Estimation of Signals and Systems

Prof. S. Mukhopadhyay
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

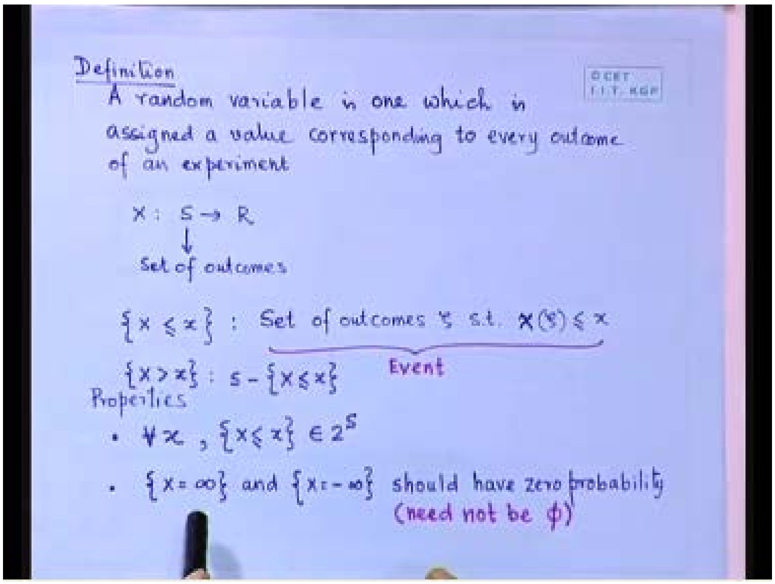
Lecture - 03
Random Variables

Yeah good morning, so in this lecture we are going to talk about random variables. Now as usual we will have to first learn what is a random variable. Actually it so happens that suppose you have a you have an experiment right, so we have seen in the last class that if we have an experiment then we can assign... it will have certain outcomes and we can assign probabilities to its various outcomes, we can even define events which are a subset of all the possible outcomes and we can assign probability to them.

Now it is not always necessary that we will have to assign probabilities only to outcomes rather we can start assigning probabilities to functions of outcomes. For example, we can say that suppose you cast a die and now if the cast if you cast a die then the probability then what are the outcomes outcomes are the 1 2 3 4 5 or 6 and you can say that whatever the phase whatever phase comes up I will give you I will give you 10 rupees for each mark. So if you if a if a if a 4 comes up I will give you 40 rupees. Now how many rupees I am giving you now becomes a function of the outcome, is it not?

Now I can now once I define this function now I can start talking about what is the probability of you are getting a certain amount of money. So we can always... we need not always directly speak in terms of the outcomes but in general we should speak in terms of functions of outcomes so that is the basis of having a random variable.

(Refer Slide Time: 00:03:12 min)



So random variable is nothing but it is a variable which is assigned a value corresponding to every outcome. So, for every outcome you are assigning it a value which is the real number. So it is basically a mapping from the set of possible outcomes to real, right?

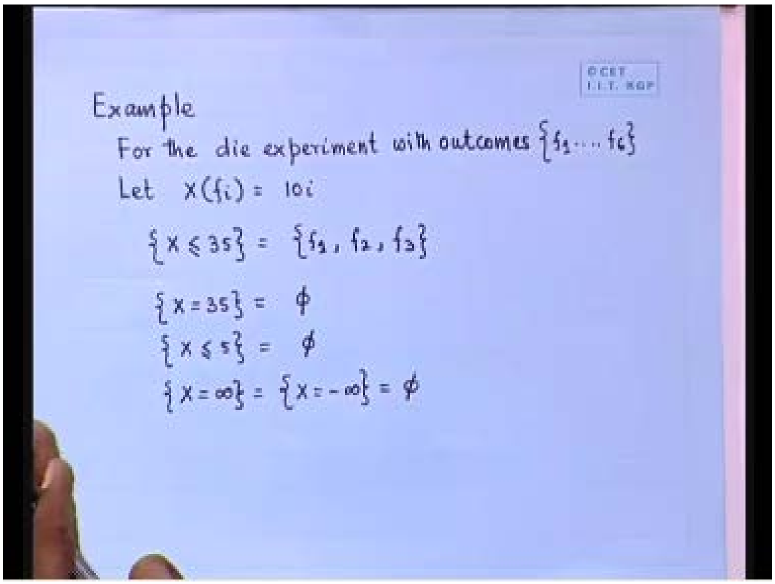
So now once you... so so now you have a variable which will get values corresponding to every outcome. Now once you have that then then we can talk about then we can make statements about this random variable and then talk about their probabilities. So we can say that what is the probability that you will get less than 35 rupees, you can say make such things. So what is the probability that this random variable's value is less than or equal to small x. So what does it mean (Refer Slide Time: 4:20) this means that this particular notation that is within curly brackets, random variable and then a then a then an expression concerning random variables it means it is the set of outcomes of the experiment of the of the experiment that we have such that x of psi according to this rule... because according to this rule corresponding to every psi belonging to s you will get a number which is x of psi. So what is so it is all these psi's like that X of psi is less than or equal to x. So so this denotes a set of events or a set of outcomes.

Similarly obviously we can say what is what is the... which are the set of outcomes that X is greater than x, you can make many such statements, that will be obviously the total set minus X less than or equal to x. so we can make various such statements.

Now this mapping that we are talking about (Refer Slide Time: 5:27) this mapping should obey in general certain certain mild properties. So the so the property is such that first of all all statements of this kind should belong to should be an event; that is for all values of x for all values of small x the the set of events that will correspond to this should be an event. It may be an it maybe an impossible event, this set may be an all set or it may be the full set s or it may be any other set but it should be a set in general. This is a mild condition because when you are mapping the set of events onto a real number real 9 etc there are there are there are... we have noted that there are some peculiar combinations to which you cannot assign a probability and it will not be a be a valid event.

So, valid events are only sort of you know countable unions and dissections of intervals on real line, these are some technical mathematical thing so so that is why this is needed and it should so happen... you should assign this function in such a manner that the value of X at infinity and the value of X at minus infinity should have zero probability. This is the...these are the property that you should obey, these are these are very very mild properties you can you I mean most of the functions will actually obey them to assign this. So so this is a random... basically simply speaking a random variable is a function of the outcomes of an experiment okay. So having said that we can see an example.

(Refer Slide Time: 00:07:19 min)

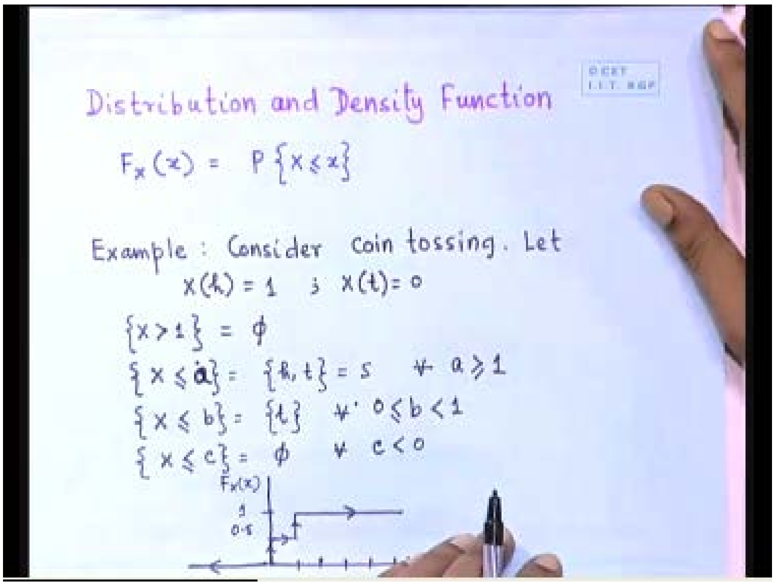


For example for the die experiment the outcomes are f 1 to f 6 and as as I said that let X f i is equal to 10 i. So I am saying that I will give you 10 rupees corresponding to every point that comes up on the die. So then what is the what is the set what are the set of events such that X is less than or equal to 35; obviously it is f 1, f 2 or f 3 because for f 1 X f 1 is 10, for f 2 X f 2 is 20 and this is 30. So the set of events for which this is satisfied are f 1, f 2 and f 3, correct?

If I say what is the what is the set of events that X is equal to 35 it is obviously the null because there is none of them will give me exactly 35 rupees so it is also an event of it is also a valid event, it is it is the impossibility event but it is a but it is a valid event. Similarly X less than or equal to phi will also give you the null event and just to see that that property satisfied X equal to infinity and X is equal to minus infinity are both phi so that property is satisfied. So this is a varied random variable.

Now now we start talking about... so far we have just talked about the set of events, now we have to... our our main interest is a probabilistic characterization of the of this variable. So now we have to start talking about the probability that capital X is less than or equal to because because that is our main interest. So now we talk about start talking about... once we have defined a variable we can define what is known as distribution and density functions.

(Refer Slide Time: 00:09:13 min)



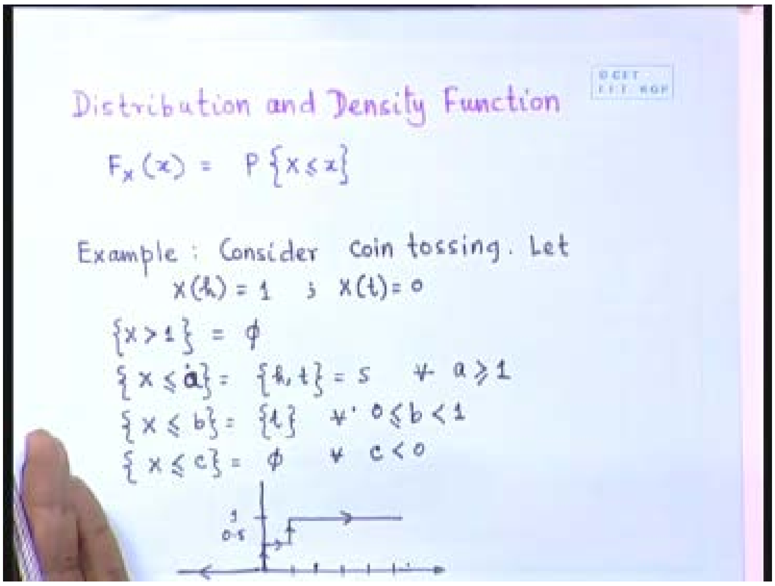
So this is a distribution function density to which we will come later. So what is a so the distribution function actually characterizes the probability that X is less than equal to x. This without this it is the set of outcomes as we have noted. So whenever you have a set of outcomes you have a corresponding set of you can assign a probability to that set of outcome, basically this becomes this is like an event because it is a set of outcomes, all sets of outcomes are events. So now we can talk about that the probability of the event okay. So only you are talking in terms of X.

So the function which characterizes the probability of the events for which capital X is less than or equal to X is called the distribution function of this, it is a distribution function of the random variable X and this value for the value small x is this (Refer Slide Time: 10:16). Sometimes we will we will drop this subscript X whenever it does not cause confusion. But in general you can talk about F subscript capital X of w which will mean that capital X less than or equal to w, so they need not the same variables but when we are using same variables we sometimes will drop this subscript.

So, for example or simplest experiment is coin tossing. So for coin tossing what is the distribution function? So suppose we say first of all coin tossing is an experiment, we know what are its outcomes; suppose its outcome is that its outcomes are head and tail, now suppose I have I have defined a random variable such that X of head is 1... first of all we have to define the random variable that is define the mapping of these outcomes to the real line. So suppose I have defined this mapping that that X of head is 1 and X of t is 0 this is my null, it will satisfy all probability all properties of a random variable. So now we can again make statements.

For example, we can say that X X is greater than 1 this is the null set phi, it can never be greater than 1. Similarly X is less than or equal to a what is the set it is h and t if for all a greater than or equal to 1 because X is always going to be less than or equal to 1 because the only values of X it takes are 1 and 0 so it is always going to be less than or equal to 1. So the set of events are h and t that is a whole set so its so its probability is going to be 1.

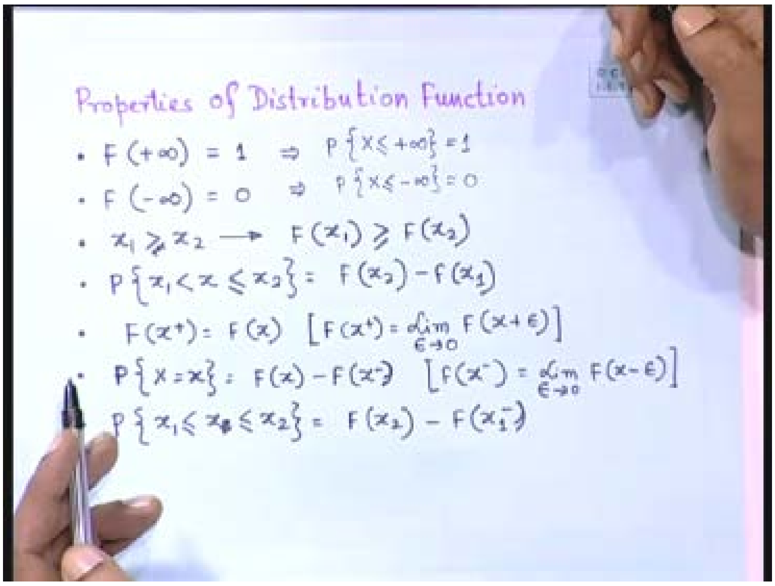
(Refer Slide Time: 12:16)



Similarly if X X is less than or equal to b when b lies b is less than 1 and greater than or equal to 0 is the is t, only for this tail outcome the random variable the value of the random variable satisfies this, correct? So the event corresponding to this is this is the tail and if c is less than 0 then this is again phi. So now I can say that what is going to be my distribution function with respect to X. So the moment c goes less than 0 f x is going to be 0 because because it is a phi event then the probability of the phi event is 0; for from 0 to 1 it is going to be this t so so if we if we assume a fair die that is the probability of head is 0.5 and probability of tail is 0.5 then it will be 0.5, if it is an unfair die then we can assume any other value that is if Suppose the probability of head is p and the probability of tail if 1 minus p in which case this will become 1 minus p and then at 1 it will rise. So this is the this is the probability distribution function (Refer Slide Time: 13:51) of this experiment under this of this random variable, right?

[Conversation between student and professor: 14:01 – 14:25] Yeah sir sir X is less than a for X less than or equal to a one head and tails... no no no if you leave the h... no no no X is less than or equal to a and a is greater than or equal to 1what does it what does it mean that basically let let a be 1.5 so x is less than or equal to 1.5 this is always satisfied because X can be either 0 or 1 so the set of events which will satisfy this is h and t.

(Refer Slide Time: 15:54)



Now suppose X less than or equal to b and b is less than 1 suppose it is 0.9, if you say X less than or equal to 0.9 that is only satisfied by the event t because X of t is 0 so so the probability will be the probability of the tail event which may be as you define it it maybe 1 minus t, correct? Understood? Okay.

So now what are the properties of the distribution function?

The properties of distribution function should be first of all that f of plus infinity should be 1 this is natural because what is f of plus infinity, it is the probability that X is less than or equal to plus infinity. So obviously it is 1 where X is naturally going to be less than or equal to plus infinity. It will... whatever real value it takes it is going to be less than or equal to plus infinity.

Similarly, f of minus infinity will be equal to 0 so this is also natural. This means... and f should be a non-increasing monotone rather non-decreasing monotone not not non-increasing non-decreasing monotone which means that if X 1 is greater than X 2 then F x 1 will be greater than or equal to x 2 F x 2 right. What does it why why greater than or equal to is because suppose because the probability function especially for discrete variables... we will see what discrete variables are; like the like the coin tossing experiment what did you see that the probability density function rises its steps. So suppose this is X 1 and this is X 2 then F x 1 is equal to x 2 here is equal to F x 2.

And on the other hand, if you have this as x 1 and this as x 2 then rather the opposite this is x 1 and this is x 2 then you have F x 1 greater than F x 2 and these are the only two cases. So, if you have discrete jumps then x 1 greater than x 2 will imply that f x 1 is greater than or equal to f x 2 so it either increases or stays at where it is, it does not decrease that is why it is called a non-decreasing monotone, monotone means continuously rising right. So right so this property should be satisfied.

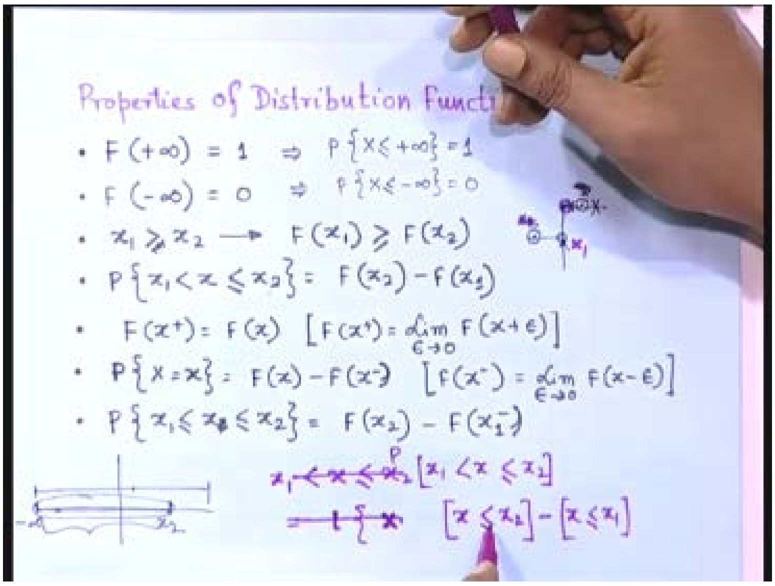
Then [conversation between student and professor: 17:45 – 17:48] Sir what what variable we have to what x 1 and x 2 variable... no it does not depend on that and their probabilities... no it not no it does not depend on that because as I am as I am moving X on the axis I am actually increasing the set of outcomes all the time so this so the probability cannot decrease.

See what is the range I am considering I am considering from minus infinity to suppose x 2 here and and x 1 is greater than x 2 so first of all first time I am considering this set, second time I am considering this set so I am always covering this so the set of outcomes corresponding to this are are anyway covered in this and I possibly I am including more outcomes so the probability will increase okay?

Now then it says that probability of x 1 if x is less than x 1 or less than or equal to x 2 what does it mean? Remember that we are having less than x 1 and less than or equal to x 2 this is to take care of discrete junks. If the probability distribution function was absolutely continuous this is a discontinuous function okay, it takes a jump exactly at some point. So x is greater than x 1 means that if if this is the... if this is the x 1 point then x greater than x 1 and less than or equal to x 2 means that it is just it just starts from after x 1 okay and is less than or equal to x 2 that will be F x 2 minus F x 1; why? Because x 1 this can be this interval can be written as the interval that is is is equivalent to the... we will not use this notation; suppose let us take the okay let us take the interval x 1 x 2 this is the this can be written as the interval x less than or equal to x 2 minus this is the interval corresponding to this (Refer Slide Time: 20:27) less than or equal to x 1.

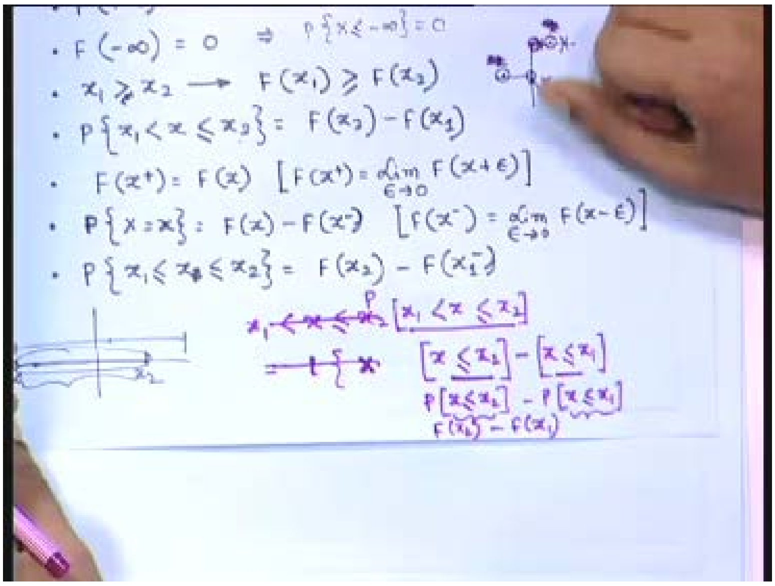
If we subtract from this interval this you will get this and and because you have less than or equal to so the point x 1 is actually included in this interval. So when you subtract from here the point x 1 will be excluded from that interval therefore it is less than and not less than or equal to 0. So so obviously the probability of this is going to be... since these are these are disjoint sets, no these are not disjoint sets but this this probability okay so this is going to be exactly the probability of this is the set subtraction so for for set union you have to take you have to disjoint but for set subtraction need not be. So this is simply because this is this is some some outcomes, this is some outcomes so you have to you have to subtract so so this become F of x 2 minus F of x 1.

(Refer Slide Time: 21:07)



[Conversation between Student and Professor – Not audible ((00:21:42 min))] Yes yes yes sorry actually actually I should write in this write in the centre. So f so so this is how it comes because this interval (Refer Slide Time: 21:56) I am now breaking up as a... I am subtracting this from this interval this and so the probabilities will become F x 2 minus F x 1 right.

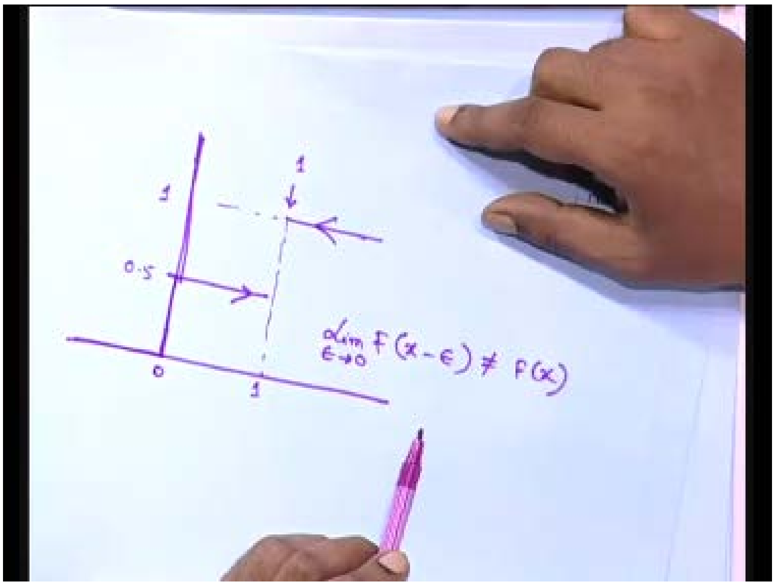
(Refer Slide Time: 22:04)



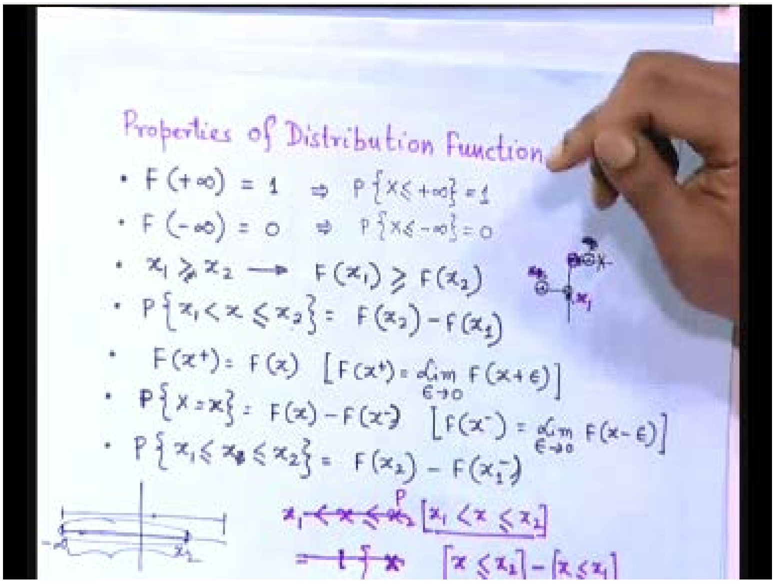
Now it it turns out that F this steady this distribution function is a right continuous function that means that the right hand limit will tend to the function that is suppose for example for the for that for the coin tossing experiment this is the exact point (Refer Slide Time: 22:34) let us say 1, see this is the coin tossing experiment, this is 0 so here it is 0 we have drawn and this is 1, this is 0.5 we have just plotted it, this is 1. So just at 1 it takes a jump so till this this comes to the limit it can be 0.99999 any number of 9s and still the probability will be 0.5 and the moment it will reach 1 it will go up to 1.

So the left hand limit that is F x limit of x minus epsilon epsilon tending to 0 is not equal to F x. Because however small you make this epsilon the probability of x minus epsilon is going to be 0.5 whatever epsilon you make as long as epsilon is nonzero but the moment epsilon becomes exactly 0 that is it becomes F x it jumps this is this is because you were considering discrete jumps and the function is discontinued. But form the right side F x plus epsilon as epsilon tends to 0 will be will come to the function F x that is why it is called a right right right hand continuous function.

(Refer Slide Time: 24:05)

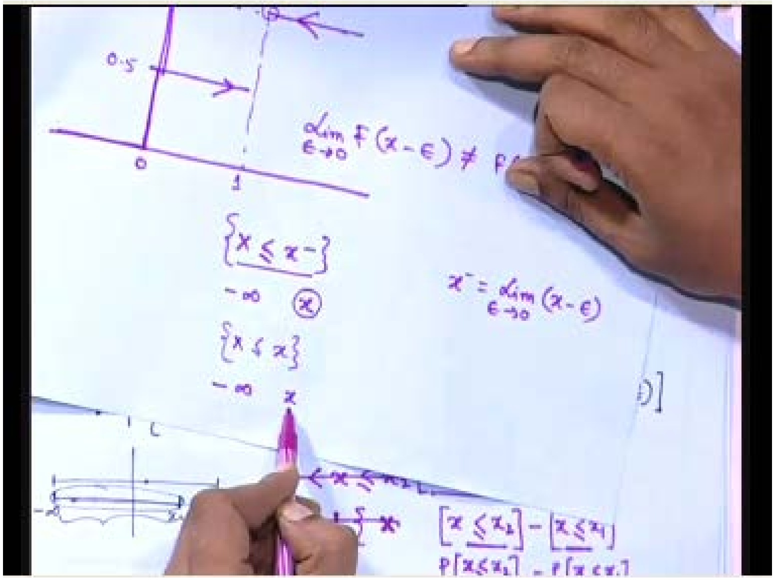


(Refer Slide Time: 00:24:16 min)



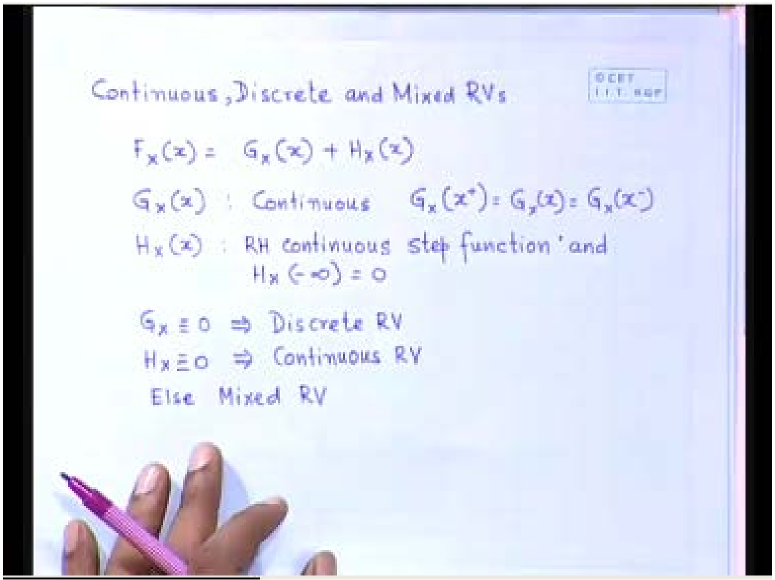
So now now obviously what is probability that x is equal to x? it is F x minus F x minus because F x minus I mean F x minus is the set the set X less than or equal to x minus where x minus is equal to limit of epsilon tending to 0 x minus epsilon is the set of all points from minus infinity to x but leaving out the point x, just that point if you take out all other points are infinity and if you put X here these are simple limit concepts. So if you if you put x here then you are the you are considering minus infinity to x including the point x. so if you subtract this from this then you are left only with the point x.

(Refer Slide Time: 00:25:17 min)



So that is why probability that X is equal to x is is F x minus F x minus and so on. These are simply derivable properties and they are easy to follow.

(Refer Slide Time: 00:25:43 min)



Now you may have in general you may have a continuous random variable, a discrete ramdon variable and a mixed random variable. For example, if you are measuring temperature of this room on several different days then that variable will be a continuous random variable so it can take any value. If you are measuring if you are having a coin tossing experiment then you are having your fully fully discrete variable, if you are having... now you can in general you can have a you can have an experiment where it is a mixed random variable.

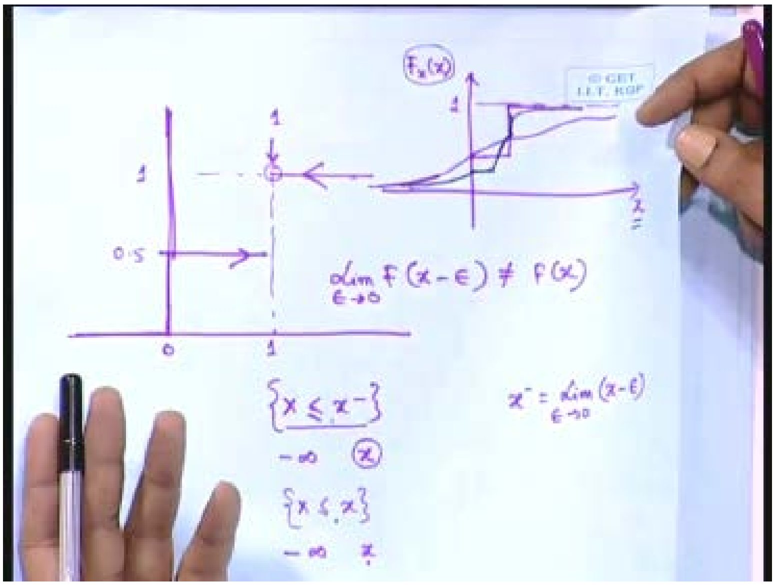
Now, for example, you can you can suppose you have three different thermometers and each thermometer has a has some fixed bias and you you are using one thermometer you are you are you are picking up a thermometer and you are coming to measure the temperature of this room. So this picking of the thermometer will be will will constitute the discrete part of the experiment because there is only three thermometers there are discrete set of outcomes and then you are measuring a temperature so this thermometer temperature combination it will again be a random variable that will be that will have mixed properties right. So you can have... this is often an example I have given.

So you can in general have that is distribution functions may look in different ways. for example, they may they may look like this, they may look... see they have to go to distribution functions they have to start from 0 because F minus infinity is 0 and they have to eventually finish in 1 because F plus infinity is 1 and it is non non-decreasing. So therefore distribution functions generally look like this. So this is it till gradually as you increase x this will approach 1 and it will it will less early have to start from 0. So this is a continuous function. If if we have as we have seen if we have like this this is a purely discrete one.

We can also have a we can have it like this that suppose it is we can have a form which is like a three different continuous segments. in general distribution functions can be piecewise continuous that is there will be a finite number of step jumps corresponding to the discrete part of the experiment and in between those step jumps there will be continuous variation. So it is a it is general in general it is a piecewise continuous function with a finite number of step jumps.

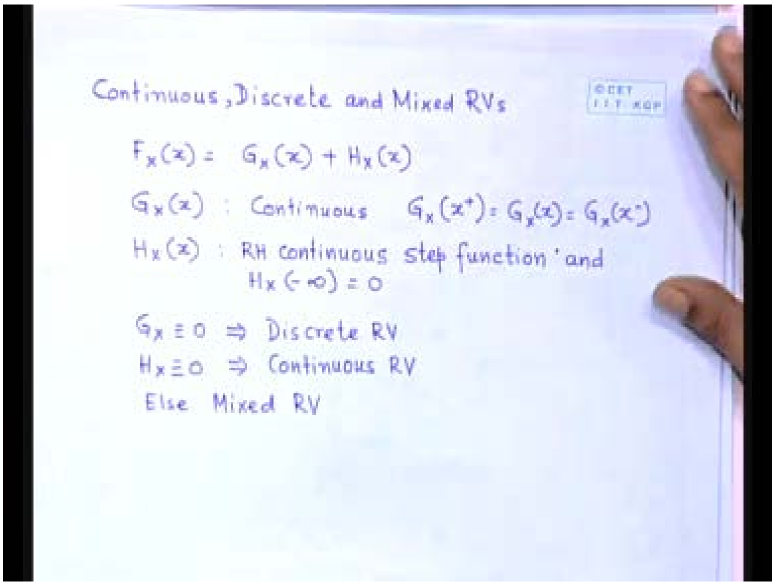
[Conversation between student and professor: 28:57 – 29:05] Sir, what do you understand by negative value function, probability density function, yes function or probability distribution function? Probability distribution functions cannot be negative. You have shown that if the axis considered as 0..... no no this is x here (Refer Slide Time: 29:14) this is F x; F x is always positive, that is the y axis okay Sir okay.

(Refer Slide Time: 29:27)



Probability distribution function cannot be negative simply because their probabilities so probabilities cannot be negative that is why.

(Refer Slide Time: 00:29:31 min)



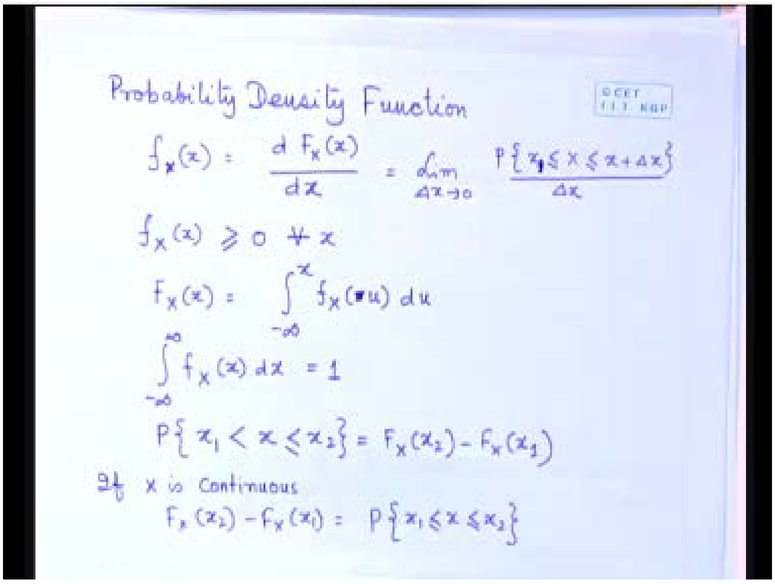
So this is the way of writing that... so any probability distribution can be written as a... it is this is called as Lebesgue's D 2 I mean Lebesgue's decomposition theorem. Anyway let us not get into technicalities; you can understand things very simply that any such function which is piecewise continues can be written as a sum of a continuous function and a sum of some discrete jumps so they can be broken up. So this part this continuous part denotes the continuous random variable part of the experiment from the continuous part of the experiment and this denotes the discrete part of the experiment.

So, so for the continuous part G x x plus this is the definition of continuity of any function that the that the left hand limit and the right hand limit converge. So G x x plus is equal to G x x is equal to G x x minus this is that is why it is called continuous and for H x this is there this these are some right hand continuous step functions like the one we have just now seen okay.

So if G x is 0 then you have a purely discrete random variable, if H x is 0 we have a purely continuous random variable and in general we have both right. So these are the various kinds of random variables that you may have.

Now let us come to probability density function. Probability density functions are nothing but rates of change of probability with the value of x okay.

(Refer Slide Time: 00:31:01 min)



So so how is it defined; it is defined as d of F x by dx. Now obviously since the function is... now which means that it is x less than or equal to x that is the limit. Basically what I have done is d of F x of x I have I am writing in terms of probability, it is a probability that x lies between small x and small x plus delta x that divided by the ratio of the internal delta x and limit of delta x tending to 0. So this is how it is defined in terms of probability.

Now obviously where this F x will take jumps this function will become some derived delta function. We know about derived delta functions from our network courses. So obviously since for all positive increments x since since F x is a non-decreasing function so therefore this must be greater than or equal to 0 all the time it cannot go negative because F x is never decreasing okay.

And how is Fx defined?
It is minus infinity to x that is if you have the density function you can derive the distribution function in this manner simply by integrating phi from minus infinity because from minus infinity you will get a zero initial condition okay because we know that F of minus infinity is 0 so you will get just the function F x. The the other function which will come for this limit will be 0 right. And obviously now if we put you know that F x plus infinity is 1 so minus so this function should be such it should be such a function that when integrates integrated from minus infinity to plus infinity it should evaluate to 1 right and this we have already seen this property so we do not need to see this again.

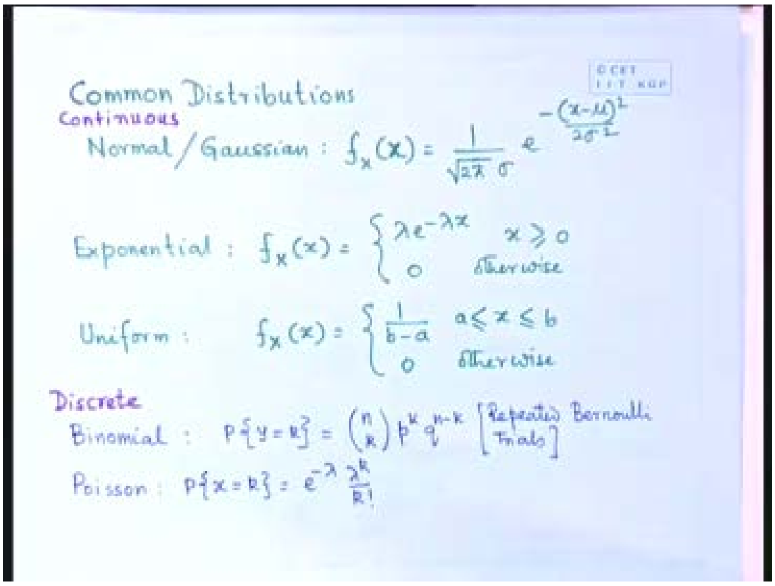
Now once you have probability density functions of the... we have we we always deal with various kinds of probabilities probability density functions. For example we have a very common or normal or or or Gaussian which are of the form like this okay, this as you know this is a very celebrated distribution function because it has because of two reasons, because of a particular property called the law of large numbers which says that if you take if you have any distribution any random variable from any distribution that if you take a large number of samples from it and if you average them then that average tends to obey the normal distribution that is a very useful thing which is often used for generating Gaussian random variables.

We will discuss this; we will, at the end we have discussed how in the computer we can generate random variables okay, many of you will will require it. So this is this property this law of large number property is used to generate Gaussian random variables.

Then you have an exponential function which is typically used in queuing theory. There are various kinds of for example there are there is the uniform function. when you when you do not know anything about how the probability probability distribution of a variable is you can assume that it is uniform that is in an interval within an interval the probability distribution function is flat that is the probability of getting any value of x in the interval a to b is the same, they are equally likely. So when you have do not when you do not have any information it is it is it is reasonable to assume that they are they are equally likely okay and corresponding you can have there are some well-known, there are there are many many distributions which are used in various contexts.

For example, in reliability context people use ((00:35:40 min)) distribution, in communication context people use (ra.....35:42) distribution so there are various a kinds of distribution. So I have listed only a few of them so this is a binomial distribution that is in what is the probability that in that in n fair trials you will get k balls of y let us say. So the probability of the... if the probability of drawing a white ball is p and the probability of drawing a green ball is 1 minus p then like that okay. So basically this comes from a simple... these are basically terms of the binomial series.

(Refer Slide Time: 36:22)



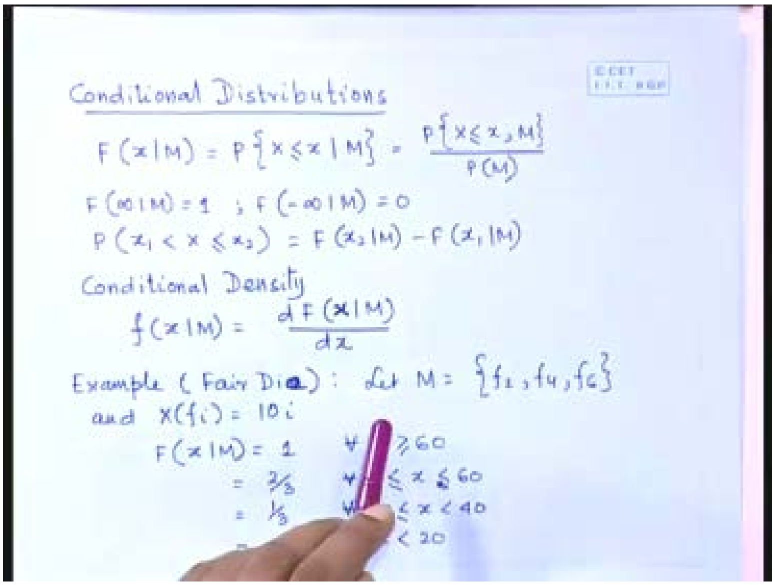
And similarly you can have a Poisson distribution which are, which are distributions which characterize waiting times between discrete arrivals, so they are also used in queuing theory. So this is just an example of various kinds of distributions. Now we come to what are known as... so we can now talk of... previously we had talked about conditional probability so now we can talk about conditional probability distribution. We are going in the same way, okay.

So what is F of x given M?

M is an event that is suppose you you know that some distribution F x is given to you okay. So now if you if an additional information is given to you that capital M has already occurred that is whatever has occurred satisfies the event capital M then can you can you make a more intelligent can you now now estimate what is going to be the probability that capital X is going to be less than or equal to x already if it is told that the event M has occurred right.

The event M may be related to that same random variable or even otherwise, whatever it is. So that is again... so this means that the probability that capital X is less than or equal to x given that the event M as an operator so that is again by the same rule it is the probability that this and this jointly both occur divided by the probability of M, these are the standard condition probability rule. This is one event (Refer Slide Time: 38:02) this is another event so probability of a given M is probability of a and M divided by probability of M okay. And this is a this has to be a probability distribution; F F of x given M is also a probability distribution just like F x so it must obey all its properties so obviously F of infinity given M must be 1 and F of minus infinity given M must be 0 and probability of and you can I mean from this thing you can say that probability of X which is greater than x 1 and less than or equal to x 2 is again F of x 2 given M minus F of x 1 given M. That is whatever probabilities x 1 and x 2 generally satisfy, the same properties this also satisfies.

(Refer Slide Time: 38:53)

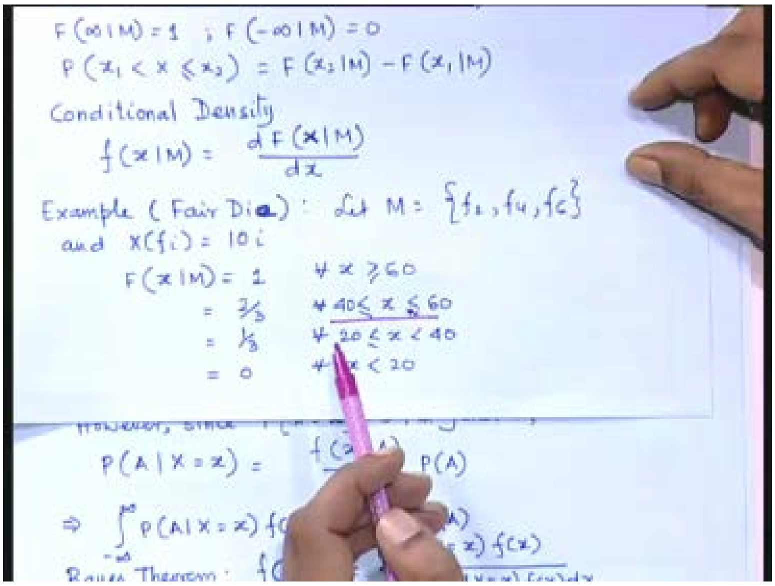


You can also examine let us say you can also define its density function. Once you have defined the defined the distribution function the the density function is simply derivative. So you can define this as D of F x given M by x in the same way.

For example, consider the case some of fair die and suppose it is told it is told that the an an an event phase has occurred that is the event M that an event phase has occurred. So what is M? M is f 2, f 4 or f 6 either 2, 4 or 6 have occurred when M is told. And suppose the random variable x is defined as in in our previous way that is for every this thing you get 10 rupees for every point on the die. So now how do you characterize this conditional distribution function? It is 1 for all x greater than or equal to 60 naturally it is 2/3rd for all x between 40 and 60, why? Because... what is the probability that... this is what...

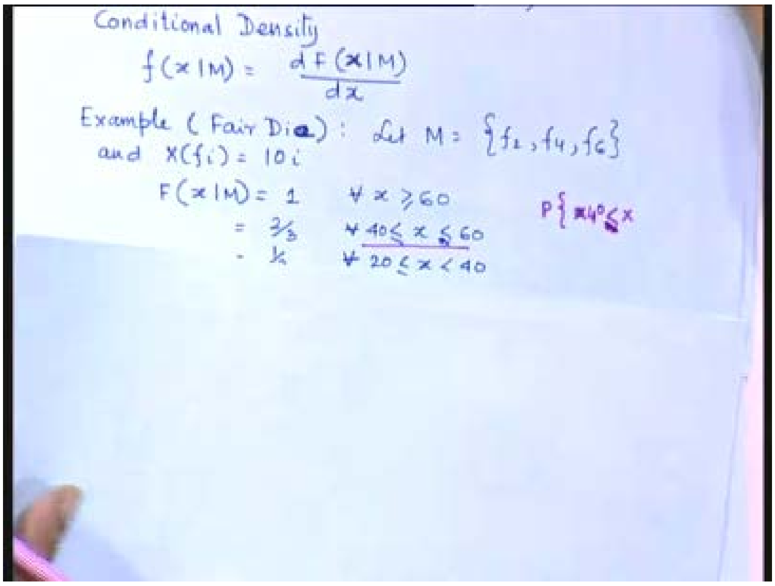
[Conversation between student and professor: 40:21 – 40:23] it is first of all... find out so we have to find out what, we have to find out the probability that X is X is greater than equal to greater than 40 let us say... actually we should write it this way...

(Refer Slide Time: 40:22)



[Conversation between student and professor: 40:52 – 40:54]

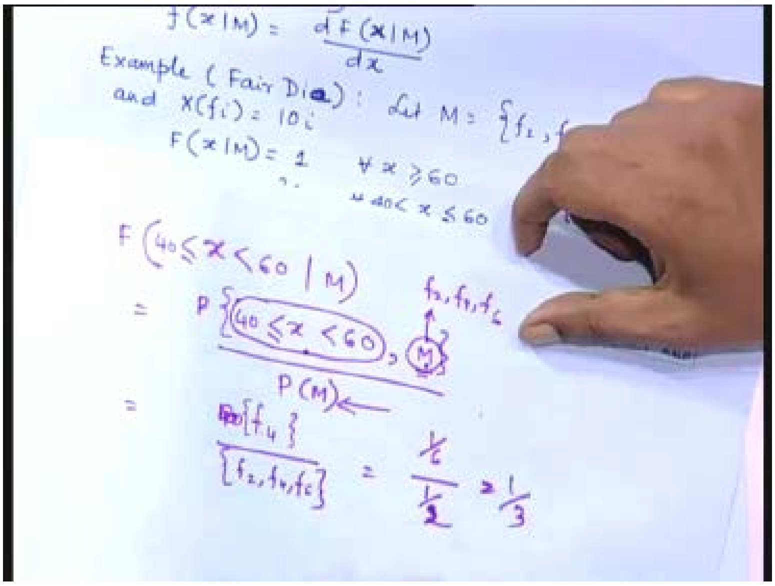
(Refer Slide Time: 41:06)



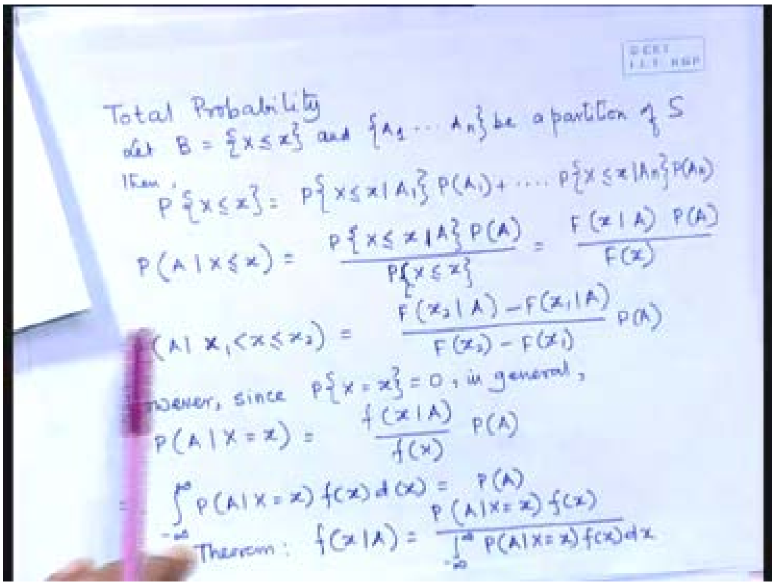
It is less than less than or equal to... it is it is it is less than 60 because we have already included greater than equal to 60 here right so it is less than 60 okay does not matter. So this is this we can separate it, it was less than 50, no no no why greater than or equal to 60 then it should be 0... wait wait no no no, so what is this? We we want to determine F x less than 60 greater than or equal to 40 given M this this is my problem, okay so that is equal to probability that x less than 60 greater than or equal to 40 and M, here is my probability of M this is my definition.

So now what is the probability that x is less than 60 and greater than or equal to 40 given that M is f 2, f 4 and f 6, it is only 4 correct? x is less than 60 what is what is the set of event? probability probability will evolve later. So this and this what is the set of event? The set of events is naturally 40 that is basically f 4 should not say 40; the set of event is F only because this contains f 2, f 4, f 6 this contains only this also contains only f 4 and what is the probability of M? it is 1/2 it is 1/2 because if you assume a fair die then it is 1/2 so it is because this is f 2, f 4 and f 6 so it is the so this one will be 1 by 6 divided by 1 by 3 1 by 2 so it is 1/3 I have written wrongly. If if if this is the case (Refer Slide Time: 43:17) this should be 1 by 3, if I include this as equal to it will be 2 by 3 right? So similarly this is going to be 1 by 3 and and this will be 0 agreed?

(Refer Slide Time: 43:12)



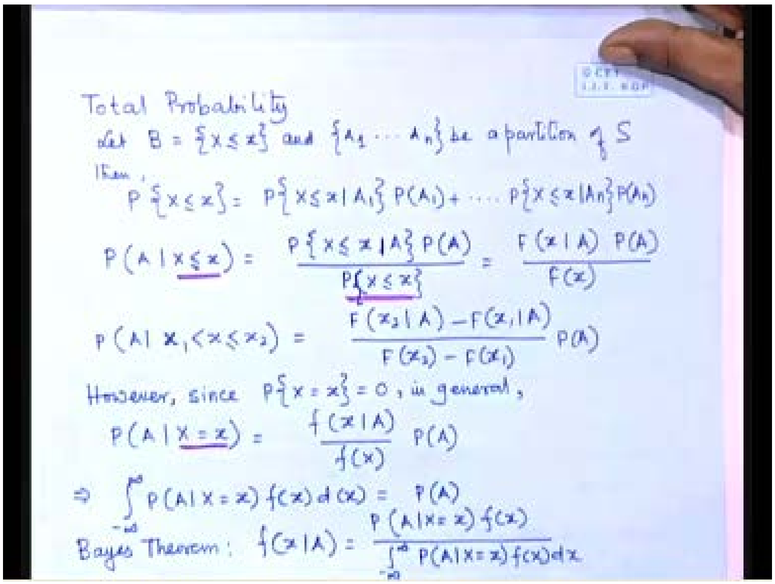
(Refer Slide Time: 43:52)



So now we can define the other things that is total probability and we we ended the last class with total probability and Bayes theorem. So we can also define now the total probability theorem and the Bayes theorem in terms of random variables; previously we had defined in terms of events okay so that will be like this that is let capital X greater than or equal to x be an event b and again A 1......... A n be a partition of S as usual then we can we can again say that probability of X given x that is probability of B is probability of X given x X less than or equal to x given A 1 into probability of A 1 plus dot dot dot; all the terms. This is the standard breakage of probability that we have done.

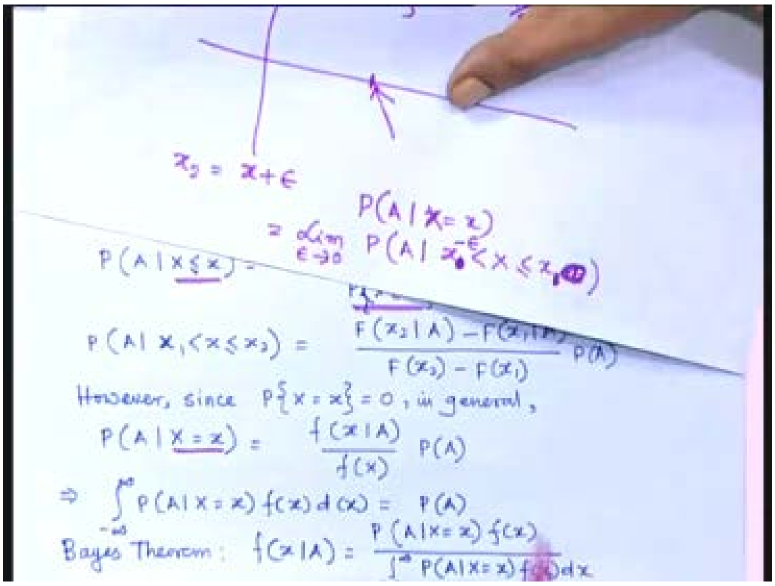
And again we can... this is my condition in probability statement (Refer Slide Time: 44:52) in if if I have a if I have this condition on probability that probability of an event A given X less than or equal to x; I am relating on both sides. Probability of an event A, basically we we are talking about probability of an event A given another event B so so I mean sometimes I am putting what is given in terms of the random variable, sometimes I am evaluating the probability given some other event, both you can do. So that is related like this standard thing because this is going to be... just simply treat this is an event A given B and then then apply the condition probability formula you will get this. This comes directly there is there is no problem, only problem comes so these two are very simple to derive. If you just put the probability expressions then you will get them; only thing is that here there is a subtle problem because you cannot define probability of A given X is equal to x in this form because of the fact that you cannot write it in terms of probabilities because probability... when you have a when you have a real line when you have a real line real variable for example then then the probability of having X having a particular value is actually 0. Now it is infinitesimally small so therefore you cannot put it in the denominator.

(Refer Slide Time: 00:46:25 min)



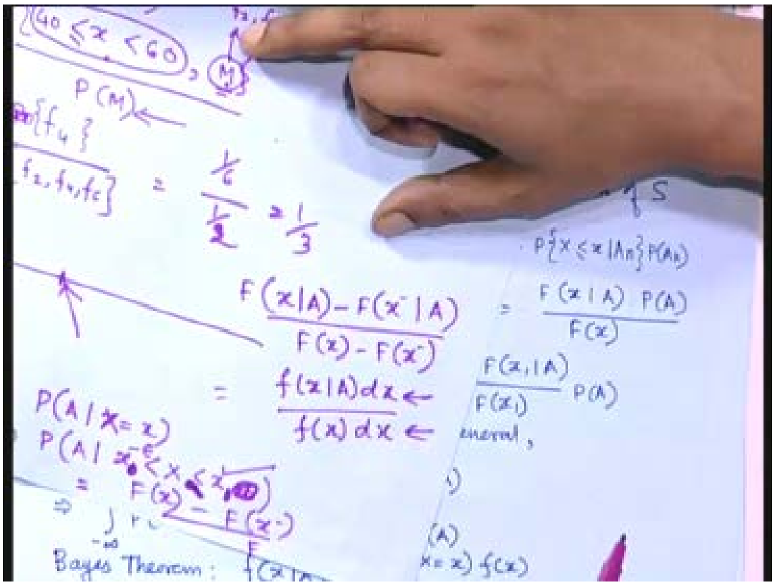
See, if you want if if you have if you have this sort of a clause (Refer Slide Time: 46:28) then in then then it comes in the denominator in my evaluation. So if you put instead of... this logic you can put in the denominator because this probability is going to be finite. If if you put this in general you cannot define it like this right. So so in that case what you have to do is that you have to define it in terms of distribution function rather density function because that probability you can... actually this will turn out to be... you you can write it exactly like this for example x less than or equal to x 2 and greater than x 1 what will be the value you can you can write... suppose you write that suppose you write that suppose you write that x 2 is equal to x plus epsilon x rather if you take yeah x plus epsilon and then you take you you you you put this formula so probability of A X is equal to x can be written as limit of epsilon tending to 0 probability of A given okay so X is less than or equal to x plus epsilon and it is greater than x.

(Refer Slide Time: 48:26)



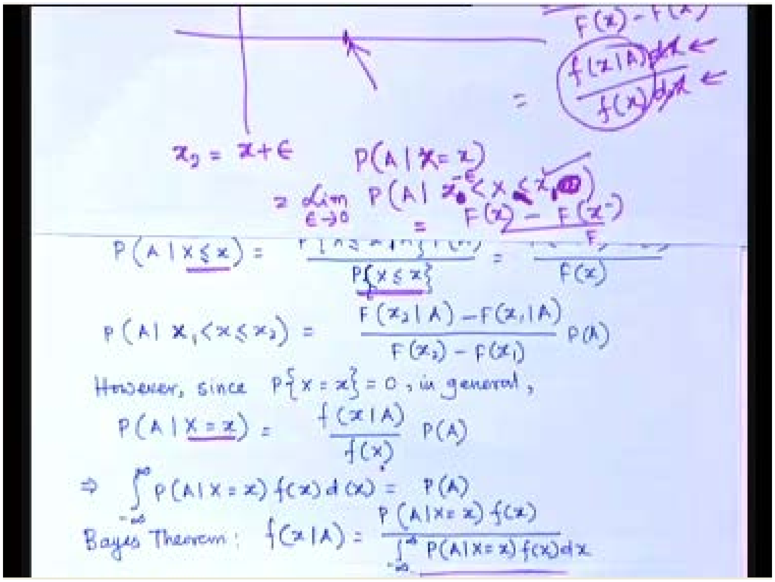
Now if it is greater than X then x will not be x, actually it should be put in the other other way; it should be X minus epsilon, X is greater than epsilon and x minus epsilon and it is less than x, no less than or equal to x. then you have... what will be this this will be... by this formula it will be this so it this case this has come in the second way so this will be this going to be x and this will be F x 1 minus A so it will be minus F x minus by F x 2 minus F x 1 right? F x, no no no no no... F x given A minus F x minus given A divided by F x minus F x minus. So this you can write as F x given A dx divided by F x dx.

(Refer Slide Time: 49:33)



Actually both are small so some something like a hospital rule situation is arising here but but when this dx dx will get cancelled you will get a finite value F x given A divided by f x that is how this is coming, both these as X is equal to x both the probabilities are tending to to very close to equals so it is very so you will basically get a get a 0 by 0 form if you want to write it in terms of distributions but but if you write in terms of density is then then see this dx’s will get cancelled you will get a... that is this 0 by 0 form will have a will have a definite limit if you write it in terms of densities. In terms of distribution it may look like 0 by 0 but it will have a limit. It is like the case of sin x by x; as x tends to 0 sin x by... both top of sin x goes to 0 x goes to 0 but that does not mean that limit of limit of x tends to 0 sin x by x is not defined it is defined as 1 right?

(Refer Slide Time: 00:50:30 min)



So because this will go to (Refer Slide Time: 50:32) both will tend to go to 0 you will have to write it in terms of density functions okay. And then you have a you have you have the corresponding finally you have to corresponding Bayes theorem. That is basically what they have done is they have they are the same thing you are writing, only the only the denominator is now being expressed in terms of all the posterior probability.

So f x given A so you are writing x give A in terms of A given x that is that is the technique and that is what Bayes theorem achieves that is probability of A given B can be written in terms of probability of B given A right. So exactly that is being achieved that is probability of X given A is probability of A given X, it is the right hand side is written fully in terms of A given x. this is nothing but probability of A, this this denominator is nothing but probability of A. So f of x given A into probability of A is probability of A given X into f x. Finally that is the expression.

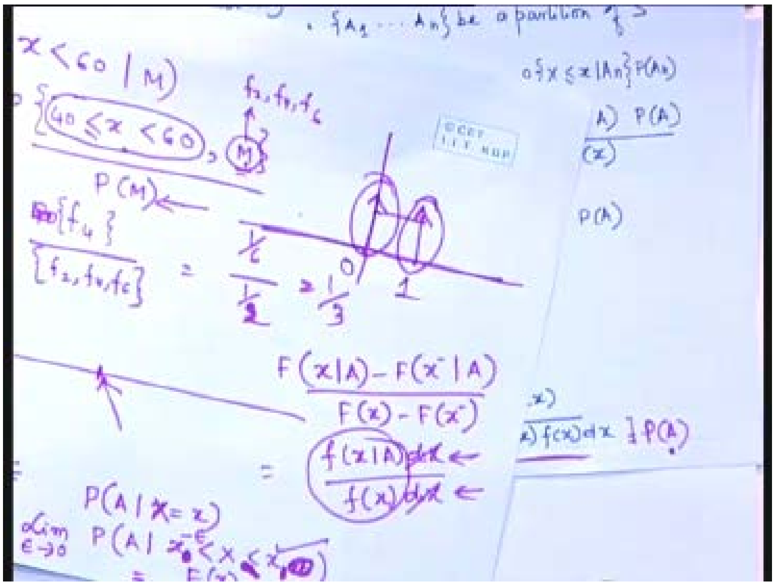
Now so the denominator is nothing but probability of A. So basically f of x given A into probability of A is probability of A given X is equal to x and into f x. So you are writing you are inverting the two sides. You have to take a you have to take a little time and think about this probably solve some problems. I will give you a problem sheet in the next class.

So basically what we have done in this class is that we have defined a random variable as a function of the outcome and then re-derived all the results that we have done in the earlier class. Earlier class we have characterized probability of the events, now we are characterizing probabilities of functions of events that is what we have done, that is all.

[Conversation between student and professor: 52:43 - 52:57] Yeah probability density function for defining as G f x G x, yes, limiting of the distribution function, derivative of distribution function, so if the random variable is a discreet one then it will have jumps.

For example our our random variable of that coin tossing experiment what will it have? It will have one impulse here and it will have another impulse here (Refer Slide Time: 53:08) so the probability density function will have two impulses; everywhere it will be 0, only at these two points at 0 and 1 it will have two impulse of size 0.5 that is all.

(Refer Slide Time: 53:17)



Thank you very much.