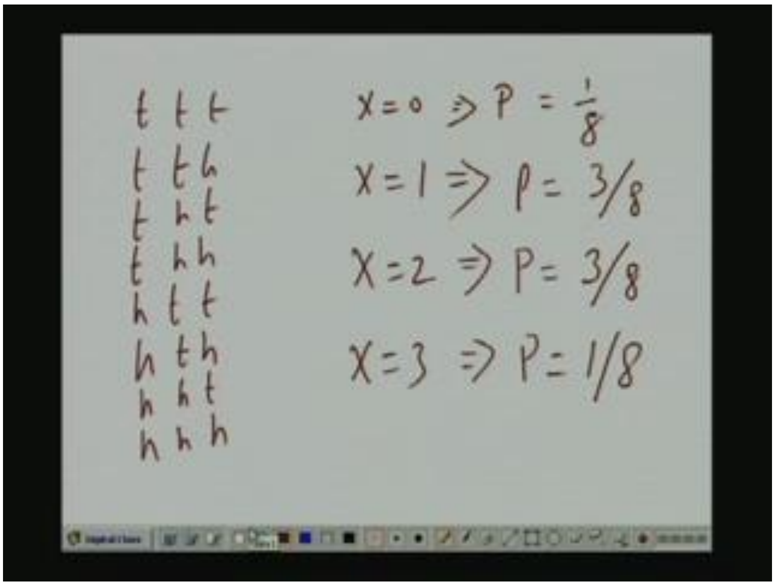
Probability & Random Variables
Prof. M. Chakraborty
Department of Electronics & Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 8
Function of a Random Variable (Contd.)

You know in the previous class, we consider a problem, then I was working it out, but then I left it half way through. I thought you would give it a try. So, let us solve that problem first and then we consider the subsequent topic. The problem that I gave was this; that there is a fair coin, means the probability of having head or tail, in either case is half and then coin is tossed thrice and the random x equals the total number of heads. So, I have to find out the probability distribution and therefore, probability density functions for the random variable x.

(Refer Slide Time: 01:39)

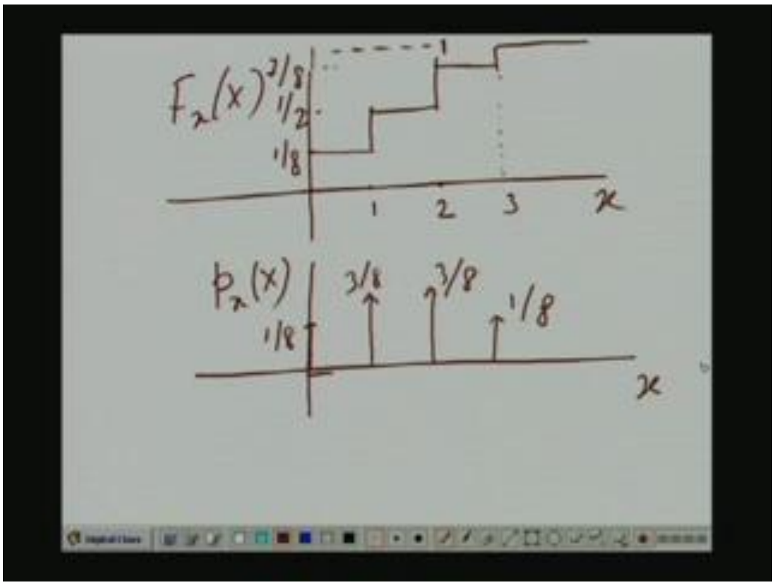


So, that I will consider this, so there these are the possibilities you know. Now, we can have tail tail tail, then tail tail h tail h tail; 1 2 3 4 5 6 7 8, 8 combinations. So, only in this case no head occurs, means the random variable X takes the value 0. Probability of this taking place is; probability of tail is 1 by 2. So, 1 by 2 into 1 by 2 into 1 by 2, so that means, X equal to 0, probability is 1 by 8. In this case, we have got head occurring once.

So, the random variable x takes the value 1. In this case also, h occurs once, then in this is another case and in all other cases, h occurs more than once. So, on 3 cases only, we have got a combination, were h occurs only once and tail occurs twice. So, occurrence of h has got a probability 1 by n 2, other 2 has probability 1 by 2 each. So, again 1 by 8 here, 1 by 8 here and 1 by 8 here, here, here and here. So; that means, x takes the value 1 with probability what? 1 by 8 plus 1 by 8 plus 1 by 8, because these are mutually exclusive these combinations, so probability is 3 by 8.

Then considered the cases, where x takes the value 2, that is let us look for those combinations where h occurs twice. So, this is 1 of them, this is 1 of them and this is 1 of them which means, X equal to 2 has a probability 3 by 8 and then X equal to 3, means all the 3 trials give you h only. So, we have got only 1 such combination, clearly the corresponding probability is 1 by 8. Now, we have to convert this information into probability distribution function and there after probability density function.

(Refer Slide Time: 04:22)



So, let us consider this X taking the value 0, it has got probability 1 by 8 and this will remain, so until and unless X becomes equal to 1. When X is 1, that time the probability is 3 by 8. So, what is the probability that X takes value either equal 1 or less than 1? That is 1 by 8 plus 3 by 8. So, there is a jump here by an amount 3 by 8.

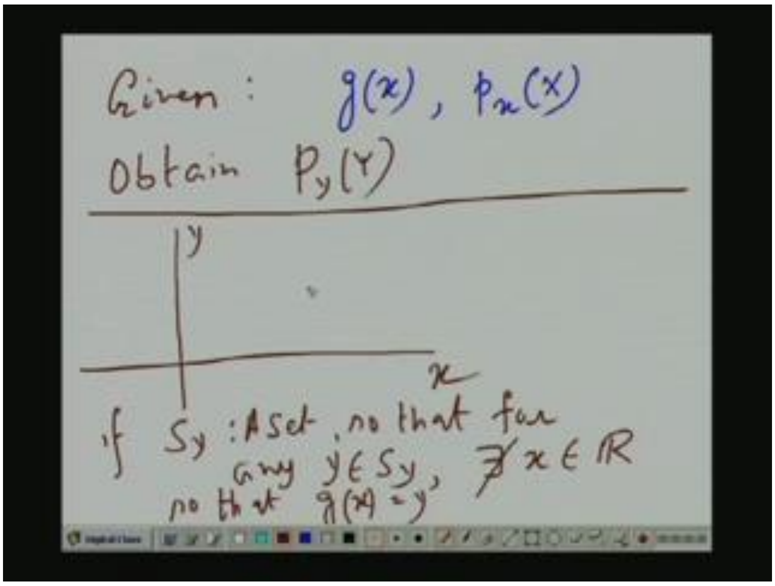
So; that means, this point is, if you add 1 by 8 plus 3 by I mean, with 3 by 8 you get 1 by 2. So, this is 1 by 2 and this remains, so till you have got this point X equal to 2, where X is equal to 2, what is the probability that X equal to 2? Again 3 by 8 which means, what is the probability that x takes the value less than or equal to 2; that means, we have to be jumped by another 3 by 8 here. So, that makes it 7 by 8. And this will remain, so till 3. Here there will be another jump by 1 by 8 and then it will saturate to 1.

So, this is the probability distribution function for X; obviously, the probability density function will be the derivative of this. Now, these flat regions will give derivative 0 and at these discontinuities, we will get impulses, each impulse will have a strength equal to the difference between the 2 levels. So, you have got 1 impulse here, another here another here and another here. This has got strength 3 by 8 3 by 8 1 by 8 1 by 8. This is not a very difficult problem and I guess all of you have been able to solve this nevertheless, in case somebody had doubt, 1 can just they can relook into this and then clarify his doubts.

So, let us now consider the next topic that is, in the previous class what we did is this just to recap, X was given to be a random variable and there was a function deterministic function g of x which was evaluated, every time the random variable X takes some value. For convenience, we all consider real value random variables and functions, which are real valued. That function was called y actually y equal to g of x. So, since x is random y also, is random probability distribution function for x was given, g x was given. Our task was to find out the probability distribution of y.

So, we consider various ways; we started with some particular y, we tried to find out the possibility or the probability of y taking values less than equal to that value, this we evaluated for various cases. So, this we know now. We have do a similar exercise for the probability density function. In fact, we have find out, we have to derive some clues from relation, where y giving the function g of x and giving the probability density of x. How to find out the probability density function for y?

(Refer Slide Time: 08:24)

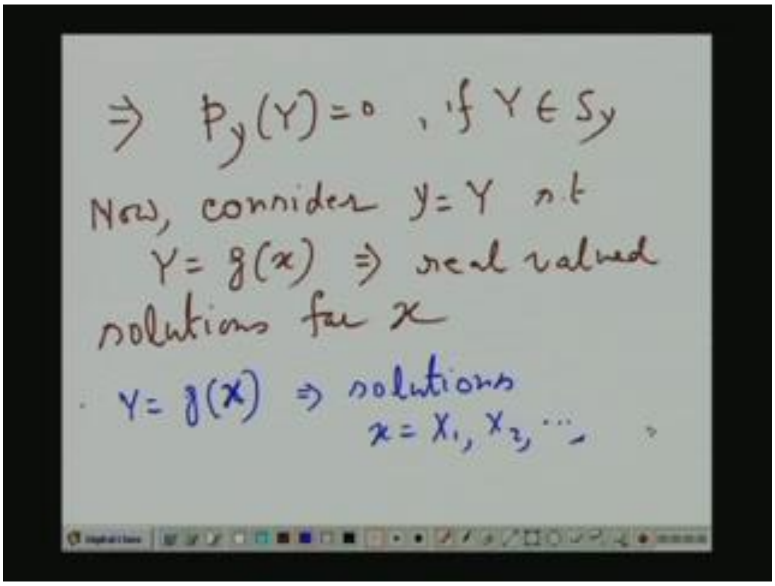


That is obtain now, in general as in our case, rather as I told you x takes real values and gives a real value output y. So, In fact, you can have a relation between like this y x and within a particular g x you can have the plot. Now, x takes real values from minus infinity to plus infinity that is given, that is x takes real values, all possible real values. Now, y has this axis. So, let us start this way there, suppose there are certain region on this axis, which are not tasked by this function g x, that is you put any value x, evaluate g x you can never get values in those regions. If such regions exist for instance, suppose g of x is such that its maximum values is 5.

So, naturally y greater than 5, that is if you point find out the point 5 here and look above along this axis, that entire region is something for which, I mean you cannot get a g x that is you cannot find any x for which g x will be equal to that. So, those actually give rise to improbable event, events which have probability 0. So, let me write down formally, that if set so that, for any y, element of Sy, there does not exist, this means there does not exist any x. Element of this real line this is for real line so that, g x equal to y. If that be, so that is there are some segments on this axis, together they form a set say Sy.

The set Sy is such that, you whatever x you take, evaluate the function g x that can never be a member of that segment or that set. If Sy be such a set then; obviously, for any y belonging to Sy, the corresponding probability will 0, because these are impossible events, because the you cannot get any x for which, g of x can give you a value for within that set. So, these are impossible events. So, for that probability will be 0.

(Refer Slide Time: 12:12)

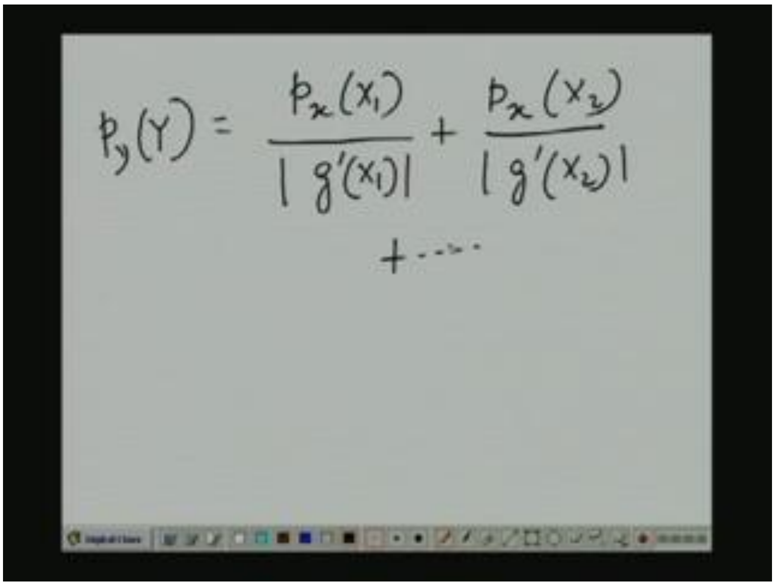


That I write in the next page. In that case, this value Y actually is taken from Sy. That is a trivial case. So, let us now consider only those segments or that part of this y axis, for which at least there exists at least 1 x, if not more on the real axis on the x axis. So, that g of that x gives rise to this value Y. So, these are not improbable events. So, consider, so that, if you really want to solve this equation. Then put this value you have chosen, this Y and find out for which x you get these, you will get real solution for x; that means, there exist some value, 1 value or many values, 1 the real axis for which the function, when evaluated on those values gives rise to Y. That is it leads to real valued solutions for Y for such case now. Earlier we consider such I mean, this Y such that, even if you try to solve it, you do not get any real value solution for x.

So, on the real axis, you cannot find the x for which g of x gives rise to this Y or this are improbable events, these are probability 0. Now, we are considering the other case, well instead of just getting a general result or maybe, I give the general result first and then instead of proving it for general case, it will be as it will be perfectly all right, we will just take a special kind of function, any be arbitrary function and show how the real comes, then you can always generalized that. I say that suppose I am solving this, these 2 solutions. Again for the time being you can assume that, you have got finite number of solutions, as if this is a polynomial of some finite degree. We have got the solutions, for a particular Y you are solving this equation, the solutions are given by x equal to say X dot dot dot dot. May be I will not take the number of solutions to be finite, let it be dot dot dot dot. I mean it does not matter whether you have got infinite solutions, these are the solutions.

So, g of X1 is Y, g of X2 is Y, g of X3 is y and like that.

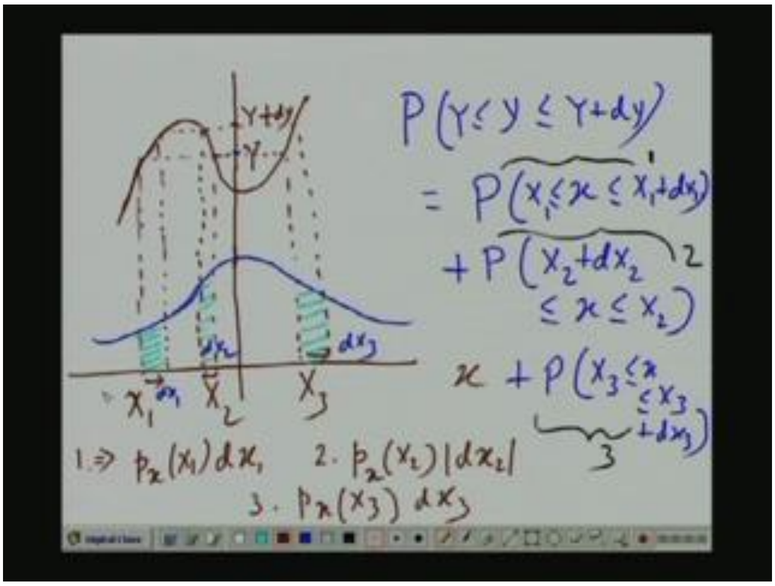
(Refer Slide Time: 16:14)



In that case, for a given Y the corresponding probability density will be, we put that 1 solution here and divided by the derivative of the function g x evaluated at this value and take its mod. Then, again px put the second solution X2 divided by the same derivative mod value of the derivative, but evaluated at X2 and dot dot dot dot. Do it for all the solutions. This will be the probability density.

Now, instead of taking the most general case, we just consider a particular form of g x, where we have got just 3 solutions X1, X2, X3. From that only, we will try to obtain this result. You know from that, we will try to obtain this result and that logic will be clear to you. So, that you can always I mean generalize that result to this context.

(Refer Slide Time: 17:35)



Now, suppose I just consider a function like this or may be, this is the function g x. This is the x axis and the corresponding probability density for x that also we plot. Now, suppose you are given a y, y here and y plus dy here. In fact, I just rewrite it from value Y Y and plus dy here. So, I have to find out what is the probability of y, taking values between Y and Y plus dy. That probability will be the probability density times dy, we have to evaluate that. But, let us see while you have this Y, it intersects this function g x, this is g x at 3 points. That means, it has got 3 solutions, you can then call them it is X1 X2 X3.

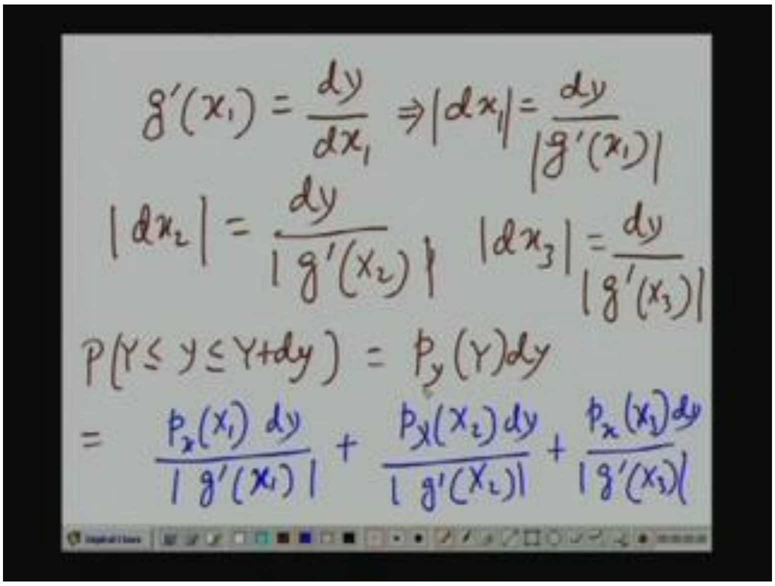
So, and as we go up, let us see what happens here. Here you have got this part. This part is dx 1, but here this part, this is dx 2 and here this is dx 3. You can see dx 1 is positive, this is a positive slope. So, it is positive where as y can increase only if x increases. Here also it is positive, because since x increases this increases, because the slope is positive. But here the slope is negative, that is why dx 2 is in this direction negative.

So, what is the probability of y the remaining between Y and Y plus dy? That is same the probability of X remaining between X 1, that is let me write down this way. Probability of y, this will be what either X should be here or here or here. In fact, this area, this area and this area, summation of the 3 areas will be the probability of this; will be the value of this. That is, this will be this time X2 plus dx 2, that is this point X 2 plus the this point, but that is to the left plus this is the probability.

Now, first you evaluate this what that will be? Let us consider term 1, this is term 1, this is term 2 and this is 3. Number 1 will be what very simple, number 1 is probability density at this point times dx 1. Number 2: number 2, will be what? Probability density at X2, but not times dx 2, because dx 2 is negative, I am only interested in this area. So, as for the definition, this is the probability density here, multiplied by this time mod dx 2, because probably cannot be negative and number 3 is very simple px X3, no need to put a mod, it does not make any difference though, dx 3.

Now, you can always put a mod across dx 1 also dx 3 also, because even otherwise they are positive, so does not matter. So, to constitute the, to construct the general result, I will put a mod across them anyway. So, this probability will be what? This plus, this plus this. This is still not what not equal to the result we quoted, but it is not very difficult to obtain that from here. Let's find out what is dx 1 before I go to the next page, because that time I will not this figure with me what is dx 1? dx 1 will be, now I know that if you take the derivative here; g prime x at X1. I call it g prime X1 that is nothing but dy dx 1, which means dx 1 will be dy by g prime x 1 likewise.

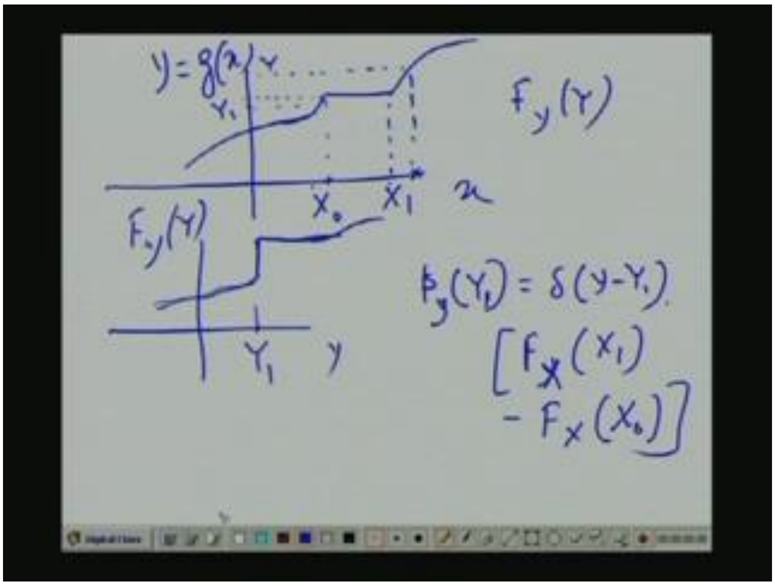
(Refer Slide Time: 26:52)



That means, which means; if you want to put a mod across this, you have to put a mod across this, but dy is always positive, dy was in our control. We took y equal to some value Y and then went up Y plus dy. So, dy is positive. So, you need to put mod only here, because varied can be both positive and negative. Similarly, g you can write down.

So, now, if you put them back, we get the probability of y remaining in this range, which by definition is nothing but dy that is equal to this into dx 1 mod, but that is equal to this. So, dy Px again dy. So, you take out dy from both sides and you will get the result.

(Refer Slide Time: 30:05)



So, for so good, just only 1 thing I do I need to add here, that suppose the function y equal to g x is such that, it has got a that it has got some flat potion, something like this. As we discuss in the previous class, there will be a discontinuity due to this, with the probability distribution function. Just for convenience you can even take a recap of that. Suppose, you take your Y to be here, then your Fy Y what will that be? The probability of x remaining less than or equal to this value. Then, slowly suppose, Y is brought down and this point also moves to the left this decreases; obviously, this is a non decreasing function, because it is an increasing function most of it.

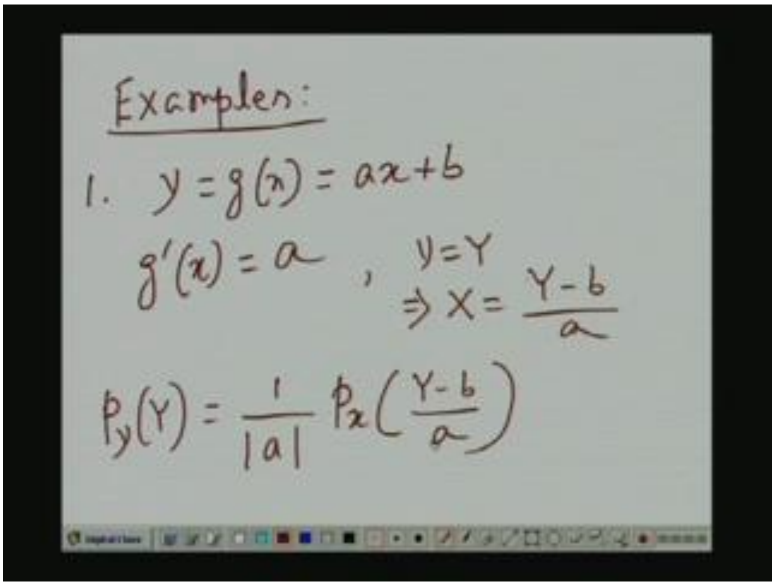
So, this will decrease smoothly, as this starts falling and this goes to this left point, this starts falling down smoothly, gradually smoothly, continuously. After it reaches here that is, when I am here. So, this is particular value Y1. Then, I have got some value for Fy Y 1 also that is, the total probability of X taking values, either equal to or less than X1 that is fine. But as I try to slide down, immediately I cross this distance and I move here. So, there is a huge jump, because I am crossing this distance. Earlier when I was here at Y1, I was crossing the total probability of x taking values up to X1. Immediately after that I am going up to X0 or even to the left of X0 and progressively far and far to left of X0.

So, there is this jump, jump by how much? That is probability distribution of X up to X1 minus probability distribution of X up to X0 that is a total probability of X lying within this range by the time jumping, so that there is a discontinuity. So, if such there is such discontinuity in the probability distribution function, if there is such discontinuity in the function g x; obviously, if you are looking at Y1 if presented Y here Y1, the corresponding derivative. So, there infinitely may be many solutions. Earlier we took y equal to Y; we had 3 solutions X1 X2 X3.

Suppose, function gx is like this and you choose this Y1 this level only. So, there are infinitely many solutions, in this interval X0 to X1 and for each of them you know this derivative is 0. So, there will be division by 0, that result is not valued. You have to for such cases; you have to look at it from different angle. Since, the probability distribution function at those points will be, I mean if you plot Fy versus as Y, there will be some discontinuity at Y1. It is going up, it suddenly it is going up I mean, it is jumping and going up, settled into 1.

Since, probability density function is nothing but derivative for these with respect to Y. At this point there will be an impulse that is in this case, Py Y, I would say Y1 will be nothing but at impulse at Y that is delta y minus Y1 means; there is an impulse here at Y equal to Y1 and of what strength? Strength is this which is nothing but what is the by amount by which there we have the discontinuity? What is the amount of discontinuity? That is probability distribution up to X1 minus probability distribution up to X0 which means; equivalently that, probability of X lie within this range. So, that is this is what a question.

(Refer Slide Time: 35:22)

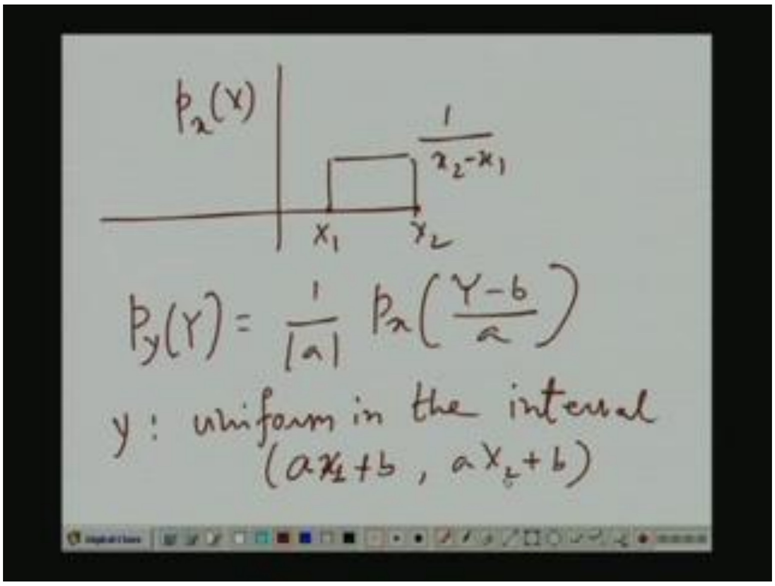


Now, we take some examples. 1: y equal to g x equal to ax plus b. In this case as you know, it is a linear function for giving any particular y equal to Y, there is an unique solution only. Some X and that will be giving by Y minus b by a. This is unique solution. g prime x is a and given y if y equal to Y, it give rise to a solution X is equal to from this equation Y minus b by a. This is the only solution.

So, you get straight away apply the formula, you get this 1 by mod of this, at this point X, but then this is independent is constant, this is independent of X. So, just put a mod a here a and

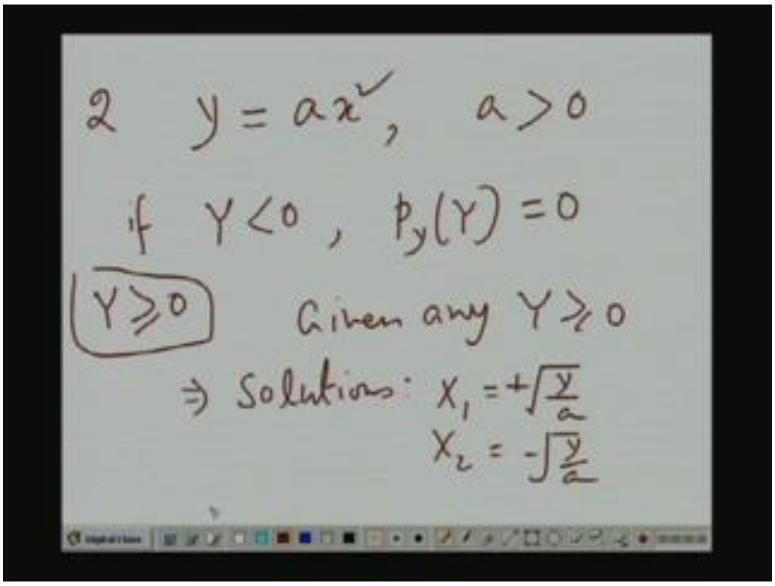
Px what value? Replace x by this express the solution in terms of given Y, put that here. As for example, suppose this x is a random variable now, it is not taking the entire real axis, it is within some range. It is a uniform random variable within the range X1 to X2.

(Refer Slide Time: 37:31)



That is suppose, px X is like that, what is the height? Areas would be equal to 1; this is a rectangle once this this side is X2 minus X1. So, height will be 1 by X2 minus X1. Suppose this is the case. So, X is not taking the entire real axis, X is taking only values within X1 to X2 and within this range, it is probably density is fixed. In this case Py Y by that formula 1 by mod a Px Y minus b by a. Since this is uniform and this is just a constant, this also uniform, this is also uniform and in what interval? If x takes the value X1, y takes the value Y1 equal to how much? a X1 plus b and if x takes the value X2, y takes the value a X2 plus b, Py and Px they have the same form which is just a constant. So, these are thing, it is a equal y also will be uniform in the interval. So, these are simple example.

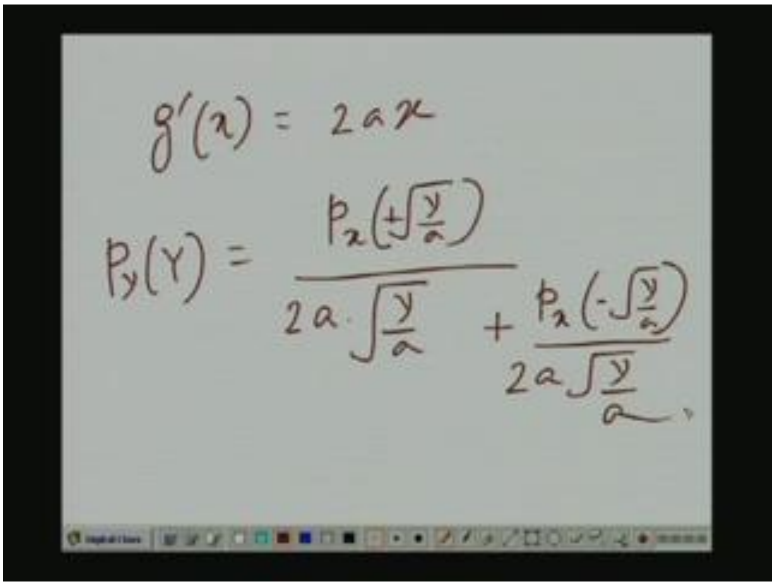
(Refer Slide Time: 39:47)



Next we consider slightly more complicated case. y is equal to ax square. So, this is your g x, I am giving that a is a real constant and positive. First x is of course, as I as I told you in the beginning, we are taking real valued x only. So, x square whether x is negative or positive, this is always positive, a is positive. So, if on the y axis, we are considering the negative range of y, then that amounts to an impossible event, because there cannot be any real x, for which y by this formula will become negative.

So, if Y is less than 0; obviously, Py Y is 0. So, we will not consider this, this is a very trivial case. So, now, we consider Y greater than equal to 0 this case. Here for giving any y, we have got solutions, how many solutions from this? This is a quadratic equation; we have got 2 solutions with the plus sign and X2 with say minus sign. However, we could call this X1 and X2, because it might figure X2 was greater than X1; it does not matter though these are the 2 solutions.

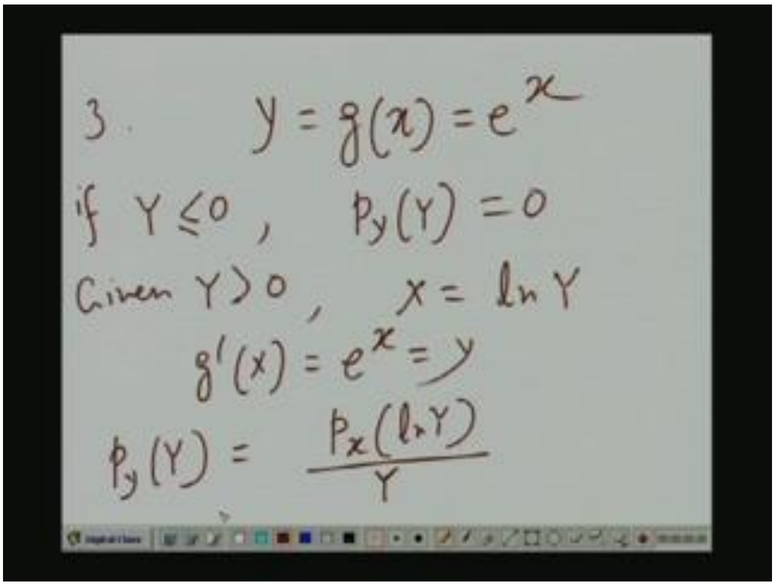
(Refer Slide Time: 41:49)



So; that means, and what is g prime x? g x was ax square. So, g prime x will be twice ax

So; that means, given any Y, this will be what? They have got 2 solutions, so there may be 2 terms. 1 term is; put the first root with the plus sign divided by twice a into this, and then since we are putting mod even if I have minus sign, the mod of operation will take care of it say. In this case, you cannot say that if this is uniform x is uniform this also will be uniform, there are 2 terms. If you plot these probability density plot, this probably density at superimpose, they may not be uniform, because you will not get uniform in a.

(Refer Slide Time: 43:39)



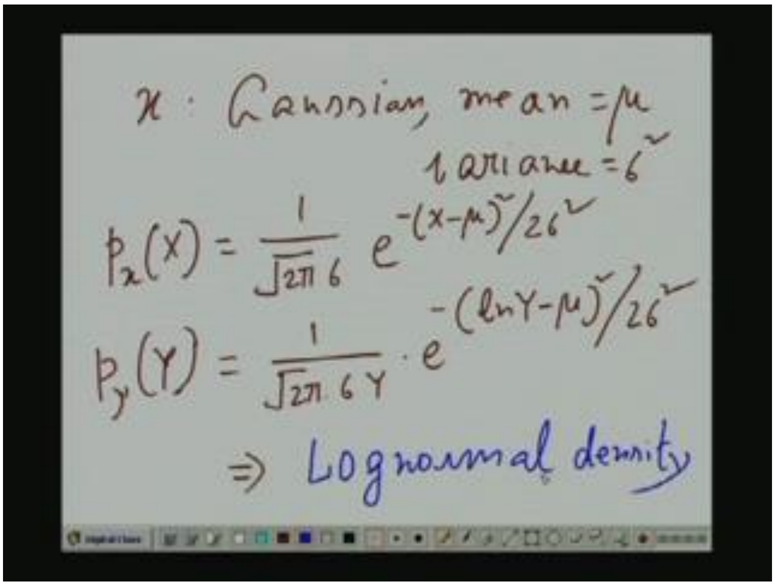
We can see the next example now. Clearly for any x negative or positive, this is the e to the power x is a positive number. So that means, if Y, in fact this can never be, if you are not considering minus infinity this can never be 0 also. Anyway if Y less than 0; obviously, PyY is 0, because that is an impossible event, there is no x for which you can get e to the power x equal to some negative number.

Then given Y greater than equal to 0 what is X? Solution is ln Y is number 1 and what is g prime x? That is e to the power x which is Y itself. In this case and this is only 1 solution, given Y there is only 1 solution by the way, it is not a periodic function or a function of such kind, where you know you have got multiple solutions possible is no. Given up Y you will get only 1 x.

So, it is very simple case Py Py Y this will be what? Px you have to put the solution. Now solution is this; ln Y divides by mod of this. In fact, I would say rather that, of this equal to what and put this. Otherwise no there may be a division by 0 according here

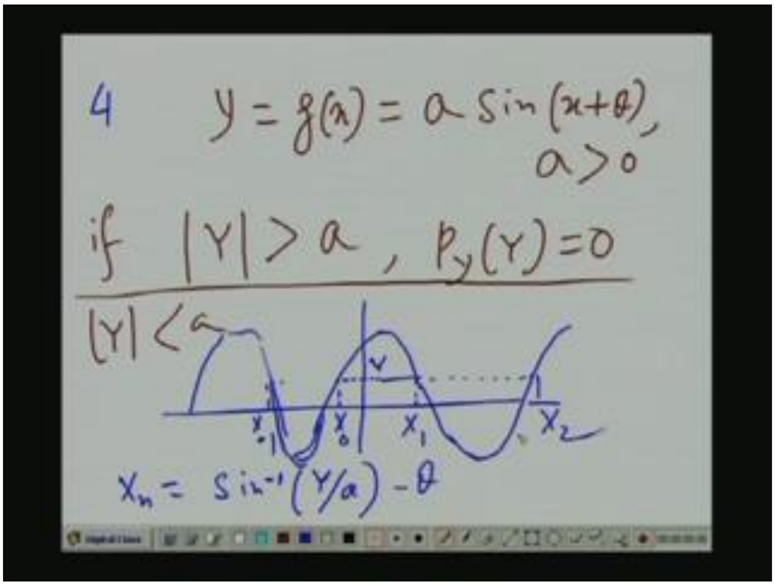
So, either way mod of g prime x that is, mod of Y Y, but Y is already positive. So, we will have this thing, this is a probability density.

(Refer Slide Time: 46:38)



As for example, suppose x is given to be Gaussian of mu that is, say x is given to be normal or equivalently Gaussian, both the terms are used, mean that is this we have already covered in the class, this particular form. In that case Py Y will be what? By our previous formula, there will be 1; there will be 1 Y here. We have to replace this by the solution that is ln Y, rest is same. So, this becomes root 2 pi sigma Y e to the power minus ln Y minus mu whole square by twice sigma square. This is called a log normal density function.

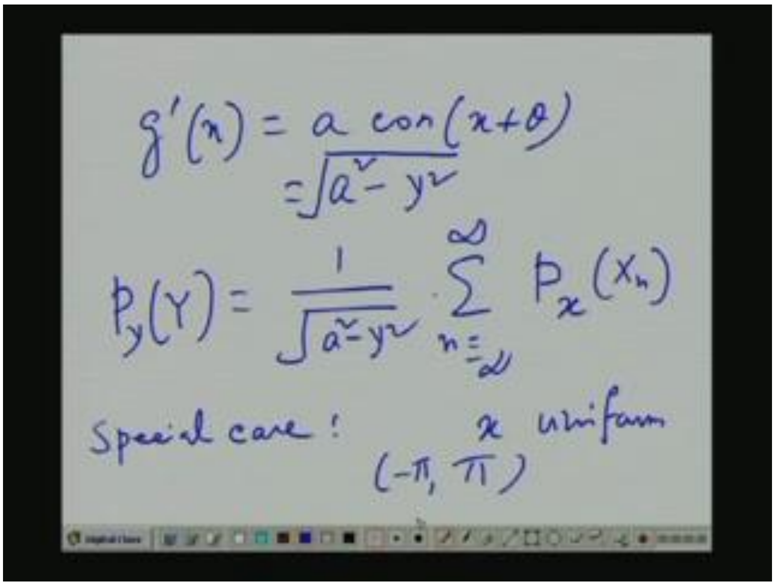
(Refer Slide Time: 48:43)



Next example, again this is slightly more complicated. Suppose, Y is g x is given to be something like this; a sin x plus some constant theta, where a is real, theta is real a is greater than 0. So, this function Y; obviously, it cannot take values beyond plus a or minus a. Y cannot take values beyond plus a and minus a; that means, if Y is less than there is a mod value of this is greater than a, if the mod value of this is greater than a in that case; obviously, Py Y is 0.

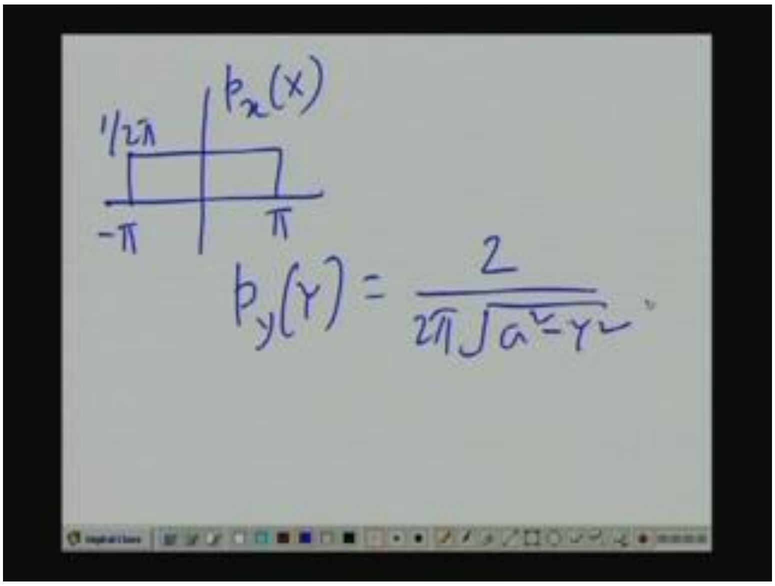
On the other hand, so we will be considering Y mod Y less than a. In this case, if you put Y here, since this is periodic, you will have infinitely solutions like this. Suppose, your Y is he, this is your Y. We will have 1 value here, another here, you can call it X naught now, this is X1, this is X minus 1, this is X2 likewise. So, we have got X0 X1 X2 X minus 1 X minus 2 and like that, so infinite solutions. So, you have to use that formula for each of these solutions. What are the solutions? In general Xn is nothing but sin inverse of what? Y by a minus theta. This sin inverse is an angle, this at this periodic is it does not have only unique value. You can have value various values; that minus theta will give you this. So, there are the various solutions.

(Refer Slide Time: 52:13)



And what is g prime x? g prime x will be a cosine x plus theta which is nothing but what? Very simple y is a sin x plus theta, this is a cos x plus theta. So, what is this? If you square them up and add you will get a square, which means; this is nothing but a square minus y square. In this case what is Py Y? Is nothing but this is common for all x, for any solution. You have to put various solutions for x, but for any solutions this is common and this is independent of the solutions. So, this is a common factor and then summation n from you can say minus infinity to infinity Px Xn. Special case; x is uniform between minus pi and pi. So; that means, it will have a probability like this.

(Refer Slide Time: 54:26)



That is 1 by 2 pi, so that the area is 1. So, this is independent of x and quite clearly, if x is taking values only between minus pi to pi and in no other range, then I can have only 2 solutions in minus pi to pi, from our previous figure. You can call 1 X0 another X1. But for both of them the corresponding probability is 1 by 2 pi. So, in that case Py Y is very simple, this is just 2 times whether summing twice 2 pi 1 by 2 pi again 1 by 2 pi and this factor a square minus y square is common. So, that is all for today and we will take up from here in the next class.

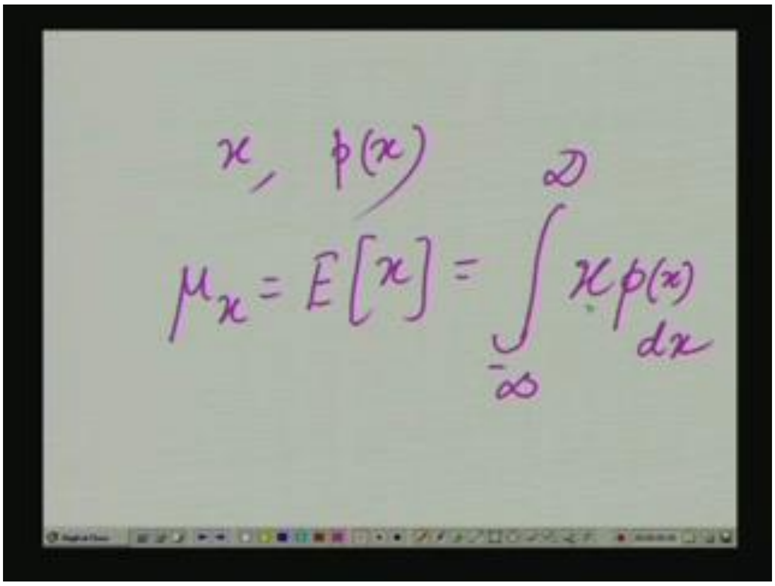
Thank you very much.

(Refer Slide Time: 55:27)

(Refer Slide Time: 55:36)

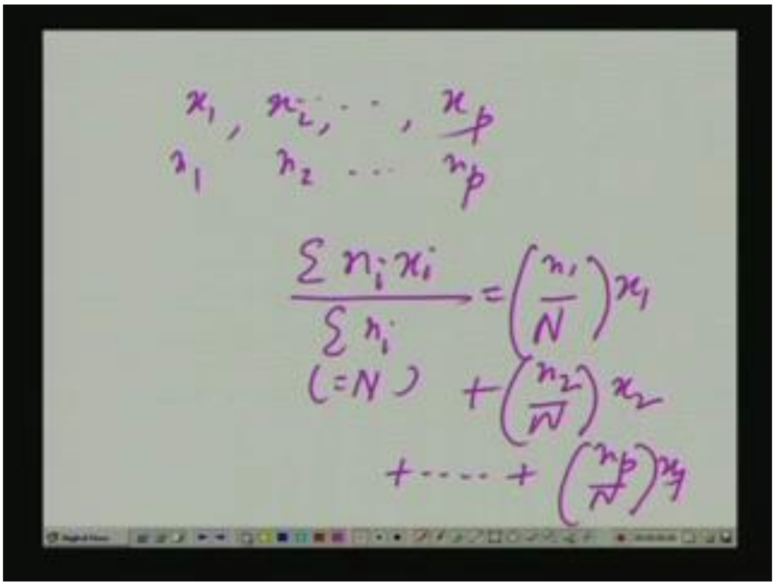
So, today we discuss some important concepts like mean and variance.

(Refer Slide Time: 56:00)



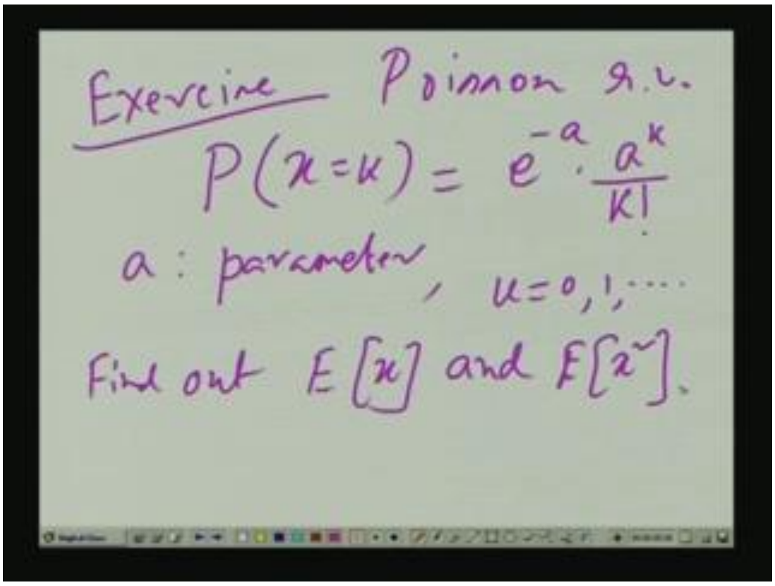
We are given a random variable x and its probability density function is px, x for the time being is continuous random variable. Then we define its mean mu x as, also expected value it is an operator. E actually means, expectation it is an operator, it is just a question of notation expected value of x as. Now, you can I mean how this formula comes, maybe we can give a frequency interpretation.

(Refer Slide Time: 56:48)



Suppose, in the discrete case suppose, I mean x takes some values like you know x1 or x2 up to say xp, this is n 1 times I mean if you observe x, may be on n 1 location you get x 1, on n 2 location you get x 2 dot dot dot dot, on np location you get xp. So, what is the arithmetic average? Net value is summation ni xi divided by summation ni, this is the total number of trials. If you can you can call this as N, in that case what you have is this; ni by then equal of today.

(Refer Slide Time: 57:41)



Suppose, I consider, so this is an exercise. I consider a Poisson random variable; Poisson distributor random variable, where the probability of x taking a value K is given by let me see, e to the power minus a a to the power K divided by factorial K. a is a parameter and K can be 0 1 dot dot dot dot like that, for this find out E x. So, that is all for today. In the next class we will take it up.

Thank you very much.