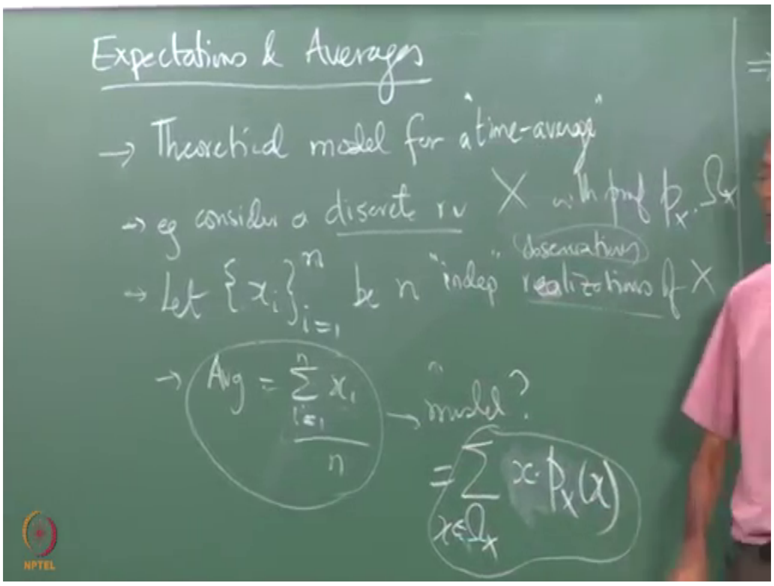
**Probability Foundations for Electrical Engineers**
**Prof. Aravind R**
**Department of Electrical Engineering**
**Indian Institute of Technology, Madras**

**Lecture - 63**
**Expected Value or Mean of a Random Variable**

(Refer Slide Time: 00:17)




So, there are two schools of thought here. When now the question is when to introduce this topic right? Should we do it alongside the PDF and the PMF in the PDF or should we wait right in my version of this course I have prefer to wait until we finish at least some modicum of description of the PDF's with.

Then only bring in this because I think it is more economical doing in this way, but in many cases people might expect you know this even earlier, but any way you are not as long as you get it at some point in the course I guess it is fine right. So this is a very important idea which starts of looking very simple, but then quickly becomes more involved alright.

So, what we are trying to do is I try to get a theoretical model for a long term average right. Long a long term arithmetic mean maybe for a time average what do I mean by time average?

So, I am just going to say a consider a discrete random variable some let us say X with some with PMF P x and let us say you take you do the underlying experiment some small n times.

You say I equal to 1 to n b sorry small n b n independent realizations this is the standard set up in probability theory right that is in realizations or observations whichever way you want to call it right. It is value that it takes at any trial is either a realization or observation both terms are used right especially the word realization if you right be comfortable with that ok.

So, of course, I mean in probability repeated trials are always considered to be independent never right dependent. So, this is these are n observations. So, this is basically now a time series if you think of it that way right you assume that at one instant of time you can only do take one observation.

So, therefore, you take n of them you can string all of these samples in a along the time axis right and you get a wonderful discrete time sequence d xi essentially. So, now if I want to calculate this average let us call it what should we call it.

So, let us just can let us not call it anything except average. So, let us say then define the average it is the sum of x i simple average right. If you do this average equal to right x i by n is there any theoretical model for this.

So, this can be calculated only if you know these observations exact, but a priori before without doing any calculations sorry without doing any experimentation right can you model this long term this kind of an average for especially for large n. So, is this is there a model for this and the answer is yes it is the so called expected value which you can write down in terms of the probabilistic description P x.

To understand that let us say you divide this sum I am saying I am I am specifically saying is discrete to make things simpler. So, each of these x a x sum i's has to be one of the values of X right and of course we have used as x i in a different way earlier we used it to a described right x 1 to whatever x n to be there n possible values.

Now, n is a totally different number here it is a time series right, but supposing let me instead of writing it let me just sort of try to verbally explain it can you not group this into the first value plus the second value plus the third value and so on. Let us say you are you are your taught yeah X is the outcome of die single die so you have only 6 possible values.

So, let us not confuse the 6 and the n. So, this is x each of the x i is only going to be a number let us say from 1 to 6. So, I group I split this right I instead of just doing it as 6 plus 1 plus 3 plus 4 and I am let us say I pull all the ones to one side I pull all the twos in the mirror and then like that so I can always write this in by splitting the numerator is not it.

And then for each of them which is 1 per 1 value of x if I what will that be that will be some n i n 1 times that value x and then you have an n is capital n small n in a denominator.

So, if you look at that at that portion of this fraction it will tend to P x of that value as n becomes large multiplied by the value that you are multiplying that the random variable is taking. So, I am just saying I am just going to say that the model is basically this.

For a discrete random variable I want you to look at this theoretical sum over all possible values that X can capital X can take X element of omega x right PMF P x and then of course, you have omega x also right which are the set of values in the X capital X takes right. So, then the model for this average is basically I am claiming is exactly this.

Of course these two are not going to be equal in it right this is some random quantity which you can keep fluctuating, but the claim is that if you make this n large enough and if this right if X actually has this probabilistic description, alright. If this is a then these two should get closer and closer as n this should right the way the fluctuation of this should decrease as capital n increases and ultimately this should become very close to this.

And so, if you make this n go to infinity for example, if you do this to the die experiment you what answer you will get you get 3 is it 3.5 or what is it if you say each face is equally likely yeah.

Student: 3.5.

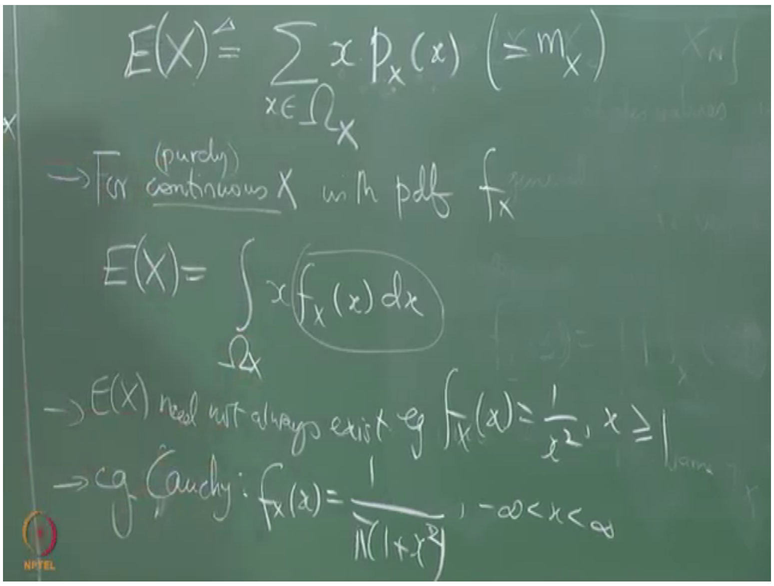
3.5 right. So, this number this long term average of some of die average of all the die throws should approach 3.5 which is this 1 by 6 plus 1 into 1 by 6 plus 2 into 1 by 6 plus 3 into 1 by 6 plus 4 into one by 6 plus 5 right. So, what is that so if you pull all the 1 plus 2 plus 3 plus 4 plus 5 plus 6 plus 6 into 7 by 2 which is 21 by 2 that divided by 6 is 21 by 12 which is 3.5 sorry.

Student: 21 by 6.

21 by 6 only not 21 by 2 21 by 6 it is 3.5 right so that is why we are motivated I do not want to the thing is this time series is only a step. So, I let do not get confused I know maybe this notation is not the best I should have put you know like the DSP textbooks, do I should put what this I in brack in as a parenthetical this thing rather than as a subscript that would have actually cleaned up this, but I did not do that.

So, I do not want to go hack the blackboard at this point anyway we are not ever going to do in this course consider this thing any further the time series argument we do not consider in this generally I just I have brought it in only to motivate this for a discrete random variable this is clear. So, some substances are what I am saying is basically this right.

(Refer Slide Time: 08:41)




It is So, called mean or mean value you it is we are we write this E is going now going to represent an operation the operation of taking an in statistical average I am rewriting that thing exactly so that is very clear without any right any of the this is not clear what I have written here I have copied it back here again.

So, I want to make I want to and we also call this as E equal to mx there are 2 a right this E is a more multipurpose as we will see some multipurpose operation symbol, but this m is very specific to mean this all right. We do not we just take exactly x here multiply by wait the value that x the value by the probability of x and add that if you just do this we call you know it is mx we also write it as E of x.

Now you might wonder why you want to write in such a complicated fashion, but there is a we will become very clear in before the end of today's lecture right. So, this E is right is it is like probability right in that sense is an operation it is and probability is defined as a function, but this is this is and you use the random variable name the capital X you put that within the parenthesis right.

So, this is the basic fundamental definition of the mean value and it turns out this is one of the right it is one of the most useful extensions of the theory in for just for doing probability calculations we just added only probabilities right and of course if you leave this x out you; obviously, you will get one, but by putting this x here it turns out you can get a much more a much richer theory and that is what we are going to be studying them almost virtually the entire rest of the course ok, fine is this for everybody. So, this is for a discrete random variable.

Let me also write out what do we do when the when it is continuous unfortunately in the continuous case right you cannot so easily talk of the probability of X taking the value small x because that is going to be 0, but instead you look at some small interval around that numbers small x.

So, I am talking of purely continuous x i am not I am going to say for lack of time I cannot look at the most general definitions possible here right that is the case of that waiting at a traffic light where you have finite probability at x equal to 0 and all that that I leave you to figure out how to you know calculate the means and other things for that particular type of distribution I am we I am going to say up front right we are only going to consider 2 class of random variables in this guy either they are purely discrete or they purely continuous right where we just do not have the time to take off on other excursions on other type the mix so, called mixed type ok.

So, for purely continues X with PDF f x here also nobody is going to prevent you from doing this kind of a calculation right you can if you spin a pointer each time you get some small x i nobody is preventing you from running that experiment n times and taking the average.

So, what is a mathematical calculation you can do with this PDF fx which will be a good model for this average and it turns out that this it turns out that averaged is basically defined as an integral of this you integrate over the set omega x.

Now why does this integral make sense because this what is this fx x dx it is a probability of getting that small value so I getting a value of small x i mean an interval around that small x either x to x plus dx or whatever. So, what you doing is you adding all those you are you are taking this probability multiplying by this x and then adding this integrals is nothing, but addition. So, you can think of the Riemann sum version of this which is a sigma over small intervals and then as you make the interval width go to 0 that the sum becomes an integral.

So, it turns out exact this is exactly what we want right in the case of a continuous random variable right. So, the calculus is a wonderful savior in all of these things so you can think of averaging even when you have continued a continuous function of x does not have to be just a discrete function.

And I think this kind of integrals you may have seen in earlier like when you do moment of ma what is it a moment of inertia or find a center of mass or something did he ever calculate center of mass for a distributed mass. Did you not do integrations of this kind? then it is an analogous to that is it. So, these definitions very clear please take the time out to understand them because they will be crucial to our future any all future work in this is I mean I do not want to just look at this section of the I want to turn my exclude them and look at you people also this is alright ok.

Now, before we proceed again that as long as this omega x is finite in some sense there is only finite number of points in omega x in the discrete case and this omega x is a discrete is a finite length interval here there is no problem about the existence of E x, but if you make this omega x go to infinity either countably in finite or infinitely long it turns out that there is no guarantee that E x will always be will exist.

So, you can easily come up with examples of E of fx which are valid PDF's which do not have expected values because the sum does not converged. So, E x need not always converged right. This is one of the simplest examples is to take the 1 by x squared PDF 1 by x squared for x greater than 1 this is a valid PDF.

How did we get this we got this by taking a uniform random variable from random variable from 0 to 1, u 0 1 and taking 1 by x right that had this PDF 1 by x squared. You try sticking this definition into that omega x is now the entire intervals from 1 to infinity will that integral converge no it will not right it does not converge. Because you get an integral of long ln x oh sorry you get ln x evaluated from the limits 1 to infinity and clearly that that does not have any it is not finite.

So, so, clearly this right we have to therefore, be careful right we just this is one thing there is also another problem that we have to address and that is 2 infinity is cancelling each other that is 1 plus infinity cancelling a minus infinity that is also something we have to avoid.

Because that also makes right may makes is a very poor mathematical model in case right you get some situation of that kind right for example, if I take the Cauchy this is a classic example. 1 by pi into 1 plus x squared I think it is the easiest Cauchy I can think off no alpha parameter or whatever right it is just simple Cauchy.

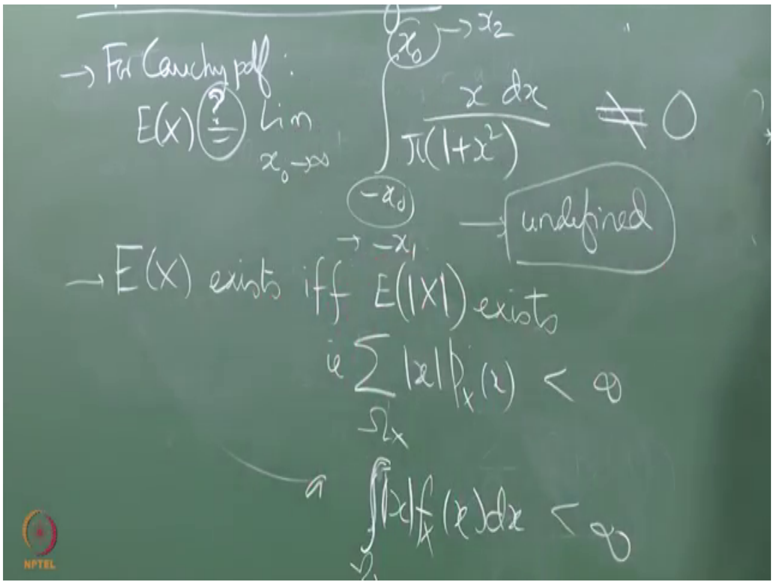
This here of course, you have minus infinity to infinity now you right stick this inside this now omega x will go all the way from minus infinity to infinity. Now you tell me if that integral should exist or not how will you decide if it exists or does not exist x divided by 1 plus x squared dx right does that integral exists from if you integrate from minus infinity to infinity the answer is no it does not exist simply because you have if you say that the value is 0 you are actually implying my infinity minus infinity is 0 which is not correct.

And in fact, you can test it out with an experiment you know the problem set 6 talks about how to generate Cauchy a random variable from uniform you stick that I write a program to do that take some long term samples of I mean a long term average with the with the Cauchy sequence right and then see if this guy converges to some value converges to 0 it will not. This is the experiment has been done and it has been reported.

So, let me save you the more time and wall effort, but of course, you are free to simulate if you want to simulate.

But if you do this when x i is a samples of Cauchy you do not get 0, it will just go the average will just jump all over the place it will not converge no matter how large you make n. So, so you shall not do so what I want to say is you shall not take infinity minus infinity and set it equal to 0. Why do I get infinity minus infinity right I get infinity minus infinity. So, let me write it here.

(Refer Slide Time: 19:13)




If I do E if I consider E x as let us say limit as some what should I say some x what is our standard notation for some large real number call it call it x naught going to infinity something I need some number right. What do I do? If I do this put a question mark here big question mark.

If you do this yes it will always be 0, if you integrate from you forget about this limit right this integral will go to 0 if you move from fixed minus x naught to plus x naught, but if you make this minus x 1 to plus x 1 and allow infinity to be reached independently separately here and here then you see that the integral is undefined.

So, this is not equal to 0, when if you change this to minus x 1 and this you change let us it say x 2 is actually undefined. So, how do we get a mathematical a statement of existence of the a test let us say for the existence of E of x right. So, it turns out that for whether it is a discrete case or the continuing case you can easily come up with a discrete count of part of this right all you have to do is consider the PMF 1 by 2 k squared or something from a 1 to infinity and minus 1 to minus infinity it is extremely simple to do that right.

So, so, basically whether discrete or continuous right we have we make this E x exists E x exists if and only if that is iff is if and only if E of mod x exists. Now what is the meaning of the this E of mod x that is I maybe, I should not have ie this sum converges. So, forget about mod x this notation for a second just think of this and this right forget about this just in this definition replace x by mod x in this definition also replace x by mod x.

So, the condition for this E E x to exists is that the corresponding sum or integral be absolutely summable or absolutely integrable. This sum itself can be positive or negative depending I mean that they in the individual components here that the sign of this x tells is there problem here this is always non negative.

So, how do you ensure that this whole thing is going to be positive you just changes x 2 or you take the absolute value the whole thing which is simply mod x times this so you want that sum to be finite why we call the d of mod x i will come to a bit later. So, let us leave this is leave this aside just for a second just think of this.

So, if this is fine then you say that E of x exists this is just a simple test right nothing more than that. And you can see you can; obviously, see that if I put mod x here no matter even if what minus x naught x naught is still not going to converge it is going to go to infinity right is not it is very clear that there you no longer get positive infinity cancelling negative infinity both negative x and positive x you get plus infinity and it will happily diverge so no problem so in that sense.

So, you do not get this ambiguity of whether it is going to you know if it is going to 0 or not. So, this is just a warning right that I mean this is an important enough situation I mean the thing which I have to say so basically you know that and the fact is that we do deal a lot with Cauchy PDF's right the Cauchy PDF has such as a has a very strong tail right which is falling off only as 1 by x squared and no PDF which has a tail which falls over 1 by x square has a mean value whether it is one sided like this or two sided like this is clear that is all I want to say.

So, basically what does the take home message from all this the PDF should die down somewhat should I faster than 1 by x square should go to 0 faster than 1 by x squared; 1 by x squared may be sufficient for the PDF itself to be integrable, but when you multiply by this x the resulting function is no longer integral I mean this is not the integral is not finite if you go all the way to infinity.