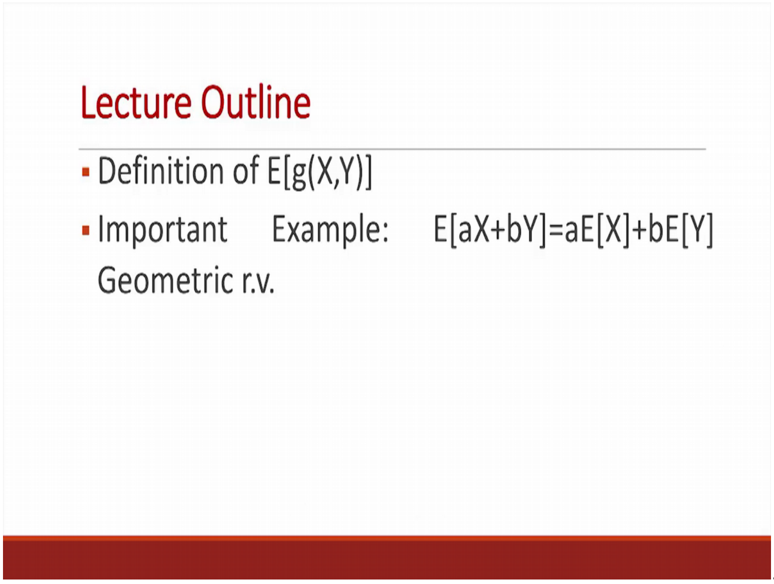
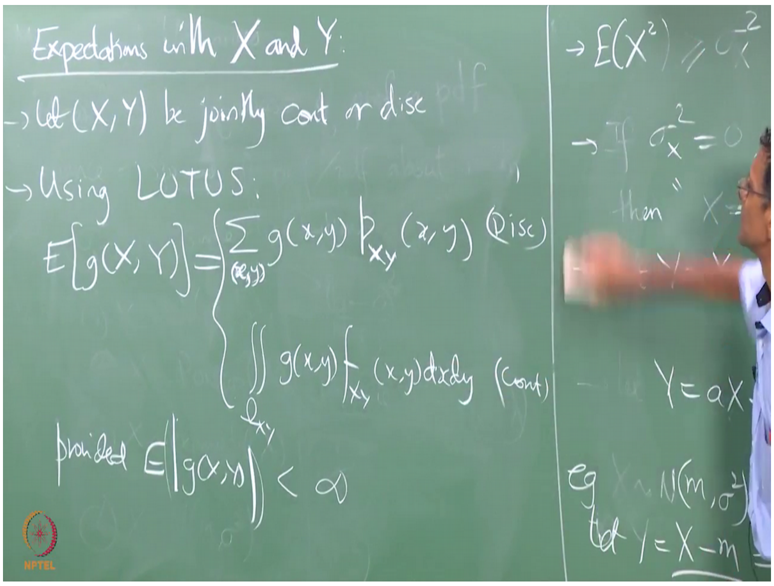
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 69
Expectations with Two Random Variables

(Refer Slide Time: 00:13)



(Refer Slide Time: 00:17)



Supposing, I have two random variables jointly defined with the p joint PMF or joint PDF, or discrete does not matter directly continuous or discrete remember in the case of a discrete PMF the computation always involves two sum and in the continuous case. You have to do an integration right, then I can directly write out using the yellow lotus theorem write the ordered lotus itself is a law right.

So, using lotus this E of a single scalar function of two random variables there so; so, now, I am just taking a scalar real valued function g of X, Y which could be X plus Y could be max X, Y could be anything as long as this single scalar valued function single real valued function right. So, this is basically in the in the discrete case basically this sum.

So, over all allowed pairs x, y in omega x, y I think this may not come. So, clearly in the recording, but cannot help it. So, x comma y so; that means, that you are doing the integral, I mean the sum over all points in the in the in this in the space omega x, y then of course, within for the continuous case what do we do? How do I write it out?

Student: (Refer Time: 02:39).

Right. So, g of x, y.

Student: (Refer Time: 02:44).

The joint in PDF, again the right these quantities on the right hand side make sense if and only if, they absolute value what you get by putting the absolute value the result is finite. In other words provided E of mod g X absolute value of this function is finite. So, of course, there are some other random variables who are pairs which this is not right; the most general formulation, but for more general formulations, we need to use some crazy notation and all that which are which have not introduced at all.

So, as of now we have only stuck to E both of them either jointly discrete or jointly continuous. So, we only give the results for these two cases: likewise, I already mentioned in the; the case of single random variable that the formula said we have written do not easily apply to the so, called mixed which is I do not like the term, but I do not know what else to say I mean not random variable x with the CDF with jumps right you cannot. So, easily write down a simple expression for the mean and leave you cannot do it for the mean then, it means no variance also becomes quite a x E of x square also has problems.

Likewise, here again right you can come up with the random variables which are neither purely discrete nor purely continuous right and it is you know in only some cases you can easily write down the an expression for E of x y, but the, but in the all the examples we have done we you know for jointly the c jointly continuous this these expressions will be valid.

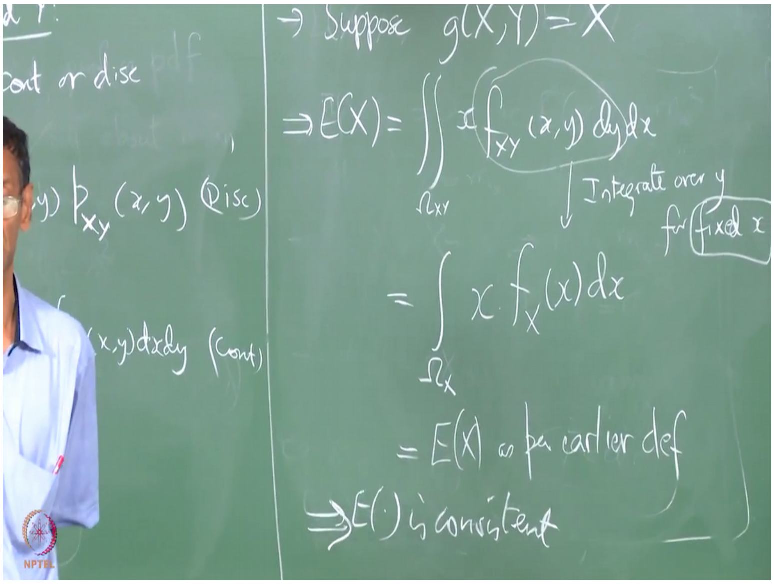
We have done. So, many for example, trinomial for the jointly discrete case the jointly discrete turns out is a more interest easy to understand and appreciate always right. So, you should always base your understanding; what happens is? Multiple random variables you know you think about the joining discrete case you only have a set of points and then try to generalize to this.

So, let us see now again some properties of writing this average again the meaning of the expected value is identical to; what we had in a what we had in the single random variable case right. So, this will be or it tries to model an average of this function g over many independent observations of the pair x, y x and y can be the height and weight for example, I correlated and you could model that is a continuous pair.

So, now if I you know I am not sure; if I can call the set of people in this room as independent, but anyhow so, if you wanted to average all those values of right some func[tion] some function g of x, y some linear combination of height and weight for example, you could in you know the what you get here is; basically a model for going and looking at each person and doing that calculation of that persons an averaging of the whole lot ok.

So, the meaning of E is always an average a statistical average no matter, how you define it. So, that brings up the question is this E. Now, that we have defined here is it consistent with me supposing you have only x; what if g of x, y is x, y as we said as we saw in the in the context of transformations right you can always define g of x, y to be x.

(Refer Slide Time: 06:53)



So, suppose supposing I just take this is this definition consistent with what we wrote earlier.

In the sense of doing do they both give the same E x; what would be our answer to that.

Student: Yes (Refer Time: 07:19).

Yes, they give the same answer why? Because if I do this for example, let us say I take the continuous case. Now, let us say I write this as dy dx not just dx dy say that there is a nefarious reasons behind us dy dx; why? Because I want to integrate over y first and then over x; what is the value of this integral? If I first integrate over y keeping x fixed and; obviously, I have to integrate over limits that are imposed by that value of x. So, this will be an integral over omega x of what x times the marginal. So, you integrate over y keeping x fixed.

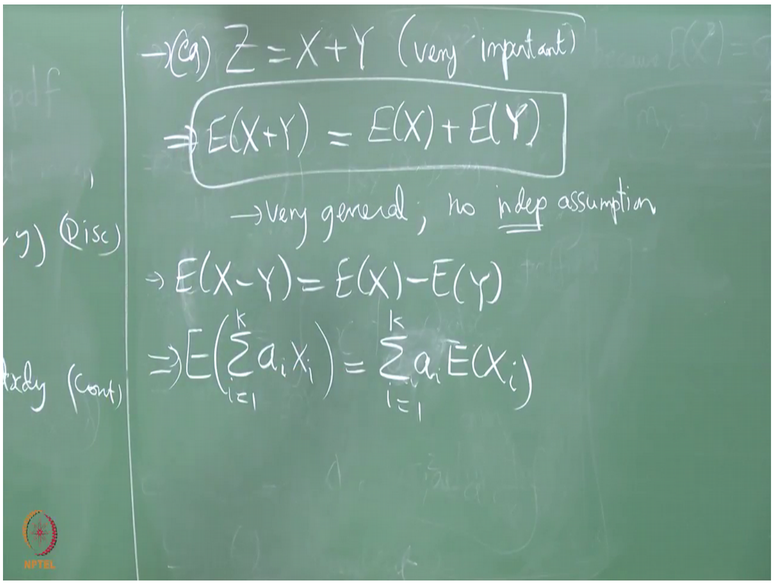
Remember, I said my very first lecture that you must be very clear about multi variable integration right and I think you have done it enough examples in some home works to appreciate this. So, without taking any specific examples I can always go from here to here not can I not if I integrate with x fixed that is this in there is an integral of this for fixed x is always the marginal density by definition and what is this E x as per earlier definition ?

So, the E operator is always consistent and the same thing can you can apply to higher density you know higher dimensional density functions also. So, in other words I never I mean that the scope of the E operator is basically defined by the what you know by, what is operating on if there is only a single random variable inside the g then you understand that to be basically involving only the statistics that random variable if you have a function of two random variables then it is the joint distribution of those two random variables and so, on.

So, E you write the same E no matter; what I am not going to write put a subscript here on E? Now maybe later on in some cases; I might write a subscript to not in this not for expectation of this kind of conditional expectations just to keep things straight I might put a subscript here to indicate the scope was at E, but for this it is not required ok. So, E is consistent that we are very thankful for this because, if it were not then a whole theory pretty much will collapse right the E would be would be I would almost be meaningless if it were not consistent.

So, no need for at least at this stage no need for subscripts; and if it is jointly discrete also I do not think there is any issue here right you they do not even have an integral to worry about you just have x of p x, y you sum over y you know automatically you get the density or the PMF of x at that value small x so, then exactly the same as before ok. So, then what is there? What is the right we did this only to show the consistency ok. Now keeping that let us take some other example of x g of x, y the most useful example of are among the most useful examples of g of x, y is x plus y always right.

(Refer Slide Time: 11:40)



Let say; this is extremely important because this is a kind of linearity of the E operator that we did not encounter in the first version which is with the single random variable. So, what happens if I substitute x plus y here or here; what do I get please do it and tell me?

Student: Expectation.

What do I get? I know I do not know I you are saying something which I cannot understand.

Student: (Refer Time: 12:26).

Ha no.

Student: (Refer Time: 12:30).

Expectation of x plus now I heard you expectation of x plus E of X plus E of Y crucial result which is not which holds universally as long as both of these are finite. So, the linearity of E takes on a yet another meaning; now you know it is not just g 1 x plus g 2 x; that we did earlier instead now we have X plus X plus Y 2 and to totally different random variables and you still you get.

So, for those of you that; are still saying. So, all I am saying is you take either formula and substitute x plus y, it turns out by the distribution of multiplication over summation right it is x times this plus y times this and then the summation if you do it twice once for x and once with y as I just did I rest right you will get E X in the first term and E y in the second term. So, please go and work out the details get for this. So, you do not require x and y to be independent. So, it is very general no independence assumption right likewise we also have E of X minus Y to be E X minus E Y and this; obviously, is m x plus m y, because right there is only one E X and that is m x right ok.

So, how do I now though I mean put this result to work what do I need yeah um.

Student: (Refer Time: 14:32).

Yeah of course, we are going to extend this to any linear combination E of sigma ai if that is what you want a I let us say I have a collection of random variables yeah; now that you have mentioned it let me let me write it. Supposing I have let us say some k random variables or something they can be dependent independent I do not care. So, this will be equal to sorry this weighted combination right.

So, I do not what I am trying to say here is this may look like a complicated thing which is. In fact, a complicated thing if you look if you want to calculate the distribution of exact distribution of this right it is not easy; in general to find right the pa PDF or PMF of this quantity of this weighted sum.

Even here you had to do a complex term I have independent you were to do some convolution and stuff like, that not the simplest of things to do to determine actual distribution, but it turns out if you just looking at the mean and the variance. The variance also is not too difficult to calculate alright turns out it is just the mean here; the mean of this weighted combination is just this weighted summation of the means and no independence. Again is required x 1 through x k can be an arbitrary collection of random variables on it on some space right ok.

Now, let we you know done this for the; and also one other thing right this quantity depends only on the means. So, this it turns out in many in practice this means are easy to estimate, because the means only involved just the PDF of PMF of just that one quantity random variable and not the others. So, in multiple ways this result is very useful.