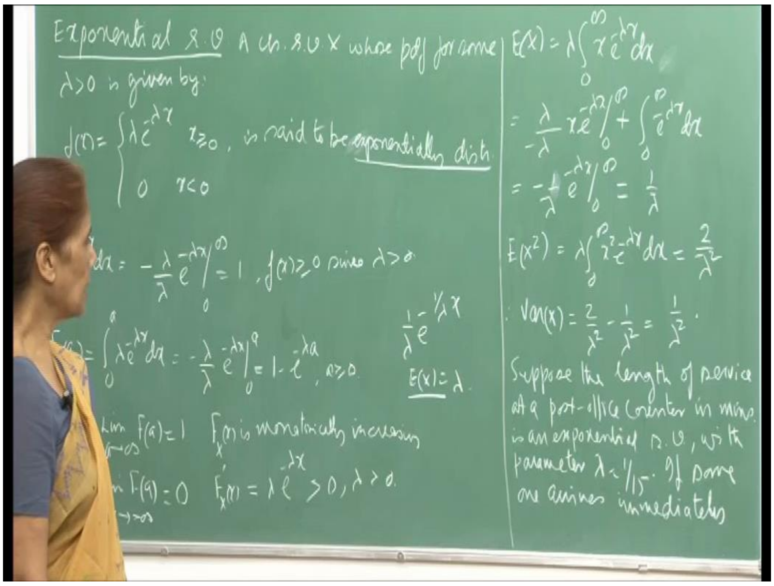
Introduction to Probability Theory and its Applications
Prof. Prabha Sharma
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture - 10
Continuous Random Variables and Their Applications

(Refer Slide Time: 00:19)



After discussing we normal random variable, we will today talk about the exponential random variable, a continuous random variable x whose PDF for some lambda greater than 0 please note that, the parameter has to be positive. Then f x is equal to lambda e raise to minus lambda x for x non negative and 0 otherwise is said to be exponentially distributed, all has an exponential distribution whichever way you want to say it. So, again we validate that this is the PDF indeed a PDF. So, it is non-negative since lambda is non-negative.

So, therefore by definition this is non-negative and the integral from 0 to infinity lambda e raise to minus lambda x d x would be minus lambda upon lambda e raise to minus lambda x 0 to infinity. So, it infinity it is 0 at 0 it is 1, so and with the minus sign, so minus minus plus, so this is equal to 1, so this is indeed a PDF, then you want to compute it is distribution function. And here this would be 0 to a, so this is probability x less than or equal to a, therefore this is integral 0 to a and this again comes out to 1 minus e raise to minus lambda a, a greater than or equal to 0, because our variable itself is non-negative.

Now, we verify the conditions that CDF must satisfy and so limit f a as a goes plus infinity is 1 from here you see as a goes to infinity this will go to 0. So, this reduces to 1 and I should have also written down the limit f a as a goes to minus infinity, see we have any way said that x is less than, I mean for x less than 0, this there is no math's and when can I say it from here if a is less than 0 than of course, this integral is not defined, I mean if a is less than 0 than this integral is 0 are you can argue that there is no math's for x less than 0.

So, limit f a a goes to this is 0, then f x is monotonically increasing, so therefore, you take the derivative of f prime of f x, which is f prime x that will come out to be lambda you differentiating. So, your f x would be well you treat a as x does not matter, so if you are differentiating this, so lambda would come here minus lambda minus lambda minus plus e raise to minus lambda x and since lambda is non-negative because, lambda is greater than 0. So, therefore, this is again non negative and so the function is monotonically increasing.

So, all the properties of cumulative density function have been are satisfied, then we find out the expectation of the random variable and that will be integral lambda 0 to infinity integral 0 to infinity x e raise to minus lambda x d x and by integration by parts I treat this as the first function. So, this will be this and now you see that at 0 this is 0, so therefore, and this is 1, so the product is 0 at infinity e raise to minus lambda x goes to or if you can write it as x upon e raise to lambda x, then e raise to lambda x goes to infinity much faster than x does, so the ratio tends to 0.

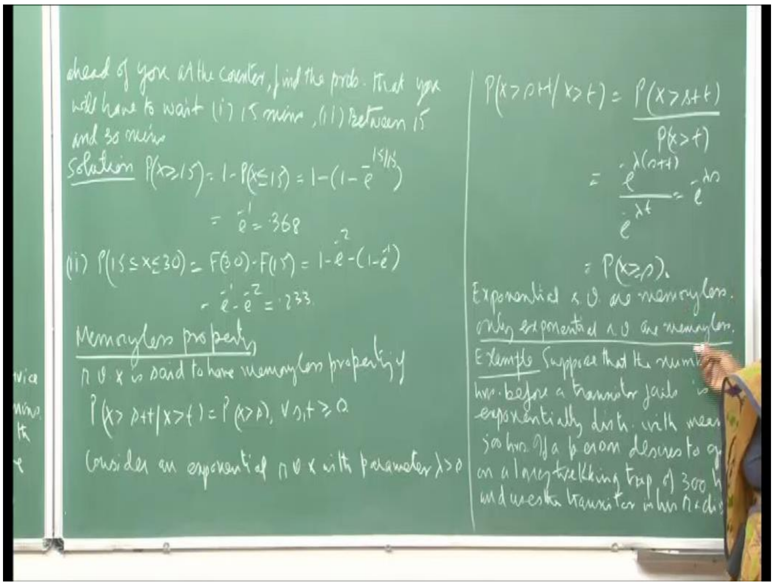
So, therefore, no contribution from this term and you are left with just this, so therefore, again you integrate minus 1 upon lambda e raise to minus lambda x there would be infinity that gives you 1 by lambda. So, the way to remember is that if lambda is a parameter for the exponential distribution and it is defined in this way, then the inverse of the parameter is your mean or the expectation some places people define it as 1 by lambda minus 1 by lambda x.

So, in that case you are expectation will become lambda, so it is inverse of the parameter whatever you use for defining this, the inverse of that will come out to be the. So, then to find out variance you find out the expectation x square and I have not done the calculations here. But, again in the same way you will have to two iterations of the integration by parts, here x square is the second function, this is the first function and you continue doing it, so then that comes out to be 2 by lambda square.

And therefore, variances 2 by lambda square minus 1 by lambda square, which is equal to 1 by lambda square. So, quick example normally this would be associated this distribution was would because, you know when you go to any public place where there is a service counter for example, post office or railway booking and so on. Now of course, it is mostly online, but still people have to go to counters for all the services, then you know the time that the clerk will take to service the customer is most of the time random variable.

And, so exponential variables do model that situation quite few this thing, so suppose the length of service at a post office counter in minutes is an exponential random variable with parameter lambda is equal to 1 by 15. So, immediately you can say that the exceptive number of time that the exceptive duration of service to a customer would be 15 minutes. Because, the expected value of the this random variable would be 1 by lambda, which is 15 minutes.

(Refer Slide Time: 07:00)



If someone arrives immediately ahead of you at the counter find the probability that you will have to wait 15 minutes between 15 and 30 minutes to find the probability that you have to wait for 15 minutes. That means, at least 15 minutes, so therefore, the event would be x greater than or equal to 15, where is x is your time for you have been to wait.

So, probability x greater than or equal to 15, which will be equal to 1 minus probability x less than 15.

And therefore, we have computed this already while computing for x less than exponential distribution. So, this would be 1 minus of 1 minus e raise to minus 15 by 15 because, your lambda is 15, so therefore, this probability comes out to be e raise to minus 1, which is 0.368. Since, as I have said that because, we have already made this computation that this will be equal to 1 minus e raise to minus 15 by lambda and your lambda is 15. So, therefore, this will be equal to e rest to minus 1 and so this is 0.368.

So, similarly for the second one probability 15 less or equal to x less than or equal to 30; that means, you are waiting time is now between 15 and 30 minutes. So, that will be again by the same computations will be f 30 minus f 15 and which again would be 1 minus e raise to minus 2 because to be 30 by 15, which is minus 2 and then this is minus of 1 minus of e minus 1. And so because this is less than or equal to 15 and so less than or equal to 15 less than 15 for a continuous distribution, so it will be this the probability.

And, so it is e raise to minus 1 minus e raise to minus 2, which is 0.233, now I want to show you another important property of the exponential distribution. So, first of all we will talk about the memory less property and we say that random variable x said to have the memory less property. If probability x greater than s plus t given that x is already greater than t is equal to probability x greater than s for all s and t non negative, which means that it does not matter how long you have already waited, if you are asking for this probability x greater than s plus t then it is same as probability x greater than s.

So, therefore, the system we are modeling does not have memory less property and so we are talking of random variable whose distribution satisfies this condition. So, now I will show you that the exponential distribution among all continuously distributed random variables are among all continuous PDF's exponential distribution have the probability of being memory less distribution, which are discrete and which have also the memory less property.

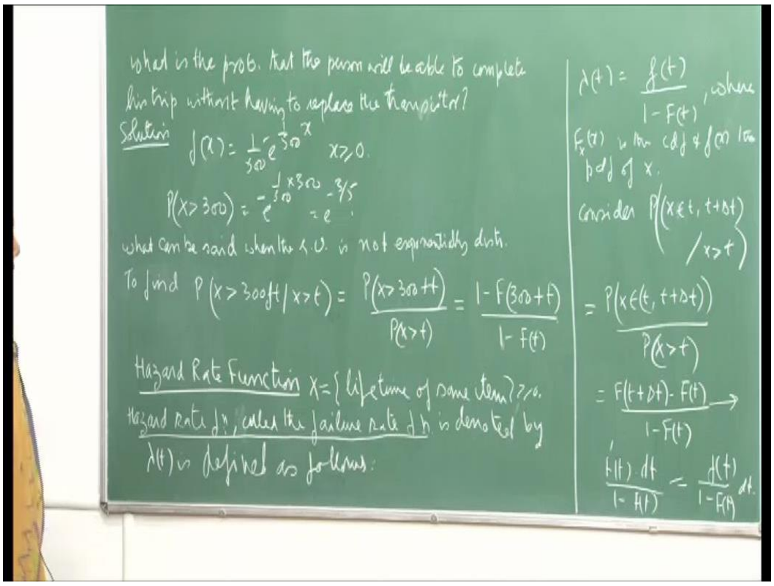
But, among continuous distribution random variables, which is exponential distributed has the memory less property. So, in case this random variable is exponentially distributed with parameter lambda, then you see if you write down this expression probability x greater than s plus t given that it is x is greater than t, then this will be this because, even the product of these two would be intersection of this two when should be x greater than s plus t, you would because, it is t smaller than s plus t.

So, this reduces to this expression x greater than s plus t divided by x greater than t, now by our definition this is because, f a is 1 minus e minus lambda s. So, 1 minus f a would be simply e raise to minus lambda a, so here it e raise to minus lambda s plus t divided by e raise to minus lambda t and which is this, which is equal to probability x greater than or equal to s. So, exponential random variables are memory less and in fact, little more arithmetic would be required or calculus to show, that if you impose this property.

Then you can actually show that the only exponential random variables have the memory less property. So, since this is being our first course I am omitting the mathematics here, but who was interested can sit down and work it out and see that, when you start with this condition for a continuous random variable, then you will see that only exponential random variables the PDF, which will satisfy this condition will actually the exponential PDF.

So, let us again look at an example, suppose that the number of hours before a transistor phase is exponential distributed with mean 500 hours. So, here the lambda is 100, if a person desires to go on a long trekking trip of 300 hours and uses the transistor in his radio, then we want to find out what is the probability that the transistor will not fail.

(Refer Slide Time: 12:27)



What is the probability that the person will be able to complete his trip without having to replace the transistor. So, he wants that the probability that the transistor will not fail for 300 hour long tracking trip that he is undertaking, so the solution is that affects is one upon 500 e raise to minus 1 upon 500 x. Because, in the problem it says that the hours the lifetime of the transistor is exponential distributed with mean 500, so as I was telling you if the mean is 500, then the parameter will be one upon 500.

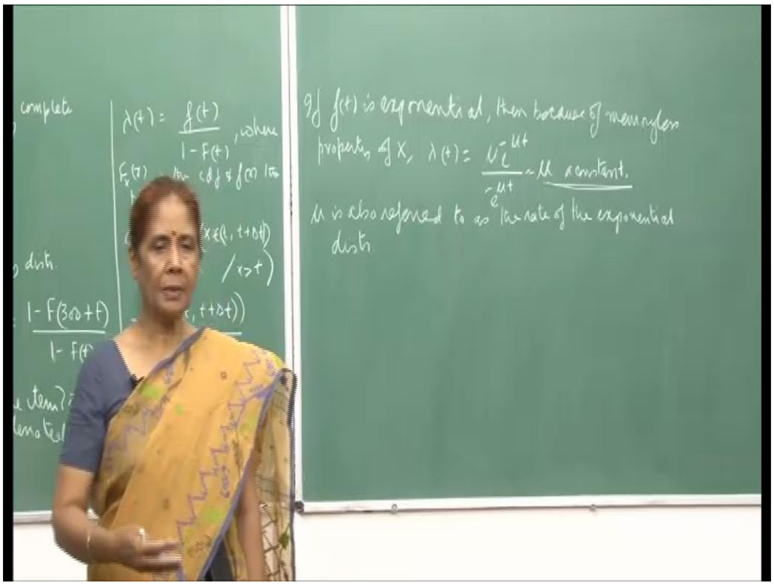
So, therefore, the PDF of the random variable representing the lifetime of the transistor would be given by this x nonnegative. So, therefore, probability x greater than 300 is simply e raise to minus 3 by 5 because, it is e raise to minus 300 upon 500 which is e raise to minus 3 by 5 and you can compute this value from the table. So, what can be said when the random variable is not exponential distributed and that case the memory less property is not there.

And, so this would be simply conditional probability in terms of capital f and so that is it you see the advantage of having random variable, which is memory less. Now, another thing that is useful and can be again for exponential random variable it gets quite simplified and that the hazard rate function or sometimes you also called the failure rate function. So, if x is a random variable which is representing lifetime of some item and of course, it does not non negative variable and the CDF the PDF is given by small f and CDF by capital F.

Then hazard rate function called the failure rate function is denoted by lambda t and is defined as follows. So, lambda t is small f t that is the PDF divided by 1 minus CDF, so and the simple explanation is possible, so here again what we are saying is that x belongs to this is the bracket here, t plus delta t, delta t is very small and then given that x is already work for t now what we are saying, that it will fail just after t because, t plus delta t. So, x lies in the interval t comma t plus delta t given that it has been functional till time t.

And, so this is probability x belongs to t comma t plus delta t, delta t is a positive quantity very small. So, therefore, when you take the intersection these two you get this event divided by probability x greater than t and so this will be by our definition f t plus delta t minus f t divided by 1 minus f t. So, now, you see the limit of this delta t becomes smaller, then you know you divide by delta t and multiply by delta t, so this divided by delta t remember limiting value of this is the derivative of f at t. So, which becomes f prime t into d t upon 1 minus f t and f prime t is your PDF f small f t divided by 1 minus f t d t. So, therefore, the definition of the failure rate because, as we see that it should fail just that t time, time t is rate has been functional up to time t then it fails.

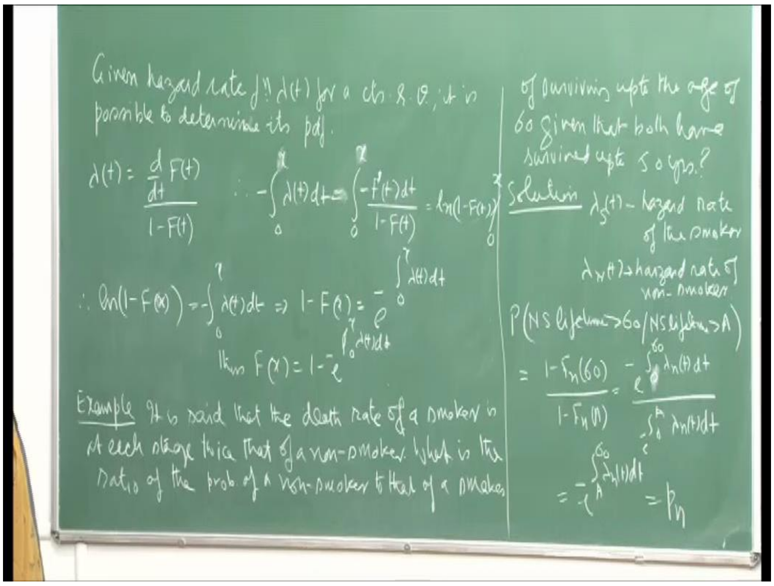
(Refer Slide Time: 16:24)



So, now if f t is exponential you see when you write out this will be small f t, so the PDF is mu e raise to minus mu t, I used symbol mu because, lambda is already being used here and this is 1 minus f t would be e raise to minus mu t. So, you see this is mu a constant, so for exponential random variable we have that rate is a constant, it is not a function of t. Otherwise, as you see that this will be a function of t, so the hazard rate function is dynamically changing depending on the time.

That means, if you are talking of lifetime, so as it should be, but for exponential distribution exponential is simple, since it is memory less therefore, the hazard rate function does not change it is a constant. So, this is the rate of failure and so because it is memory less, it does not matter how old the instrument is the probability of it is failing any time is same. And so here the rate is also constant mu is therefore, also referred to the rate of the exponential distribution. So, now, mu is the parameter this is also the rate of failure for exponential distribution and we saw that one upon mu will be the mean of the exponential distribution.

(Refer Slide Time: 17:40)



So, given hazard rate of function lambda t for a continuous random variable, it is possible to determines it PDF. So, we will just show because, you see lambda t I can write as d d t of f t 1 upon minus f t and then if I take the integral of both the side attach a minus sign. So, there is minus sign here, this is 0 to x lambda t d t and this is integral 0 to x minus f prime t, I written this as f prime t d t upon 1 minus f t, now by formula for integration because, derivative of this is this therefore, this will be l n of 1 minus f t.

So, therefore, l n of 1 minus f x because, you computing from 0 to x, so there is 1 minus f x will be and this is integral minus 0 to x lambda t d t, I hope this is clear. Because, see this will come out to be l n of 1 minus f t and this is from 0 to x, so 0 to x when you put x here, this will be l n of 1 minus f x and at f 0 is what, f 0 is 0. So, l n of 1 is 0, so therefore, the contribution from here you get is l n of 1 minus f x, you remember limit f x as x goes to minus infinity 0. So, therefore, I am just using that. So, that reduces to l n of 1 minus f x and this right hand side this is minus 0 to x lambda t d t.

So, therefore, you can say 1 minus f x is equal to e raise to minus 0 to x lambda t d t and so f x you can write down from here as 1 minus of e raise to minus integral 0 to x of lambda t d t. So, if I know lambda t, then I can integrate here and then my f x will be of this form, so; that means, it is enough if you know the hazard rate function of random variable, you can determinate distribution. So, once you know the equality density function, you can determine the PDF also.

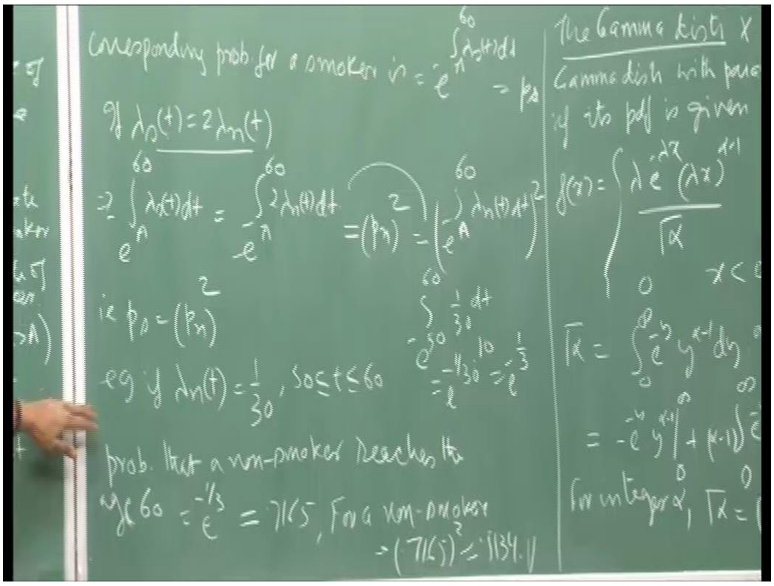
Now, I will just illustrate the concept some more through an example, this is see it is said that the death rate of a smoker is at each stage twice that of a non smoker. That means, this that your age gets reduced by half, if you are a smoker compared to what a non smoker. So, what is the ratio of the probability of a non smoker to that of a smoker of surviving up to the age of, so suppose I am just saying that, what the both of them surviving up to the age of 60 given that both have survived up to 50 years, you want to find out the probabilities.

So, may be ratio part is not important all I am saying is let us find out the probabilities that non smoker will survive up to the age of 60 given that he or she has survived up to 50 years. Similarly, a smoker the probability what is the probability of a smoker surviving up to 60 years, given that he had survived up to 50 years. So, I will define lambda s t as the hazard rate of the smoker and lambda n t as the hazard rate of non smoker. So, now, let me just right now take instead of 50 just let me take some years A age A.

So, what we are saying is that a probability that a non smokers life time is more than 60 years given that the life time of the non smoker has been more than A is the conditional probability. Because, again the intersection these two because, 60 is more than A, so this is probability that the non smoker has been, the life time will be more than 60, so that becomes 1 minus f n 60 divided by 1 minus f n A. And from here, I will see this is 1 minus f x is e raise to minus integral 0 to x lambda t d t. So, that is would be 0 to 60 lambda n t d t and 1 minus f n a will be 0 to A lambda n t d t e raise to...

So, this is what we just computed here and therefore, if you take this upstairs, then you see the integral this become e raise to A to 60. So, this will be e raise to minus integral of A to 60 lambda n t d t and this let me call as p n, so the probability of smoker surviving up to the age of 60, given that he has already survived up to the age of A.

(Refer Slide Time: 23:01)



And correspondingly for a smoker this probability of surviving up to 60 years given that he had survived up to A years is lambda s t d t and I am calling it as p s. So, if the belief is that lambda s t is twice lambda n t; that means, the death rate is twice as high for a smoker compared to a non smoker, then I substitute for lambda s t twice lambda n t here. So, that will be twice lambda n t and so this would becomes square of I would not writing the step I mean actually this is equal to e raise to minus A to 60 lambda n t d t square, which is probability p n square.

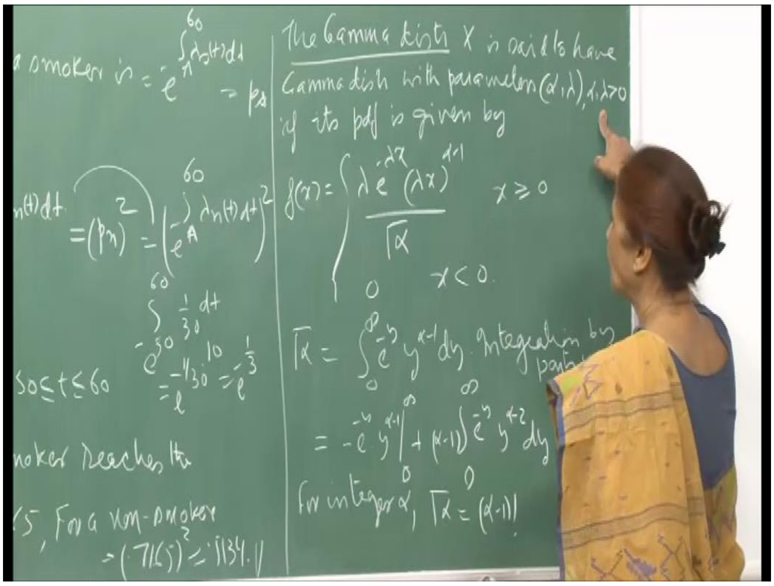
So, the effect is that the probability gets squared up for a non smoker, so the probability of surviving up to the age of 60 for a smoker is square of the probability of the non smoker surviving up to the age of 60. So, this 50 was not really because, A could be anything here, so given at that is why we said, that at any age at each stage, so therefore, it does not matter when you are making this comparison. So; however, the old both the people are after that if you want to say what is their age of surviving up to 60, when the probability for the smoker is square of the probability for the...

So, now, here as I said that if you take lambda n t to be 1 by 30; that means, remember I am taking the situation when. So, that means, this is now because, this is constants, so therefore, I am taking the exponential situation, that is the random variable the lifetime is random variable is the exponential distributed, then a probability that non smoker reaches the age 60 would be because, this lambda n t is 1 by 30.

So, you want to compute will it come out to be e raise minus 1 by 3, e raise to minus you are saying this is 50, 60 and 1 by 30 d t. So, what will that be 1 by 30 into 10. So, this is I mean e raise to minus, this is e raise to minus 1 by 3, so if this is a constant; that means, it corresponds to exponential random variable. And so this is e raise to minus 1 by 3 which turns out to be 0.7165, so for a non smoker and for a smoker it would be a square of this, which will be 0.5134.

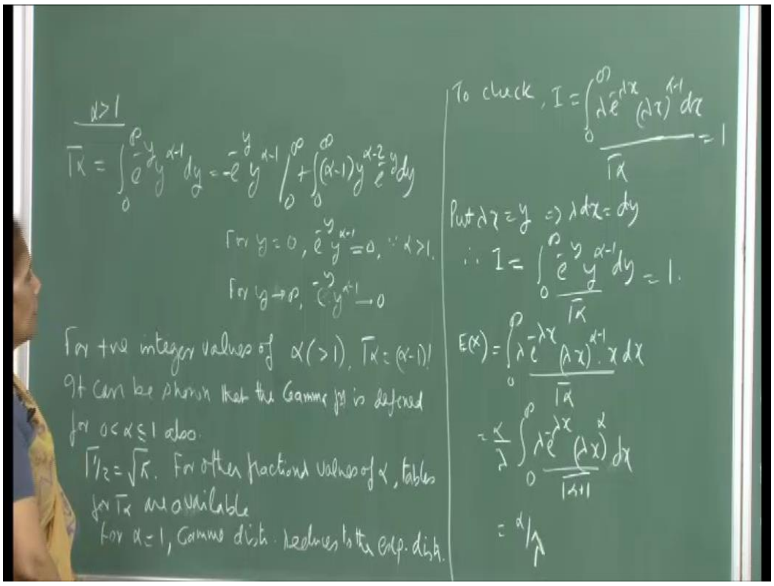
So, see how fast the probability has reduced because, person smoking, so the probability of a smoker surviving up to the age of 60 is 0.5134 and for a non smoker it is 0.7165. So, that we wanted can have many more applications and it has will go long may be I have put some problem related to hazard rate function in your exercise 4 also.

(Refer Slide Time: 26:42)



Now, we continue with some more special continuous random variables and the next one is the gamma distribution and x is said to have gamma distribution with parameter alpha and lambda, both the parameters have to be positive PDF is given by this equation. So, effects is lambda e raise to minus lambda x lambda x into raise minus 1 upon gamma alpha and that is why the name.

(Refer Slide Time: 27:05)



So, this is let me show you the calculation for the gamma alpha, so let us compute the value of gamma alpha for alpha greater than 1. So, therefore, the definition is this is equal to 0 to infinity e raise to minus y, y rest to alpha minus 1 d y right, so integration by parts and therefore, this will be the derivative here is e raise to minus y should be minus. So, here this is minus e raise to minus y, y rest to alpha minus 1 0 to infinity and then this becomes plus because, minus minus is plus.

So, therefore, this will be plus 0 to infinity then derivate of this is alpha minus 1 y rest to alpha minus 2 e raise to minus y d y. Now, let us compute the values at the end points, so for y equal to 0, see this will be 0 because, alpha is greater than 1 and that is why it is important. So, alpha is greater than 1 therefore, this is a positive power, so therefore, this is 0 and of course, at 0 e raise to 0 is 1, so this is equal to 0 and again y goes to infinity, then this goes to infinity faster than this here again because, alpha is greater than 1, so this exponent is positive and therefore, this will go to 0.

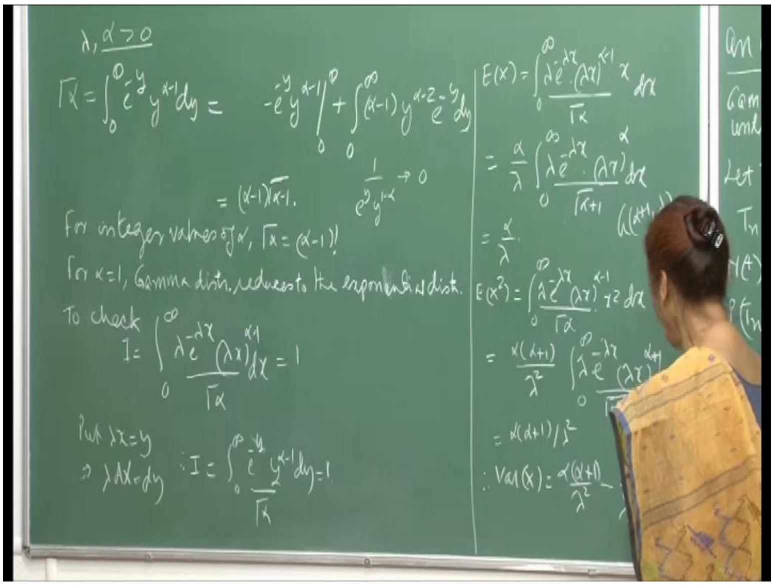
So, therefore, no contribution by this term and so your integral, so gamma alpha reduces to just this, which by our notation will be gamma alpha minus 1 should. So, therefore, I did not write, so here it relatively; that means, this integral will be now alpha minus 1 and so for positive integral values of alpha, if I you know go and doing it iteratively. So, for integer values of alpha positive integer values of alpha, this we can see is gamma alpha would reduce to factorial alpha minus 1.

You can see that because, as go on, so the finally, what you will have be this will be then I should have, this is alpha minus 1. So, this should be equal to alpha minus 1 gamma alpha minus 1 from here alpha minus 1, so therefore, as you go on the next iteration it will be gamma alpha minus 2. And so as we go on alpha is a positive integer therefore, you end up finally, with just integral 0 to infinity e raise to minus y d y and so this will reduce to alpha a factorial of alpha minus 1 for alpha be positive integer.

Now, it can also be shown that the gamma function is defined for alpha between 0 and 1, this is also possible we can also show that the integral will is defined that it will be finite value. So, for all values of alpha between 0 and 1, the integral is also defined; that means, gamma alpha is defined for alpha between 0 and 1. And one important value is gamma 1 by 2, which is root pi and this integral we will obtain this values later on in the forth coming chapters.

So, I will talk about fractional values of gamma alpha between 0 and 1 and there are tables available for fractional values of alpha. The tables available non negative fractional values of alpha tables are available and for alpha equal to 1 the gamma distribution reduces to the exponential distribution.

(Refer Slide Time: 31:00)



All of to check the see our gamma PDF is given by this, so when you put alpha equal to 1 this term is gone and so your PDF reduces to lambda e raise to minus lambda x and gamma alpha is also 1. So, therefore, you will be the gamma PDF reduces to lambda e raise to minus lambda x for x no negative, so therefore, for alpha equal to 1 the gamma distribution reduces to the exponential distribution. So, this is one relationship and then I will show you some other relationships between many other this thing.

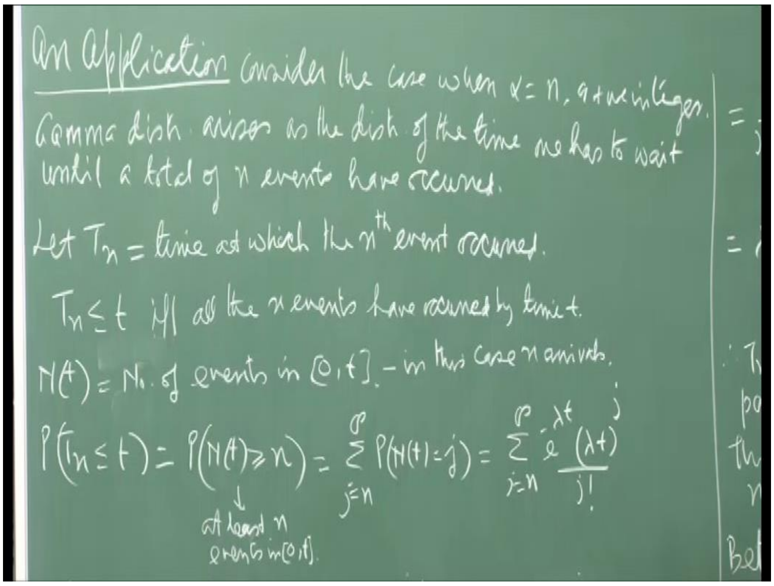
Now, you want to again check that this PDF is a valid PDF and therefore, I have to show that this integral will evaluated to 1. So, here of course, you put lambda x equal to y, then lambda d x is equal to d y and immediately this integral reduces to e raise to minus y lambda d x gets replaced by d y and this is y rest to alpha minus 1 upon gamma alpha. And so from here you see that divide by gamma alpha here and this equal to 1, so the validation is complete.

So, therefore, this is the valid PDF and now you want to compute the expectation of this gamma random variable, then you have to multiply this by x and integrate, but it is easy to manipulate because, I will add x to this and lambda also. So, I divide by lambda, there I multiply by alpha and multiply here by alpha, so this becomes gamma alpha 1 by our definition. So, this will be gamma alpha 1 and this will be lambda x rays alpha, so; that means, the PDF of gamma here the parameters are alpha plus 1 and lambda.

So, therefore, since this is again the PDF for gamma alpha 1 comma lambda, so again it will integrate to 1 and so you will be left with alpha by lambda. So, the expected value of alpha, lambda, gamma, variable is alpha upon lambda, so; that means, if the convention to write this first and then this. So, then this divided by this that is you are taking the gamma distribution alpha comma lambda, then similarly expectation x square can also be just by simple manipulation computed immediately, what I will do I need lambda square here to bring together with this.

So, I divide by lambda square and I multiply by alpha alpha 1 to make it gamma alpha plus 2 and this will be lambda x rays to alpha plus 1. So, again this is the PDF of gamma alpha plus 2 lambda, so this will integrate to 1 and I will be left with alpha into alpha plus 1 upon lambda square. And so variance will be this quantity minus alpha lambda whole square and that gives me alpha upon lambda square, so simple calculations to tell you the required quantities.

(Refer Slide Time: 34:25)



Now, let me just show you an application and this application may be I will using a concept which we have to do, but it does not matter I still start that this would be good time to mention this application. So, here see this is we are considering the case on alpha is positive integer, when alpha is equal to n is a positive integer, now gamma distribution e raise as the distribution of the time one has to wait, until the total n events have occurred.

So, it is like you go to a railway booking counter, then you have people ahead of you in the queue and each person as I have told you that the we will treat service time as random variable. So, and remember that the I shown you that the service time being random variable can be an exponential random variable, so for each person the service time is random variable and then since there different people independent people. So, each one get serviced, so then the total time would be sum of that many independent random variables.

And exponential distributed random variables this is the idea, so now, the what I am trying to show you is that, the gamma variable is actually the time one has to wait till all people ahead you have been serviced and you have also been serviced. So, the way we will measure it is that when I am saying alpha is equal to n, so here I could be counting that you also have been serviced. So, then until a total of n events have occurred and of course, later on when we do the poison process and so on, then the whole thing will become much more clear.

But, you can just get a feeling for the application that I am trying to discuss here, so now, let T n be the time at which the n th event is occurred. So, when T n less than or equal to T this event could that mean, if and only if all the n events are occurred by time T, T n is the time at which n th event occurred. So, now, T n less than or equal to t will be that by time t all the n events have should have occurred, means in our particular case all n people have been serviced by the railway ticket booking counter.

Now, let capital N t be the number of events in 0 t, see in this particular case if the people ahead of you like say n minus 1 people ahead of you, you are the nth person. Then; that means, the time span you are taking is 0 to t, then that many people should have arrived in the time 0 to t, when they are n people in the system, then only they get serviced this is how. So, in this case there are n arrives in this time and so when we look at probability T n less than or equal to t, if you want to compute this probability.

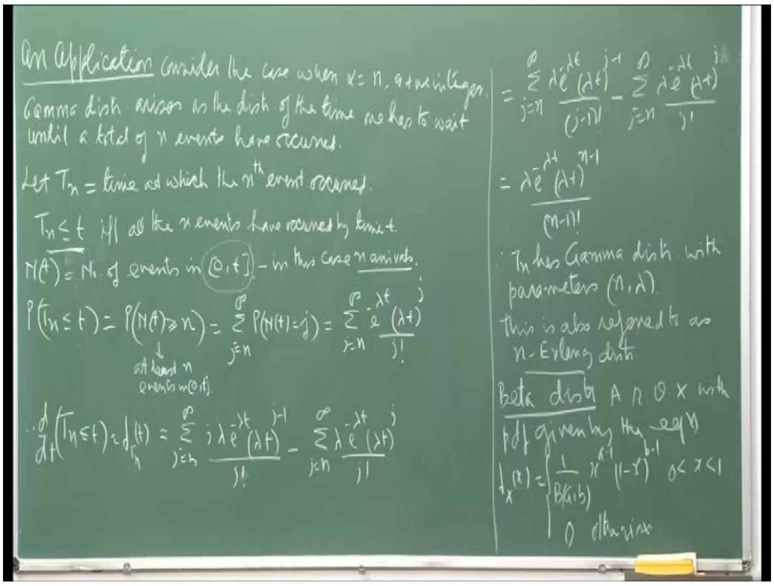
Then this is same as probability n t greater than or equal to n because, at least n events in 0 t; that means, at least n arrives must be there can be more. But, when we are talking of n people to be serviced in this span of time, then at least at many people should have arrived that many events must be there in the system. So, this is what it is and this we will therefore, because this discrete thing people arriving are discrete events.

So, j this is to be j from n to infinity probability n t equal to j probability that there are j people in this system time 0 t. And then you summit of from j equal to n to infinity, now this is the part, see very often when you talk of events discrete events occurring in span of time, under certain conditions it can be shown that these will be poison arrivals. That means, the number of arrivals in a span of time would follow a poison distribution and therefore, probability n t this I will proof oftenly and we are talking about process is also one.

So, then in detail I will discuss how you arrive at this probability when you under certain assumptions you can show that the probability n arrivals in the time 0 t would be given by this. So, you summed up j to n, so it means for n t equal to j actually n t is the poison random variable and so the mean value becomes lambda t because, you are taking the span 0 t, so this will later on we explained. So, therefore, this is your poison probability is summit up from n to infinity.

Now, this is your cumulative distribution function for t n to find out the PDF, I will take the derivative, which is f t n of t and. So, we differentiate this expression with respect to t and you see first differentiate this and lambda comes out and j, so j lambda e raise to minus lambda t, lambda t rest to j minus 1 divided by j factorial minus derivative of this, which will be lambda minus lambda e raise to minus lambda t, lambda t rest to j divided by j factorial. And here j varies from n to infinity and so just rearrange the things little bit j when cancels out here. So, it will be j minus 1 factorial lambda t rests to j minus 1.

(Refer Slide Time: 40:17)



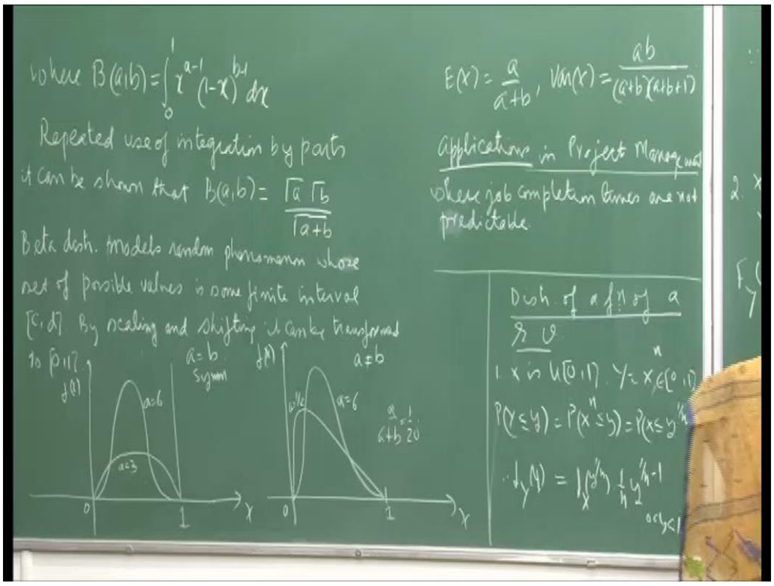
So, what you see is that here the terms are from starting from j minus 1 and here it is j, so you see things will cancel out in pairs. And except the first term here will be left out, which will be lambda e raise to minus lambda t lambda t rests to n minus 1 divided by n minus 1 factorial, that is only one because, when you put j equal to n plus 1 here that will be lambda t rest to n upon n factorial. And here also j equal to n it will be n factorial and lambda t rests to vise it j minus 1 here.

Because, when I am differentiating with respect to this one here then lambda t rays to j remains in that, so this is it. So, therefore, this term will cancel out with the second term here and then the third term here will cancel out to the second one here and so this will process will go on only the first will be left out here, which is lambda e raise to minus lambda t lambda t rest to n minus 1 upon 1 minus 1 factorial, so this is now gamma distribution with parameters and gamma lambda.

So, therefore, the amount of time a person has to wait till he serviced and if there are n minus 1 people ahead of you in the queue. And so that is random variable and just now shown that when the arrivals are poison, then this will be gamma distribution with parameters n comma lambda. So, in this case this is also referred to as n erlang distribution that is another name in literature you might some books may referred to this distribution as n erlang.

So, we have got some feeling about the gamma distribution and as we go on I will give you some more inside into the thing. Now, the other distribution continuous variable distribution, which is of important is the beta distribution, so a random variable x with PDF given by the equation f x x is 1 upon this is the beta function a comma b x rays to a minus 1 1 minus x rays to b minus 1 x between 0 and 1 and 0 otherwise.

(Refer Slide Time: 42:34)



The integral we denote by b a comma b, so this is 0 to 1 x rays to a minus 1 1 minus x rays to b minus 1 d x, where a and b both are positive. Now, again just as for the gamma distribution, we can show that for a greater than or equal to 1 and greater than or equal to 1 the integral will converge. And in fact, for integer values just like we did for computation for gamma distribution, it can be shown that this integral will be equal to gamma a gamma b by gamma a plus b.

So, therefore, I should again correct my statement here that is the beta PDF is this function divided by this. So, it becomes gamma a plus b divided by gamma a gamma b and therefore, the integral will ((Refer Time: 43:23)) to be one, so the beta PDF is actually this divided by this number and so we denote this integral by b a comma b which is. So, therefore, it has I will since defined for all a b positive, now for a greater than or equal to 1 and b greater than or equal to 1, you can show by integration by parts the integral will converge and will be equal to this.

Therefore, fractional values of a and b also it can be shown that this is defined the integral is defined and it is equal to gamma a gamma b upon gamma a plus b. Now, many useful applications of the beta distribution and one case that it models random phenomenon whose set of possible values is some finite intervals c comma d. So; that means, all possible values of this random phenomenon occur between within a certain interval c comma d.

And, but ((Refer Time: 44:24)) we have here defined the variable to be from 0 to 1, so then by scaling and shifting, we can transform this interval to 0 1. And of course, one obvious transformation is that y is equal to x minus c upon d minus c, so then all possible values of x which are within c d will now be, so the corresponding y variable will have all values between 0 and 1. So, and then we will see for applications of the beta distribution and will compute the other quantities related with the beta distribution.

Now, I am just trying to give you the pictures of graphs of this beta distribution for different. So, now, when a is equal to b this graph is symmetric, the graph of the beta function PDF is symmetric and for example, a equal to 3 it will be something like this and as a becomes bigger, the mass gets concentrated this the graph becomes narrow or and this is symmetric. So, if you draw for a equal to 10 probably it will be something like this peak the peak will be higher and so on.

Now, for a not equal to b the graph is asymmetric and skewed towards the left, so for example, a equal to half it is almost skewed towards the y axis and as a increases again the skewness shifts to the center. And this of course, the graphs are not drawn to scale, but they indicated a upon a plus b is equal to 1 by 20, so in this situation suppose a is 6 we can find out what the value of b is and so on. Now, there are situations see for example, if you have a big project, in which you have lot of jobs and so of course, a big project will be made up of number of jobs.

And this project may not have been handled completely before, so there is lot of uncertainty about the job completion times. And as a project manager he has to or he or she has to you know sometimes be have some estimate as to how long it will take for the whole project to be completed, which means that, must have a good idea as to how long it will take for each job to be completed, now in the absence of any previous experience because, jobs have not been performed.

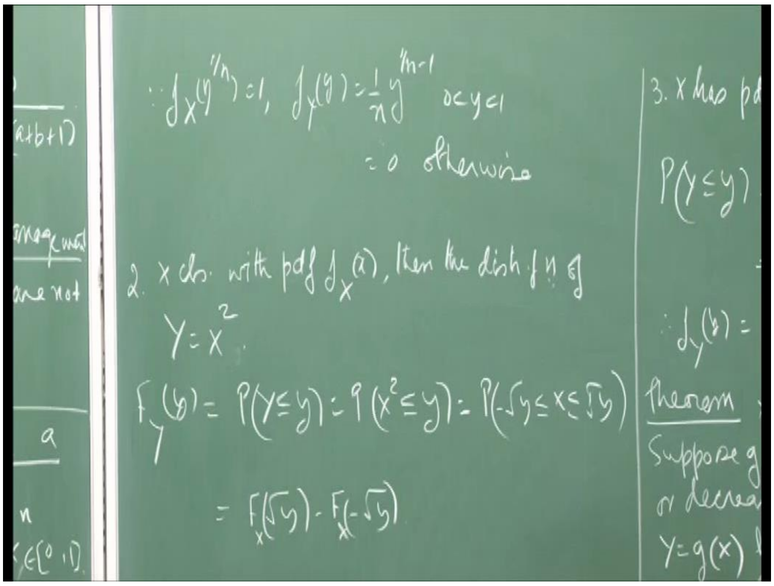
For example, of course, this is no very old example that when they would trying to put a man on the moon, then it was completely new project all the job that made up the project new. So, people had no idea about the how long it will take for the jobs to be completed, but there are certain finites span and then of course, you do not expect just like as for the normal distribution for example, it is a symmetric distribution. But, then here there was no reason to believe that the completion time distributions will be symmetric.

So, therefore, beta distributions fitted the bill very well because, here were the distributions, which have they finite span and even though it is a continuous distribution and then it was not symmetric, so and so on. So, then huge projects were then the time estimations were made using beta distributions, so interesting applications, so where job completion times are not protectable, you have no idea. Then again by integration by parts, you can show that expected value of x is a upon a plus b and where else x will be a b upon a plus b into a plus b plus 1.

So, beta distributions again you want get's time one can talk about I can give an idea, how the time estimates are done using a beta distributions. Now, as we go long we also keep coming back to distributions of large functions of a random variable and I will just do some sample functions here, and then try to give general result. So, let us say suppose x is uniform 0 1 and; that means, x is taking a non negative values, then if you define the function y equal to x e raise to n then; obviously, x e raise to n will also remain in the interval 0 to 1, that means the range for y is between 0 and 1.

And, so if I want to find the PDF of y, then probability y less than or equal to small y is the same event as probability x n less than or equal to y and this will be probability x less than or equal to y 1 by n. So, remember this is because the value here are non negative, so this is the same as this inequality the n the root and now if you do differentiate both sides, this will give the PDF of y and here when you differentiate this, this will give you PDF because, now effects right this side is f x y 1 by n. So, you are differentiating respect to y and so this effects y e raise to 1 by n into derivative of y 1 by n, which is 1 by n y e raise to 1 by n minus 1 and y between 0 and 1 so but for a uniform random variable this is equal to 1.

(Refer Slide Time: 50:11)

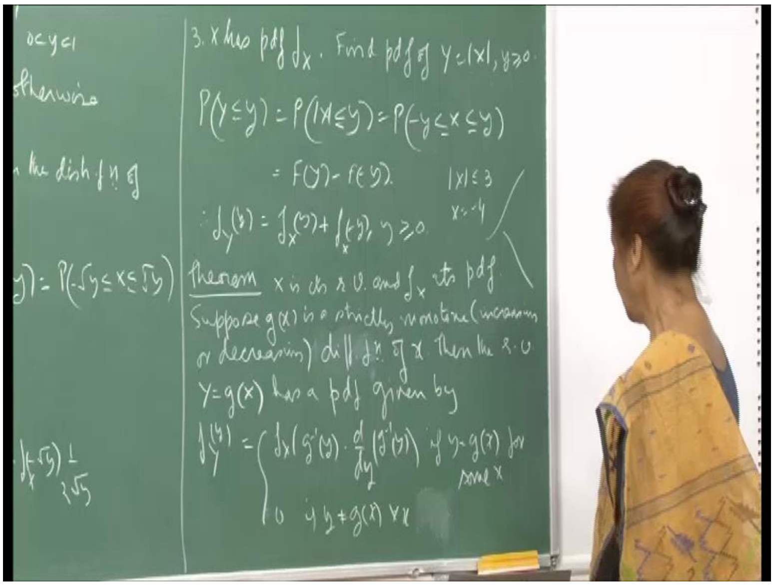


And therefore, the PDF of y reduces to 1 by n y e raise to 1 by n minus 1 when y is between 0 and 1 and 0 otherwise, take another functions. So, now, you take the function y equal to x square and this case you are not saying that x can only take no negative values, x can take negative values also. Then you see, when you write down this event probability y less than or equal to small y, which is probability x square less than or equal to y. Then this will be then equal to this event that capital X is between minus root y and plus root y.

Because, I did not if I you know sort of put the restriction that x has to be non negative, then; obviously, this it would have been just this part, this part would not have been there. But, since I am allowing x to take all positive negative values therefore, this will be equal to this event right and so by again our this thing writing the momentums of the accumulative density function, this will be affects root y minus effects of minus root y.

And then so we differentiate again respect to y and d d y f f y is the, so here this will be affects root y and then derivative of root y which usually 1 upon 2 under root y, this is x and plus affects minus root y. So, the minus minus will become plus because, there is minus coming from here, this is the minus here already, so plus affects of minus root y into 1 upon 2 root y.

(Refer Slide Time: 51:56)

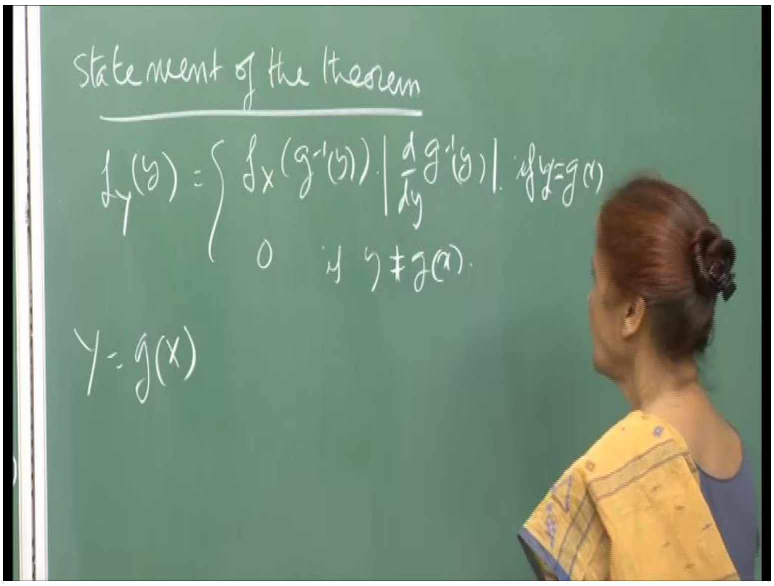


So, this would be your PDF are y equal to x square, now the third kind of function that I am looking at here is y equal to mod x. So, x has a PDF affects, then here than in this case y will have nonnegative values, even though x is negative values, so we write down the event probability y less than or equal to small y, this is probability mod x less than or equal to y, which again can be written as x between minus y and y because, the absolute value has to be less than y y is a positive number.

So, therefore, in magnitude the value x even if it is negative, it has to be higher than minus y. Because, if you are saying that mod x should be less than or equal to 3, then you are x cannot be minus 4 because, absolute value of x would be 4, which is not less than 3. So, therefore, the values of x have to be between minus y and y and therefore, this is f y minus f of minus y and when you differentiate respect to y, you get PDF of capital y, which is affects y again minus sign gets converted to plus because, there is a minus coming from the derivative of minus y.

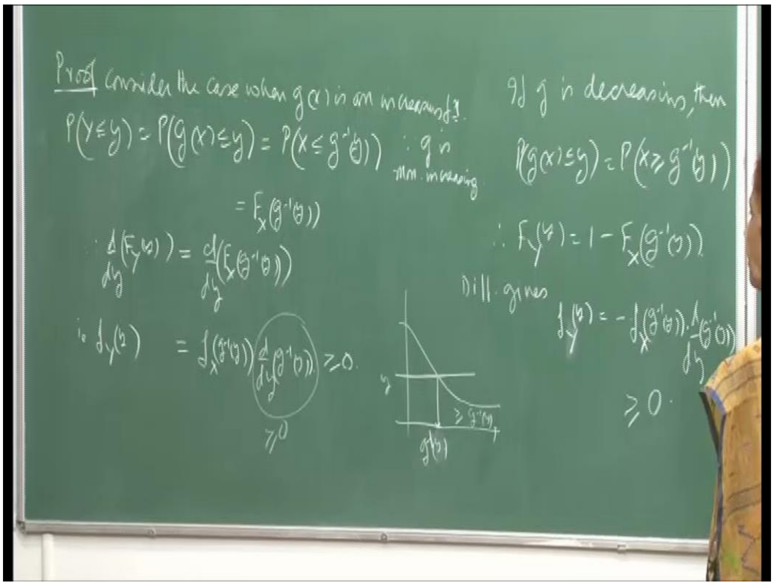
So, this y greater than y equal to 0, so one can go on, but then I will just summarize all these in this theorem. And so this is the x is the continuous random variable and affects it is PDF, suppose g x is a strictly monotone increasing or decreasing function, so this is now very clear and it is differentiable. So, strictly monotone means that it is either going like this the function like this or it is coming like this, so monotonically decreasing or monotonically increasing. Then the random variable y equal to g x has a PDF given by this and it will take a few minutes to this prove this result.

(Refer Slide Time: 54:05)



So, in this I just realized that in the statement of the theorem, this absolute value sign is missing, but that is important and I will show you why. So; that means, when y equal to when you are looking at the function g x of the random variable x and we are finding the PDF of y, then f y of small y is affects the PDF of x into at g inverse y. And absolute value of d d y of g inverse y, when y is equals to g x and 0 if y is not equal to g x and of course, here what we are saying that the Fourier's values are not being consider. Because, when you take the inverse function to be here see it is monotone function, so the relationship between. That means, this would be I do not have to worry about extra values here, so this will be fine, so now, let us look at the proof of the theorem.

(Refer Slide Time: 55:01)



So, we start with the event that y is less than or equal to small y, which is equal to this and this I can write x less than or equal to g inverse y now because, the function is monotonically increasing. So, this inequality from here, this inequality is the valid outcome because, it is increasing, so the inequality will not change a function g we have assumed is increasing function. And therefore, this is equal to f x of g inverse y is it, so differentiate both sides with respect to small y.

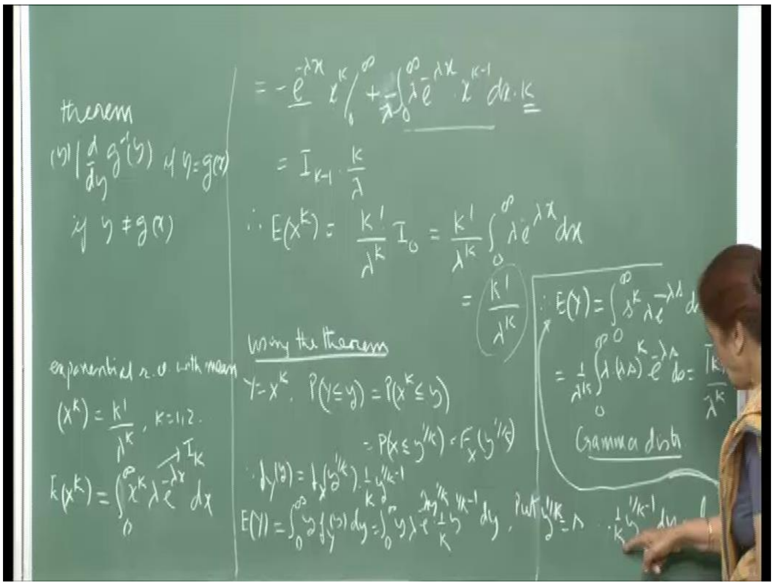
Then this f y y, which is d d y of this thing and in the step you get this is the PDF of capital Y, which is here when the differentiate capital F X you get small f x. So, that is g inverse of y into the derivative of this ((Refer Time: 55:55)) and now here because, the function g is monotone, this gain a result from calculus you can show that positive derivative. And f affects being the PDF of x this is non-negative, so the product is non-negative when therefore, this is non-negative, so this satisfies the condition.

And of course, you can verify that this is also valid PDF; that means, the integral would be 1 and so this is it. So, when you are taking the increasing function then because, the this part is nonnegative I do not have to put the bar sign here, but when g is decreasing, then you see, the probability the event g x less than or equal to y will the transform to x greater than or equal to g inverse y. And that is what I am showing you that if this is the function g x, then you are saying g x less than or equal to y.

So, g x less than or equal to y, so beyond g inverse y are the values of which your function is less than y, in the function is decreasing. And so the inequality here will reverse and that will be x greater than or equal to g inverse y and therefore, f y is 1 minus this you will write as 1 minus f x of g inverse y and then again differentiation of both sides will give you f by y, then minus f x g inverse y and derivative of this. Now, since g is a decreasing function, this derivative would be negative, so minus minus this will make it positive.

And that is why it is important that we write the absolute sign here because, your PDF cannot be negative that is a first condition. And which terms out to be because, when the function is decreasing his will be negative, so minus into this become non negative and so here again this would be positive. So, this is for the completion sake that I wrote down this theorem, but normally what we do is we ((Refer Time: 50:00)) you know do compute the PDF we write down the equivalent event, when you a function of a random variable. And then of course, differentiating both the sides you try to get the PDF for the function of a random variable, but at times it also helps to be able to use the theorem, so this thing.

(Refer Slide Time: 58:22)



Now, let me just show you as I was saying that one can either obtain results directly or using the theorem. So, if x is an exponentially distributed random variable with mean 1 by lambda, then you show that expectation of x e raise to k is k factorial upon lambda k, k varying from therefore, any finite value of k is what you had. Now, direct solution I am giving you, so solution 1 I should have said actually this is solution 1, which is a direct solution.

So, here because I know the PDF of x, so expectation x e raise to k will be 0 to infinity x e raise to k lambda e raise to minus lambda x t x. Remember we are told that the mean is 1 by lambda, so the parameter the distribution would be lambda e raise to minus lambda x t x and integration by parts treating this is the first function. So, the lambda lambda cancels minus e raise to minus why I am writing t here will be lambda x into x rise to k 0 to infinity plus minus and minus sign makes it plus 0 to infinity e raises to minus lambda x and then x rise to k minus 1 d x into k.

So, you see this integral will come out to be this and then I again multiply and divide by lambda. So, then this become my regular gamma functions, so one upon lambda, so this is we I k minus 1 into k by lambda, this integral because, when I differentiate this, this k will come out all the k should have been there, you have differentiated x k, so k is already there k is here. So, that k I am writing k by lambda and then this integral is your I k minus 1, our I k was lambda e raise to minus lambda x x rises to k and d x integral from 0 to infinity.

So, now, this integral light denote by I k minus 1 and therefore, you expected value of x k as you go on doing it by iteratively here of course, k is a positive integer. So, this is k factorial upon lambda k into I 0 finally, right if I go doing it repeatedly, so this becomes then for if I write down for I 0, this will be k factorial upon lambda k integral 0 to infinite lambda e raise to minus lambda x d x, which turns out with this, this integral is one because, anyway you know that this PDF of an exponential distribution with the parameter lambda.

So, therefore, this integral is 1 are you can directly show that this is 1 and therefore, this is the answer, now we want to use the theorem. So, using the theorem that is you compute the PDF of the random variable x is to k; that means, I compute the PDF of the function of the random variable and then through that root I try to compute the expected value. So, if I define my y's x rise to k that I am trying to compute y less than or equal to small y, which is probability x rise to k less than or equal to y, which is then x because, everything is non-negative.

So, therefore, the inequality will be converted to this that is x less than or equal to y rise to 1 by k and then f y y that is the PDF of y now will be f x of y rise to 1 by k into 1 by k y rise to 1 by k minus 1. Because, this I will be writing this is as f x of y 1 by k right and so e y will therefore, be 0 to infinity y into f f y y d y which substitute for f y y in terms of f x. So, this will be 0 to infinity y lambda e raise to minus lambda y e raise 1 by k 1 by k y rise to 1 by k minus 1 d y.

And, so e y is now here I can put y rise to 1 by k is s, so then 1 by k y rise to 1 by k minus 1 d y is d s. So, this whole thing goes to d s and so I have now here s rise to k. Because, y will be s rise to k and lambda e raise minus lambda s t s and this is now we recognize that if I take divided by lambda rise to k and combine lambda here. So, lambda s rise to k and so this your gamma PDF and therefore, did not PDF in sense that I must have here gamma k plus 1.

So, therefore, this integral will therefore, be equal to this integral v equal to gamma of k plus 1 and then divided by lambda k. So, that is another way of and since this is an integer k is an integer positive integer, so this will be k factorial upon lambda k, which we got here also, so you know whichever is convenient one can try to get the result either way.