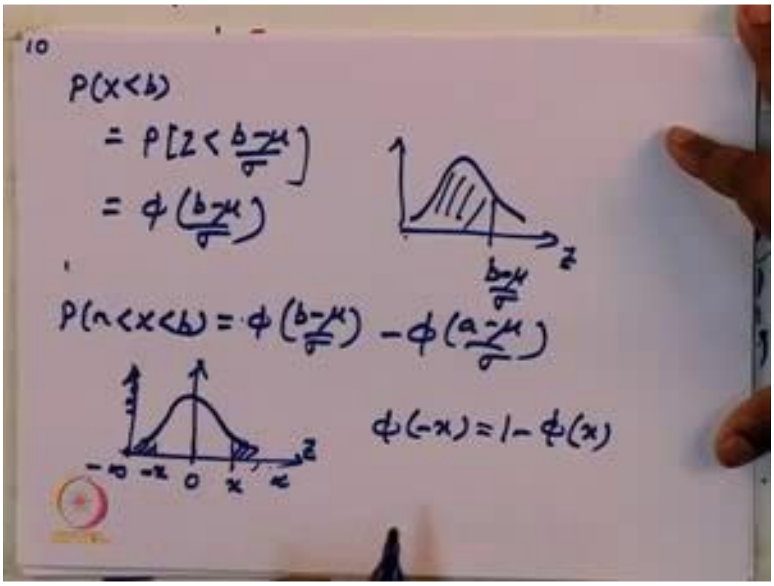


So, for a normal distribution the mean is equal to mu, and the variance is equal to sigma square; and what you can also observe that from this distribution since this is symmetric about the center. So, this is also the mode y because this is the maximum occurring value, and this is also the median because this is right because this is the symmetric distribution, and for every point to the left there is another point to the right which is occurring at identical frequency. So, their average will always give you this particular value which is nothing. So, this is nothing but mu. So, it means that for a normal distribution

So, mean is equal to median is equal to mode; now for this if I know that f of x is defined as 1 by sigma root 2 pi, e to the power minus x minus mu whole square by 2 sigma square right. So, I can define this variable z as x minus mu by sigma then f x. So, this f x I can then find out a cumulative distribution function. So, this becomes f of z becomes 1 by root 2 pi e to the power minus z square by 2, and the cumulative distribution function of phi of x, phi of x then becomes 1 by root 2 pi into minus infinity to infinity e to the power minus y square by 2 d y, minus infinity to x minus infinity less than x less than infinity.

So, this is your value. So, what is y phi of x for this distribution? For any value x phi of x is this area under the curve this is your total probability of finding any x in any value less than x. So, we can write down for the normal distribution. So, for probability of x less than b is nothing, but probability of z less than b minus mu by sigma is equal to phi of b minus mu by sigma.

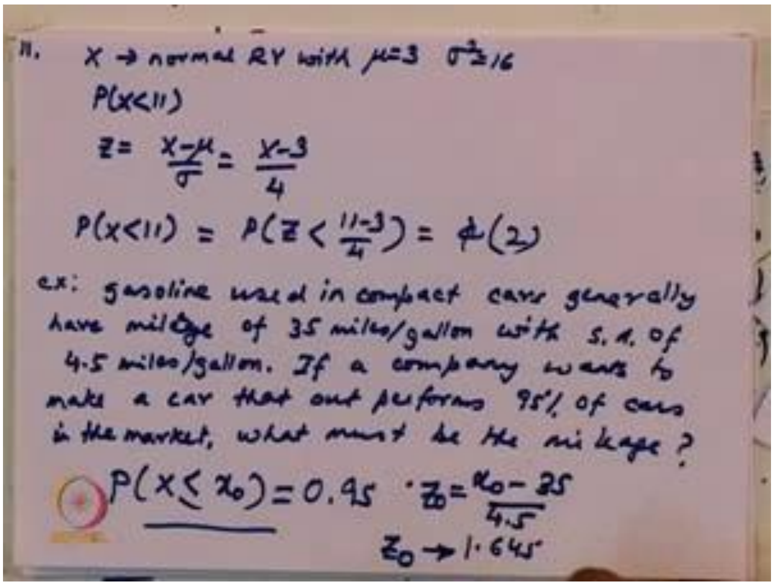
(Refer Slide Time: 21:15)



So, this is the value. So, these values these values of phi are all been computed and you can find them from tables, existing tables in any statistics book. So, I can write down. So, if this is probability of x less than b as I drew earlier. So, for any b this is my probability of x less than b. So, I can find out for the z variable, what is the correspondence. So, this correspondence to b minus mu by sigma and accordingly I can look this up value up, I can write probability of a less than x less than b is nothing, but phi of b minus mu by sigma minus phi of a minus mu by sigma.

So, one more thing because of symmetry if this is your distribution. So, this is for the entire x for phi actually you will have mean is always equal to 0 right; so for phi mean is always equal to 0. So, you have minus infinity to infinity. So, for any value x this area will also be equal to this area, in other words phi of minus x; so let us say this is value of x and this is minus x. So, this is your z axis right. So, phi of minus x you know this area and this area are exactly small. So, I can write phi of minus x is equal to 1 minus phi of x. So, this whole area is unity, entire integration will return you a value of unity and this area is nothing but whole area minus this area and this come comes because of symmetry. So, accordingly by looking up these phi values I can find out any particular value where x has a value less than equal to 1 less than equal to infinity.

(Refer Slide Time: 23:33)



So, let us just take one sample example. So, if x is a normal random variable with mu is 3, and sigma square equal to 16 we want to calculate probability of x less than 11. So, how do I calculate; I just define z is equal to x minus mu by sigma in this case it is x minus 3 by 4. So, p of x less than 11 is nothing, but probability of z less than 11 minus 3 by 4 is equal to phi of 2. So, you can find out this is by looking up this value what is the value of probability of x less than by looking up the table.

Let us take another example, let us say you have gasoline used in compact cars, generally have mileage of 30 5 miles per gallon with standard deviation of 4.5 miles per gallon. So, if a company wants to make a car that out performance 95 percent of cars in the market what must be the mileage? So, essentially we have been asked to ask for what value of x naught;' for what value of x naught do we get probability equal to 0.95. So, what do we do? We define z naught as x naught minus 35 by 4.5 and find out. So, we want to find out for what value of z naught, by looking up the table what value of z naught do we get probability of z naught is 0.95.

Once we know that. So for example, if you look up the normal distribution tables you will see that z naught is roughly 1.645. So, then you can use this equation to invert, and find out what is the value of x naught. So, that completes our discussion of normal distribution. So, today we briefly discussed about some cases of uniform distribution and then we came and described about normal distribution. So, for our normal distribution we derived that your expectation is equal to mu, and your variance is equal to sigma square.

So, you can define us another normal distribution where you define. So, you rescale the axis. So, you define z equal to x minus mu by sigma, then you will get a distribution which is centered at 0. So, it will have mean of 0 and standard deviation of minus with that you can. So, you can use then you can use this particular distributions to scale and because you have tables which have these values, you can transfer any translate any variable x into a standard normal variable z, and then look up the table to find out what exactly is the value with that.

I thank you for your attention and we will meet again in next class.