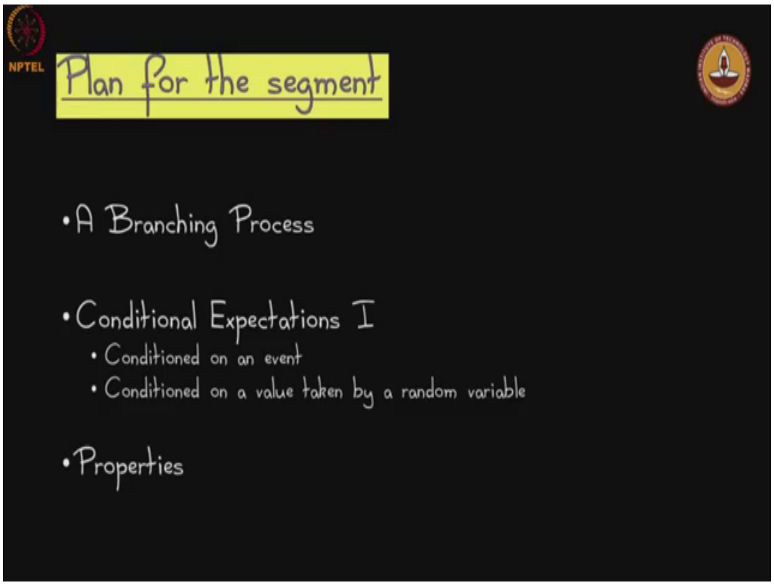
Probability & Computing
Prof. Jhon Augustine
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Module – 02
Discrete Random Variables
Lecture - 11
Segment 3: Conditional Expectation I

Let us start the third segment in our second module. So, far we have looked at the expectation of a random variable.

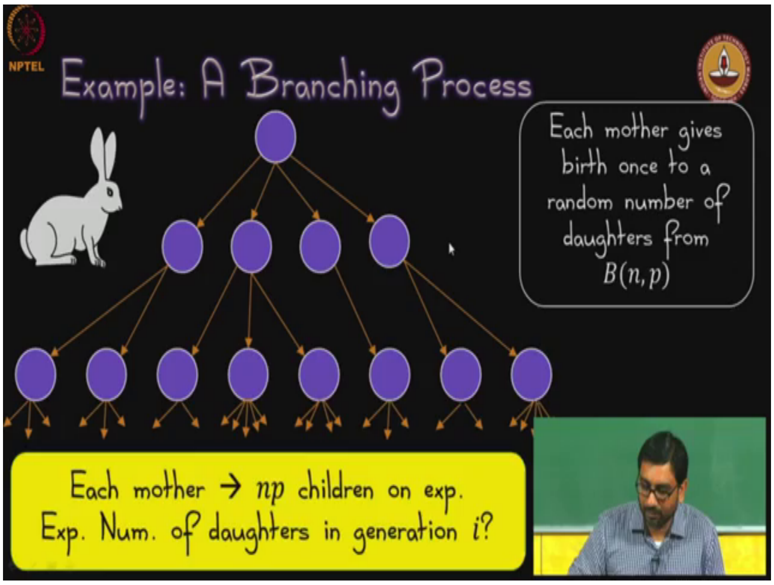
(Refer Slide Time: 00:25)



What we will do this time is look at extending the idea to include what is called conditional expectation. And there are two variants of this notion of conditional expectation. So, we look at the first variant in this class and this segment and then we look at the second variant in the there are related notions. It is just that the same term conditional expectation is used to denote two different slightly different contexts concepts.

Let me motivate that by looking at what is called a branching process. It is a very simple process, but it is a very very useful process that shows up a lot.

(Refer Slide Time: 01:00)



And so I motivate this in the context of just trying to understand the population of a species. So, for example, let us take rabbits, and normally when you try to model populations a lot of times what they do is they simplify it and just focus on the females in the species ok. So, now there is you start off with one mother if you will and so you make the assumption that each mother gives birth once and that is this again an assumption, once to a random number of daughters.

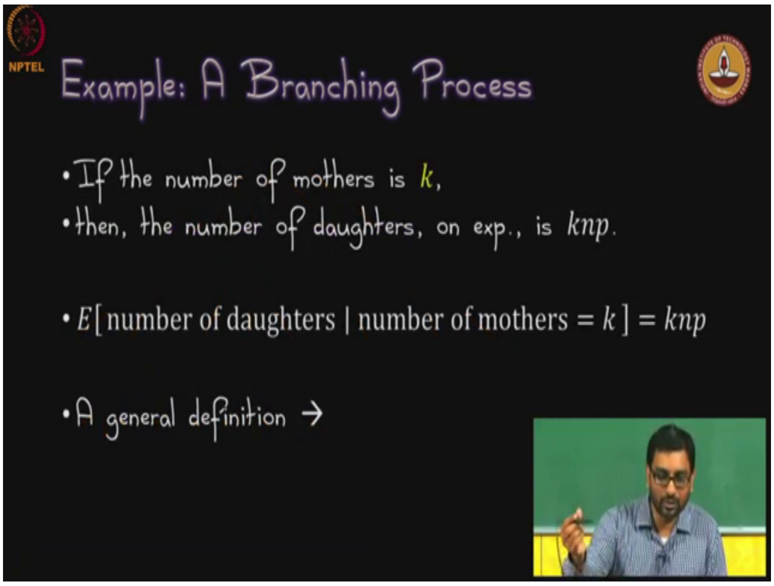
From and this random number is drawn from the binomial distribution with parameters n and p. This is just of course, a modelling assumption the real world, may not work exactly this way, but we are just going to use this to understand how things work ok.

So, let us say the first mother gave gives birth to some 4 daughters and in their generation those 4 give birth to their own daughter some 3, some 2 and so on and so you have generations where they give birth 2 daughters. And as you can see this is a very; obviously, something a population that branches out. So, it is of often called a branching process and shows up a lot in evolutionary systems even in computer sciences shows up when one function can spawn other functions.

So, what is clear is if you fix a mother you know the expected number of daughters and that is n times p, maybe already so. But what is not clear is what is the number of expected number of daughters in some generation i. So, in the in the first generation we saw we know that, but how do we extend that to an arbitrary ith generation that is the question.

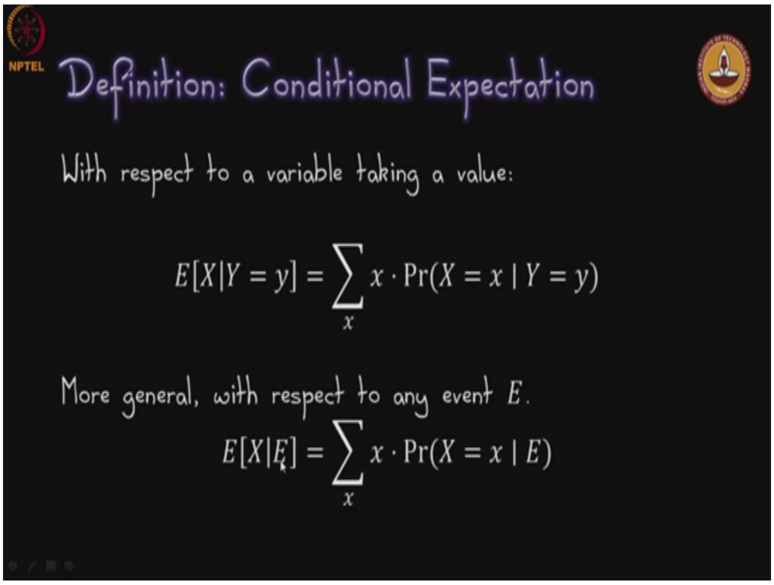
And of course, intuitively it should be n p in the first generation, n p the whole squared in the second generation and n p to the whole to the 3 in the third generation and so on. There should be some intuition here, but we do not have, what we have studied so far does not guarantee that I mean there is a formalism does not work out. The intuition might work out, but the formalism is what we need and that is what we are going to develop and this and the next segment.

(Refer Slide Time: 03:30)



So, let us see for example, if you are if you know that at some particular generation there are k mothers ok, then the number expected number of daughters is k times n p, this is applying linearity of expectation over the k mothers. So, I can write it this way expected number of daughters given that the number of mothers is equal to k in equals knp ok. And this is something we just seem to have made up, but actually this is there is a formal definition in general, so you take any two random variables X and Y you can have what is called the conditional expectation of x conditioned on y equals a particular value.

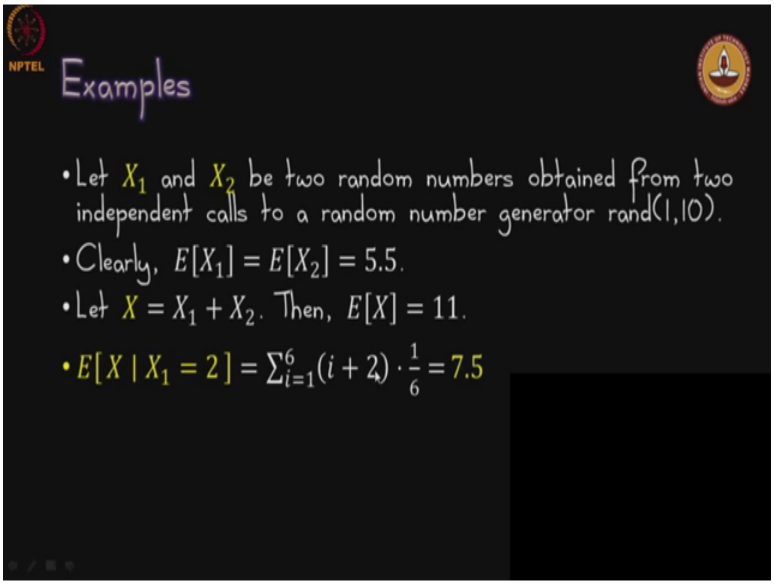
(Refer Slide Time: 04:01)



So, in this case in the previous slide Y represented the number of mothers and X represents the number of daughters ok. So, when the number of mothers is a specific value then you can ask what is the expected number of daughters. And as you would expect the formula is very very straightforward. If it were if it did not have a condition you be summation x, x times the probability of x, but because as the conditionality over here you add the conditionality to the probability as well.

Essentially what are you doing we are taking the sample space and saying let us limit our sample space to just this one portion, where Y equals the little y that is all we are doing. And you can generalize this a little bit you can remember Y equal to y this is just next an event, unfortunately we reusing E here, ok. So, maybe think of expectation of X given some other event F can also be written in a similar fashion. So, apologies for reusing the letter E here, ok. The first type of conditional expectation where you are conditioning on a particular event.

(Refer Slide Time: 05:33)



So, let us look at an example some examples. So, X 1 and X 2 let us say there are two random numbers obtained from two independent cause for a random number generator which range with range from 1 to 10. So, or you can think of it as a 10 sided die if you will. And you have expectation of X 1 equals expectation of X 2 equals to 5.5. Let us look at the sum of those two outcomes. So, X equal to X 1 plus X 2, expectation of X we know by linearity of expectation is 11.

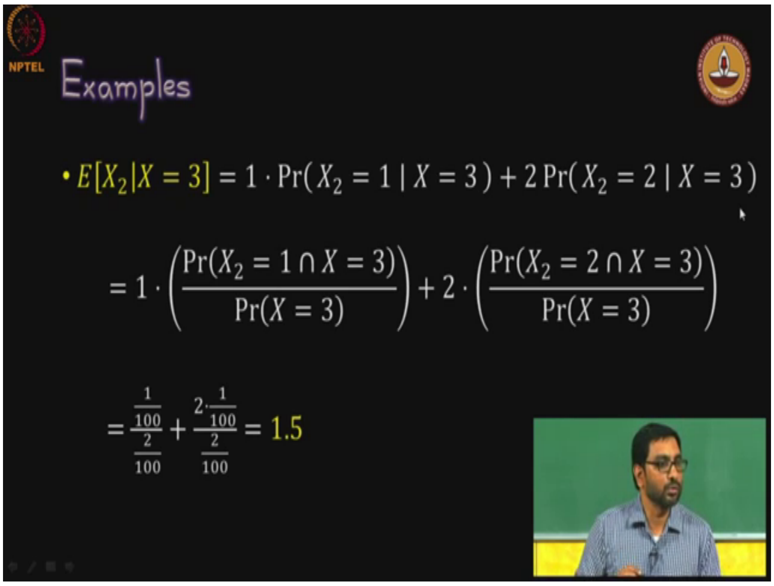
Question is what is the expectation of x given that the first random number generator, generated 2. And what is that? Well, now again you can in this case what we are doing is summing over all the values. So, what are we doing over here X 1 is fixed at 2 ok. So, the only variable is really X 2. So, what, how are we applying this? So, we are just varying the values that X 2 can take ok. So, that is ranging from 1 to 6, but then here internally we are considering what is the outcomes its and this the first outcome is fixed at 2, the second outcome is the value I and that will happen with probability 1 6.

Remember the whole sample space is limited to just the second random number generator. The first one is fixed at 2 ok, so that is why you have a one over 6 over here which is which comes out to be 7.5 which is making a lot of sense because the first one when its fixed at 2, the second one contributes 5.5 to the expectation you get 7.5

Student: (Refer Time: 07:24).

1 over 6 over here, 1 over, oh 1 over 6 because now because the first random number generator is fixed at 2, the randomness only depends on the second random number sorry that is an error. So, that should be 1 to 10, I think I just flipped in my mind from a random number from 1 to 10 to die in my head. So, while I was doing the slice. So, thanks for pointing that out. So, that is an error. So, that has to be updated. So, that has to be worked out with this going up a 10 and this also will have to be here 10, but if you work it out I think the final answer based on the intuition is correct. So, sorry about that, but the final answer is still correct.

(Refer Slide Time: 08:10)

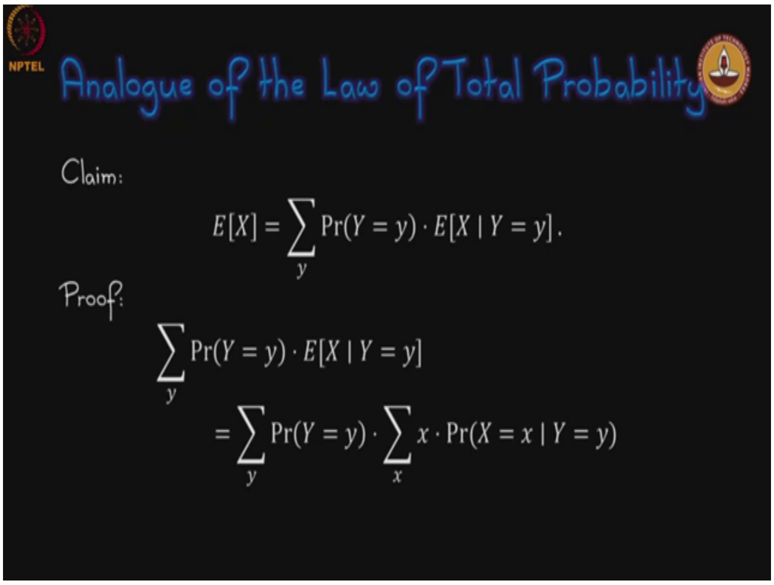


Another interesting various, so let us hope I have not made any mistakes here. Let us see expectation of X 2 given that X equal to 3.

Here what are we doing? So, now, this capital X is the summation. So, here we are not, we are conditioning on the fact that the sum is 3, if the sum is 3 we are asking what is the expectation of X 2, but now X 2 can only take a few values, it can take either the value 1 or 2 it cannot take the value 3 or more. Why because if it takes the value 3 or more it will have to be then added to the first item, first value and then if this conditioning will not work, ok. So, you are left with either 1 or 2. So, one times the probability that X 2 is value is 1, again with the conditioning time plus 2 times the probability of X 2 taking the value 2 with the conditioning.

And if you work that out just apply the formula for conditional probability it works out to about 1.5 and that should make intuitive sense to you. Given that, the sum is 3 the X 2 will either be 1 or 2 with equal likelihood and so 1.5 seems to be the right answer and that is what we get through formal verification of this.

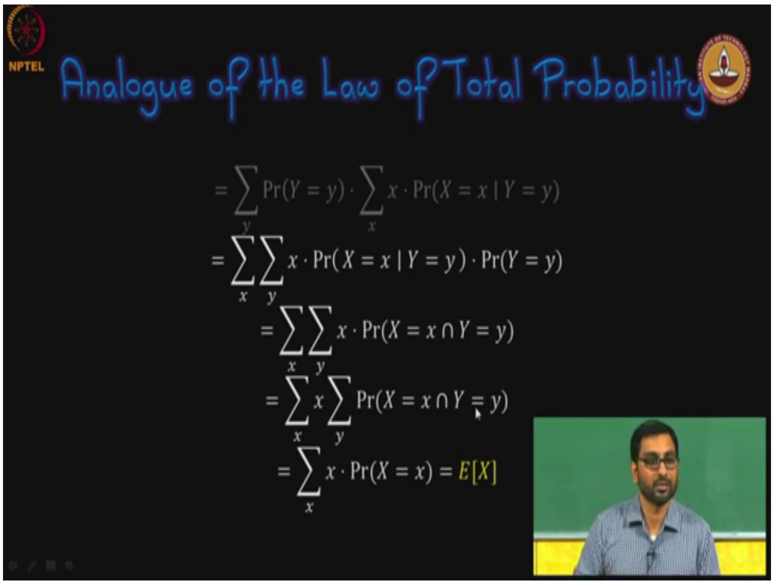
(Refer Slide Time: 09:38)



So, some quick properties of this, so basically this is just repetition of ideas that we already know, but we want to view it through this conditional expectation I mean notion. So, we know this notion of law of total probability. Does it apply when you view it from the conditional expectation point of view? Yes, as it turns out. So, the claim is this the expectation of X is equal to the sum over all values Y probability of y equal to y. So, this basically covers the entire space times the expectation of X even when you have conditioned according to y.

So, how do we show this? Let us just look at the start from the right hand side and here the what we are doing is from the right hand side we have this expectation term we simply expand that out, apply the formula.

(Refer Slide Time: 10:33)



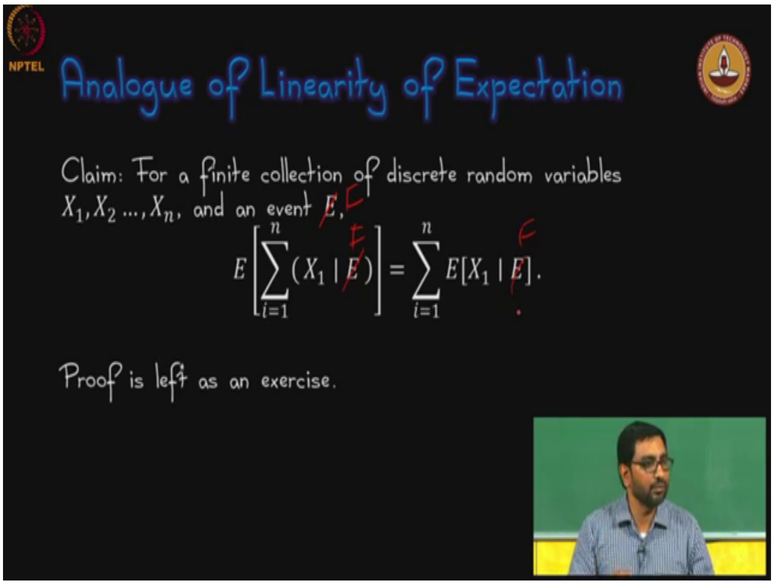
And from that expression we collect the summations. So, we get x summation x, summation y and everything is inside of it. But now what do we have over here is probability X equal to x, given Y equal to y times probability of Y equal to y that is of course, probability of X equal to x intersected with Y equal to y, conditional probability formula.

And so, now, what we are doing this x does not depend on y, so it can be brought out and what we have here is summation y and we have X equal to x intersected with Y equal to y. And what is this?

Student: Law of (Refer Time: 11:15).

Law of total probability, that is going to be just probability of X equal to x. So, you get expectation of X, ok.

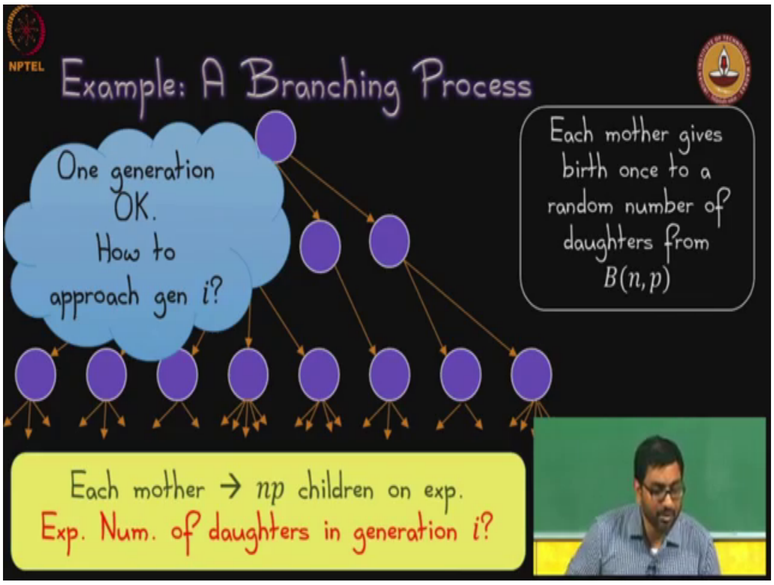
(Refer Slide Time: 11:23)



So, law of total probability. What about linearity of expectation? Again that will also hold. So, I will spare you the proof, but essentially what are we asking, so under again I have made this. So, let us assume this is some other event F ok. So, the expectation of us, the sum of several random variables each conditioned upon some event say F is equal to the sum of their individual expectations each conditioned on the same F. So, this as you would expect ok. You can work that the details out.

Basically it this is a very intuitive thing to state right. So, all that conditioning does is it, it redefines the sample space, it carves out a new sample space, but still is a sample space never the less and from which a property space is defined. Therefore the expectation should, the linearity of expectation should hold that is the intuition here, ok.

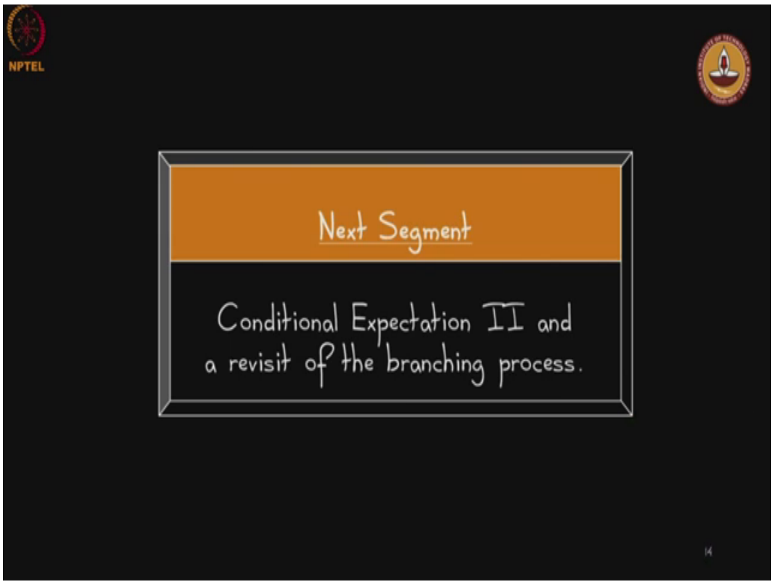
(Refer Slide Time: 12:22)



So, back to the branching process. What does this conditional expectation give you; when it when you can condition on the number of mothers you get the number of daughters. So, I what this tells you is if at the I minus 1th generation if you will, if there are some k mothers then in the next generation there will be k times n times p daughters. So, it gives you a step in the induction if you will, but it still does not tell you what happens over the course of the entire populations evolution.

So, we are still sort of trying to understand what the expected number of daughters in generation I will be, ok.

(Refer Slide Time: 13:18)



This question this red question is still not addressed yet, we know it conditioned on the previous generation, but we do not know it in general, ok.

So, that is still something that needs to be addressed and that is what we are going to address in the next segment where we will talk about the second notion of conditional expectation. And we will revisit this branching process and try to understand that the answer to that question, ok.