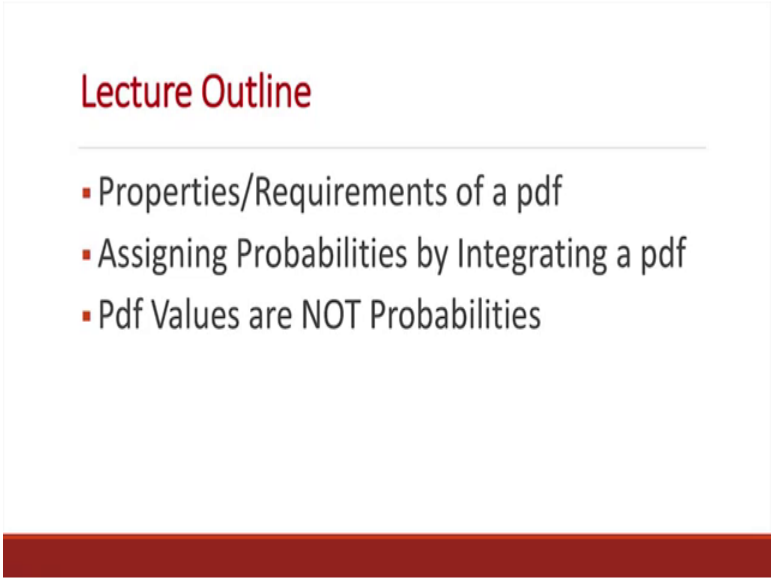
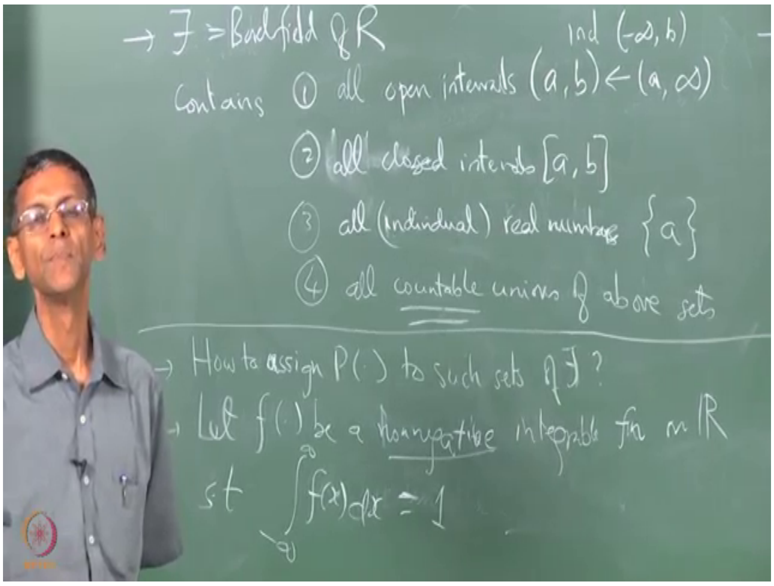
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 44
Probability Density Function (pdf)

(Refer Slide Time: 00:13)



(Refer Slide Time: 00:19)



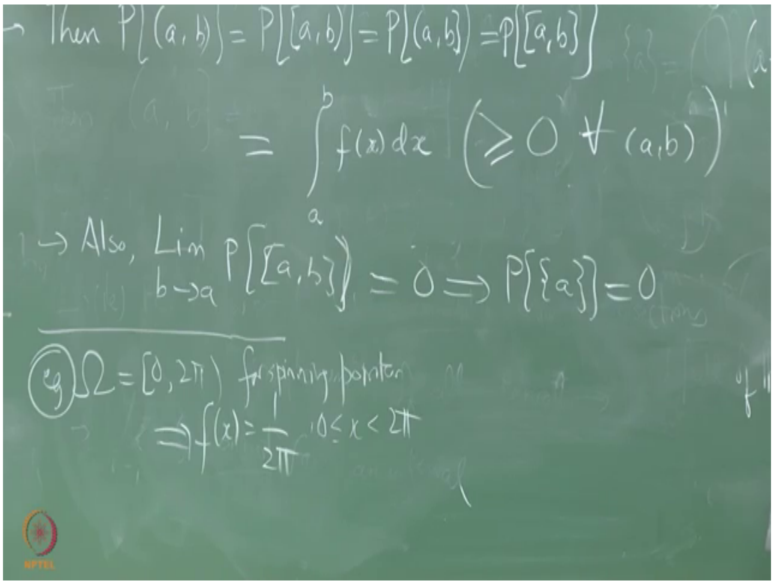
So, the question is how to assign probability. Consistent with the axioms of this of F, this F, each of these sets we need to assign a p which is consistent to the actions of probability. One the most popular way or the most important way is to start with a Riemann integrable function which is a standard integration that we have been studying all along right.

So, let some f be non negative integrable function on R on the real line it should be non negative what is the main the other requirement most important requirement that we require on f. We require this integral from minus infinity to infinity of f of x such that this is what? 1 exactly. So, I have written it correctly, please note let f be a non negative integral function fn over on the real line such that it is area as represented by the integral entire area is unity.

Based on this function how will we assign a probability to any interval? This is nothing to do again with remember we are going back to the first part of this course where we only talked about omega f p, no we are not really talking about random variables anything at this point we are just talking about the basic omega f p. Now, where we have we have nailed omega let us say as equal to R for the time being it is not a big issue to shrink it down to any subset of like 0 1 or whatever right. Then f we have said it is going to be all of this of course, you know if ever if you have omega it is only some 0 1 for example, then you may not you would not even get any anything of this kind it you are you know you will, you know all of this will be part it will be inside the omega itself 0 1 whatever. So, we do not have to worry about infinity when it is not there right.

Now, P if I mean I do not necessarily need to always integrate for I is infinity, infinity if I have a smaller portion of the real line I can live with a function which is whose area is one only on that portion of the real line I do not need to look at the entire real line. But the entire real line is useful in the sense one once you solve it for this case you have solved it for all for omega being any piece of that real line, is not it. So, what is P of a b?

(Refer Slide Time: 03:34)



Should I, let us say that open it well. There is a we have to look at ok. A distinction here do we take the open interval or do we take close to it, it turns out it does not matter. All of these combinations will have the same probability. What for things can I do? As long as a and b are finite I can look at 4 possibilities here is equal to what? Based on f what should I do?

Student: (Refer Time: 04:19).

I have to integrate from.

Student: (Refer Time: 04:22) a to b.

A to b, that is as simple as this. So, it turns out if you do this then the resulting probability measure as I call it is entirely consistent with the axioms which we started out with because if you take for example, disjoint intervals a b and c d you require that the probability assigned to the union of a b and c d be the sum, but integral a additive. If you integrate over a b and then integrate over c d if you add it is like integrating over the 2 into you know intervals or supposing you say a to b and b to c for example, it is the same as integrating from a to c is it not.

So, the additivity property is automatically satisfied for disjoint interval. And p of omega is what? Obviously 1 of course, when for P of omega we only I mean if you say minus infinity you are not going to write all these other things is just try to open interval. That by definition we have taken it to be 1 and no matter how small we make this, this is never going to be negative, because f is constraint to be to be what?

Student: Non negative.

Non negative. So, this integral can never be negative for all a b. Note that this is only I mean, this is the most popular way of assigning a problem and can we cover all of these, yes we can we have covered open intervals we have covered closed intervals it turns out that if you look at individual points you will get what probability.

Student: 0.

0.

Student: 0.

And limit b tending to a of P of let us say a b is 0. So, individual points how I mean this is roughly seeing yeah I do not I do not think, I can I mean you would have to put some more mathematical rigor into this by miss roughly saying and assuming that a is fixed in you can move b towards a or something that is what I mean by this right.

Student: (Refer Time: 06:59).

If you make the being go and fall on top of a you find you by this construction you will only get 0, so individual points have 0 probability. But does that mean individual points cannot occur? No, just says that the probable only intervals have finite probabilities, but if you look at this particular point in an interval and ask for you know the probability of observing that point measure the infinite precision by this kind of a construction such a point will have 0 probability. And not only this a countable union of such points a i is also probability assigned to such as head is also 0.

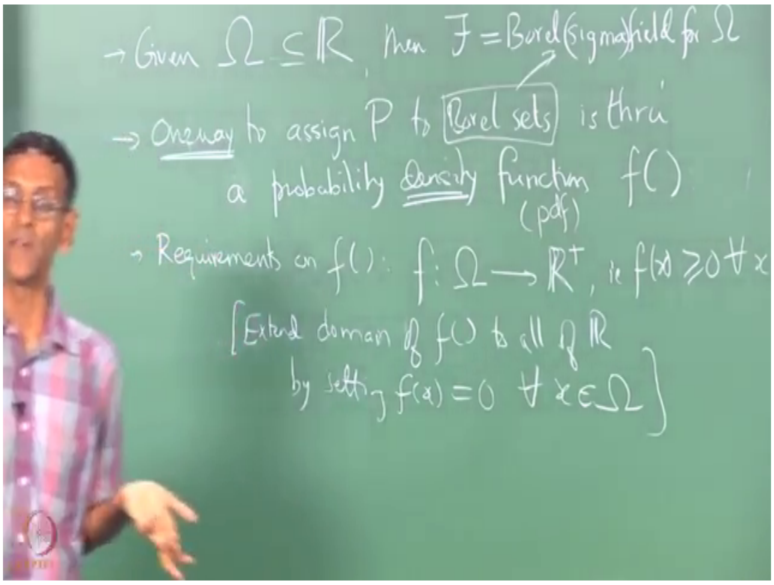
In other words this the countable union of individual points again has 0 probability. There is a probability of observing a 1 or a 2 or a 3 is also 0. Only intervals in that Riemann integration assigns a positive quantity was the measure only two intervals right. So, and obviously, there is no shortage of such functions. We will see several examples and all of you have studied them in various math scores earlier right. So, there is no shortage of such functions.

So, the question is again the match does not tell you which function to use in a particular situation. It is up to you have to figure out what is the problem a function best suited for the problem at hand. If you have a spinning pointer which can come to rest at any point on the circle you might want to assign what is called a uniform probability measure in the continuous sense not discrete sense.

If I take 0 to 2 pi in radians omega first spinning point of and note that I am leaving out 2 pi. So, then I will say f of x is 1 by 2 pi. So, eg this is just an example. This is one possibility. What does this mean? It says that no interval is going to be favored or any other interval of the same width getting between 0 and pi by 4 probability of a point R coming to rest in that interval 0 to pi by 4 is exactly the same as any other interval over which pi by 4, the interval is the probability is dependent only on the width of the interval not on the location.

In fact, this is what is called a uniform distribution in the continuous space and typically you start with 0 one rather than 0 2 pi, but since we started with spinning pointers I have to find angles I am taking it as 0 2 pi. So, there is 1 by 2 pi is a normalization. This 2 pi will go away if you make the upper limit equal to 1 that is the all important uniform 0 1 distribution.

(Refer Slide Time: 10:27)



We were taking this we are going to take the sigma algebra F to be the Borel sometimes just called Borel field and you can of course, narrow it down to this portion of the real line that you are interested in you do not have to look at that all the subsets Borel line. For example, in this pointer case we are only looking at omega equal to 0 to 2 pi it does not really make sense to look consider sets like minus 10 to minus 5 and so on and so forth right.

So, this collection of sets we are going to be assigned probabilities. How we are going to assign probabilities? One way of assigning probability is then and members of this Borel field are called Borel sets. Let me write that let me this thing out in full probability density function with emphasis on density.

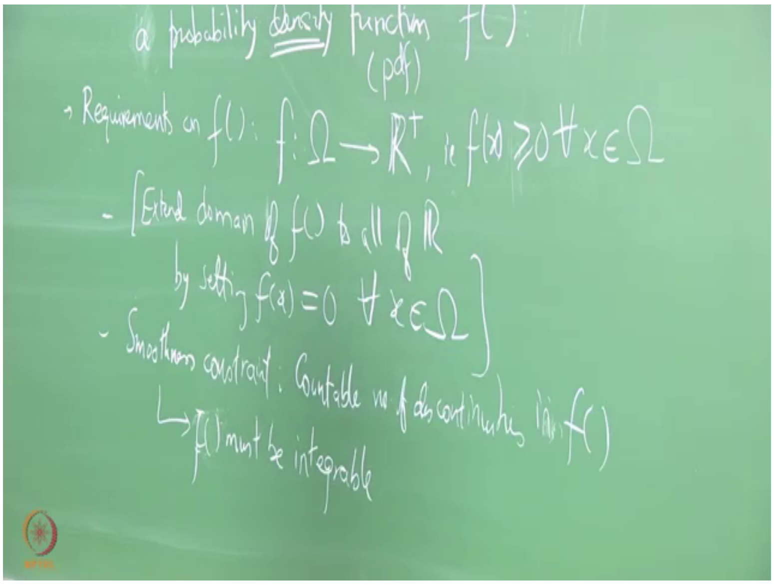
So, again (Refer Time: 12:02) I mean before l says these are elements or members of the Borel sigma field itself right. So, what are the properties of this probability density function? We are going to call it pdf from now on, we will use this lowercase small latest pdf. The pdf, the requirements of the pdf let us say emphasize is just one way of doing this and it should be first of all integrable or let me even be more start from even more basic viewpoint. It is basically the domain of this f you minimally it must map, it must be defined for all points in omega and this R plus is sometimes is a notation used for non negative real numbers right. So, this non negativity is very very important, if it is any negative values it cannot be a pdf.

So, add a minimum it should be defined for all points in omega, but most people consider an extended definition by looking at the whole real line. So, let me say that is just a convenience right. So, extend domain of f to all of R. How do you do this? By setting f of x equal to 0, for what? x naught in omega. So, this is typically done just to keep our thinking straight. We have some function which is which takes non-negative values only or positive values only in they set omega which is this which is the portion of the real line that we encountered in an experiment. Such as a pointer which means between 0 and 2 pi or any other limited portion of the real line for any particular experiment that we have to consider right.

But this it you know typically the way in which these probability density functions are defined is to define them to be 0 outside omega as well. So, that there is no confusion basically. This is just to avoid confusion right. So, therefore, if you integrate 0 functions in over any interval you get 0. So, the idea is to say that if you look at intervals outside omega even though they are not strictly speaking part of this Borel sigma field itself you can say that they are big, they are basically null events, they, for the case of a point or any interval which is in the negative portion of the real line is a null event it will never happen. So, it has 0 probability and that is consistent with this.

So, having taken care of this, so then we also want some smoothness constraint on f, we do not wanted to arbitrary jump around. So, how do we impose a smoothness constraint on f. Before that let me say. So, smooth smoothness constraint is necessary so that it becomes integrable.

(Refer Slide Time: 15:53)

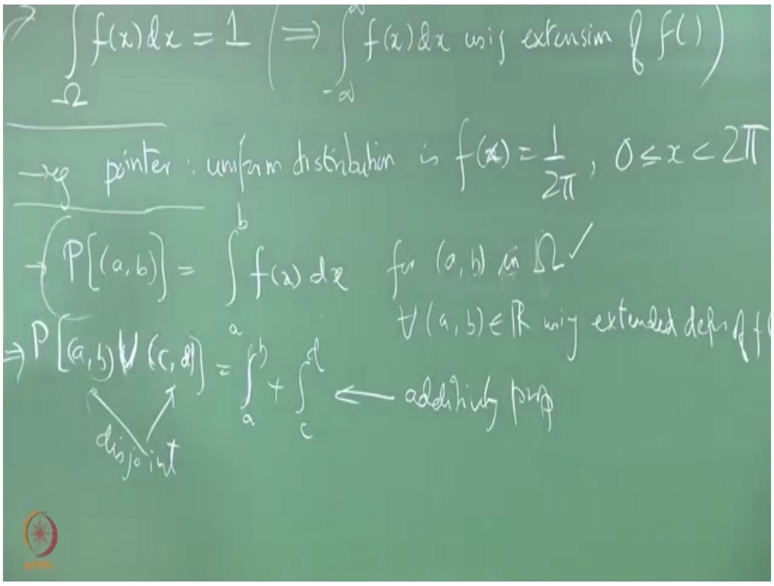


This I did not say yesterday I will put it here even before the integrable condition. So, we will want a countable you know at maximum we want let us say not even countable, but I miss have not written it I will keep that countable number of discontinuities. We do not want to have the thing jumping all over the place. I mean maybe I should have started with f here anyway f has, in I will just say in a.

So, this function does not arbitrarily jump around like for example, taking one value for power real number and rational numbers are in different value for irrational numbers. If you think of such type of functions there big big you know you know you know it mathematically it might be to define something, but is it integrable. So, this has to do with the integrability constraint. So, f must be integrable and not only I mean integrable in the Riemann sense that we have been studying all along. So, there are different methods the definitions of integration are some of you may have seen it right.

But we are going to only stick with the standard Riemann integration, but we have encountered all these years ever since we started studying integration in high school or whatever. So, this integral over omega f of x dx must be equal to 1, it must be non negative and it is integrable, integral over the omega must be 1, this is just a normalizing condition right.

(Refer Slide Time: 17:22)



Even if this integral is finite that is sufficient for us because we can always normalize by the finite value of the integral to get unity right. So, basically we can we can consider any f which is non-negative and integrable and fine and having a finite integral. So, this I know yesterday I may have written minus infinity to infinity, but fundamentally you want it is integral over the set omega to be to be 1.

Of course this automatically will mean the integral over minus infinity to infinity is also 1, using the extended definition because we have said that we can, if necessary we can extend it to cover the entire real line. So, therefore, most textbooks will only will only give this, but remember what they really mean is this what they really want is this right. In the case of the spinning pointer you want integral if you know if you do not like the uniform. So, yesterday's case was the pointer. The uniform probability distribution is what? First, just because 2 pi is the same as 0 I am just saying that the domain is closed on 0 and open in 2 pi, but that is just a technicality I just let it.

So, even if you did not want like this particular distribution if you wanted a different distribution you can go ahead, all you need is any non negative function f which integrates to one nominally or you know any finite value over the interval 0 to 2 pi that will suffice and obviously, there is an infinite number of such functions and so there is no. So, as in the discrete case in the continuous case also we do not the theory does not care about the physical meaning of or the whether there is actually gives you useful probabilities or not it is not something that clearly worries about, theory only cares about the mathematical consistency of the definitions that, is it ok.

So, this is just the example. The most important point is that in the probability as I said yesterday of any interval a b is the integral from a a to b of f of x dx. Now, here you can ask where are these points a b typically both a and b will be the interval a b will be in omega or this is a close open interval just for the sake of completeness suppose typically this will be in well I this will be in that Borel, it does not matter this interval in I just say I do not say element if. So, but it turns out that using again the extent definition you can take any two points a b on the real line. Definitely this is true, this is true for all a b element of R using extension and the point is that because you get is, you get a null event you do not have to say that it is explicitly 0 or something you can live with this integral over 0 function and you do not have to keep making special cases right.

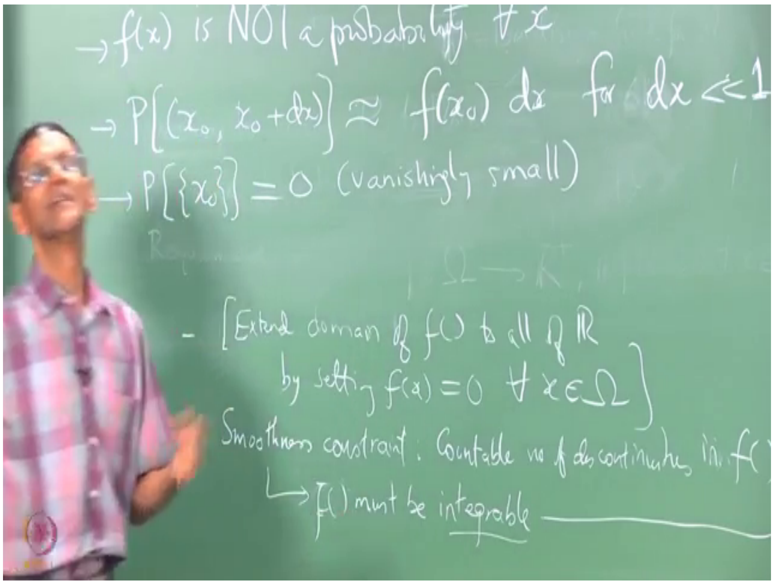
But the nonzero probabilities are; obviously, going to come only here, they cannot come from the general extended the extended domain extended definition of f and more importantly the if I have a union of disjoint sets. What will this be? This will be the sum. The sum from a to b plus the sum from c to d that is the additive property which is automatically satisfied because integrals add right. So, if they are disjoint so this is the additivity property right. So, all these things go to show that they are defining this the probabilities by this is perfectly consistent with the axioms and so on and so forth.

What are more properties of this f? Remember this f, f of x is not a probability for any value of x.

Student: (Refer Time: 22:50).

Union a b union c d. Let me go back up here and start. The values of f of x themselves are not probabilities. Am I going, so find any questions?

(Refer Slide Time: 23:16)



This is a very constant source of confusion in among the students. Yes, it is only f of x is only a probability density you have to integrate f of x to give you a probability. This is where the pdf differs fundamentally from a pmf. A pmf is a collection of probabilities. But the pdf is not a probability.

In other, only thing you know is that this f of x is non negative, but it turns out it can be infinitely large also in some cases as we will see. There is no constraint on, no upper limit on the value of any f of x, f of x, but when I say f of x I am talking of the particular value of f at some point x. Please distinguish between a function f and the values f of x. So, f of x is not a probability. This is understood by everybody. You can get any arbitrary large positive values here. Need not be bound to where between 0 and 1 because you are integrating and such examples will follow.

What is the probability of this? If I take some interval x naught to x naught plus dx for some small x, so small dx, what is this going to be? This is by our usual into calculus approximation first order approximation Taylor series whatever you want to call it is what. What is this? For a small interval x naught to x naught plus dx. What is this? What happens in that if I put?

Student: (Refer Time: 25:11) finding that (Refer Time: 25:12).

This in terms of this integral what can I simplify this integral when a is x naught and b is x naught plus delta x, it will be when I said this dx is not the same as I mean there maybe I should use some other dx here, but does not matter just. So, this is basically, typically dx is positive in such right. So, when dx is much smaller than one which means it is have a small increment. This is the connection between f and probability there is a probability of any small interval and we already said this yesterday P of the i, i this x naught for any x naught in omega wherever this is I said 0 but you can think of it is being vanishingly small. So, if I make this dx go to 0 yes the side RHS does in fact go to 0, but that is not to say that this x, x naught is a null event, it is not a null event right. So, we here we have an example of getting 0 probabilities for something which is not a null even remember because when you do the experiment you do get some real value some x naught.

But a priori the probability of getting exactly that x naught is vanishingly small a priori. But that it turns out we can actually condition on that particular observation x naught. So, that comes later. Even though it has vanishingly small probability we can. So, essentially we whenever we think of x naught we are actually thinking of some small interval like this which is also consistent with the engineering because it is allowing for some or shall we say imprecision in specifying the value of x naught or if you do not like x naught to be at one end you can put it in the middle and say x naught minus dx to x naught plus the x or whatever y you can do any of that stuff, but based essentially this kind of thing is just reflecting our imprecision in specifying or measuring exactly a real number which we cannot anyway do in practice.

So, this and why do we need and I guess it is clear that we do need some smoothness concern. We want an integrable function essentially, we do not want some crazy function which mathematicians are fairly are fond of constructing and then we do not to pull ahead or I saying how the hell do I integrate this right. So, that is why I am saying up front in our all our probability models of pdfs this f will be integrable it will have, if at all it has jumps or discontinuities. So, they will be finite in number countable not even countably infinite there will be finite, you can count them on the fingers of even just one hand not even two hands right. So, integration is not a problem essentially.

So, I think. So, as of now I mean I let me not anyway we are going to see a lot more of this probably a pdf later on. So, we will move on to a different way of specifying of assigning probabilities. Maybe I should have put delta x here to differentiate this dx from this dx, it does not matter. I do not think it is a major confusion, this you can write, you can use something else out here you can you can put dy if you want or delta anything. So, maybe I should also specify this, this is assuming that dx is positive. I do not want somebody to take get up and say what if it is, are you looking at very large negative numbers? No, I am not.