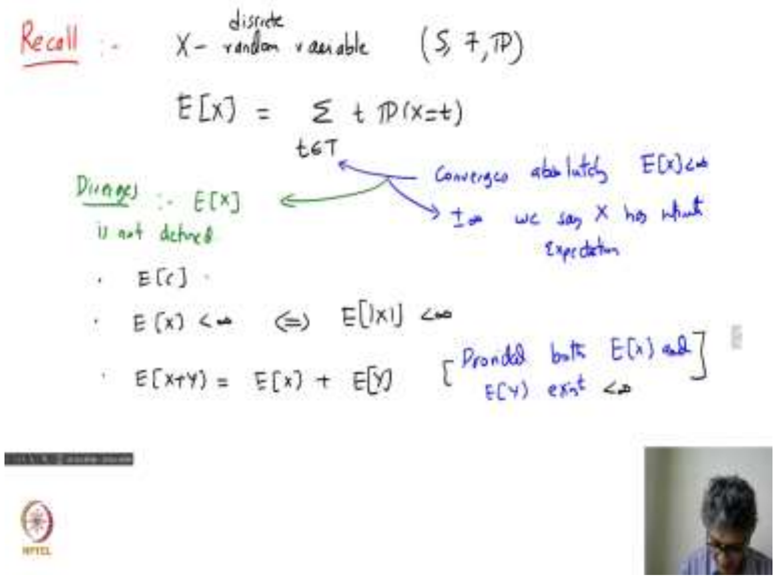
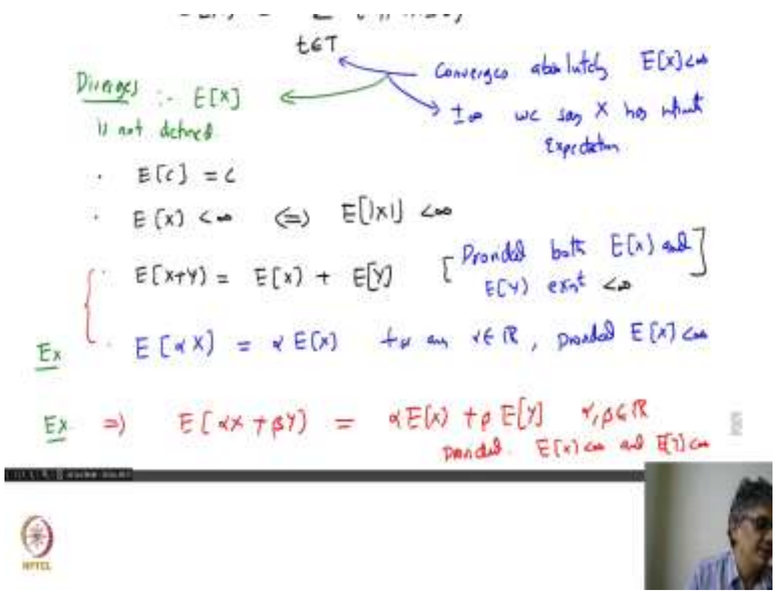
**Introduction to Probability – With Examples Using R**
**Professor. Siva Athreya**
**Theoretical Statistics and Mathematics Division**
**Indian Statistical Institute, Bangalore**
**Lecture No. 23**
**Expectation Independence and Functions**

(Refer Slide Time: 00:21)




We start again. So, I just want to quickly recall what I did last time, so (())(00:22) bigger one, so it is what we be doing? We were trying to understand, so X is random variable on the sample space S and events T, probability P, so and then the we decide, we so that is actually a discrete random variable as far we have concern and for that we defined expectation of X as the sum over t and T chance that X takes the value of t times and this we interpreted as this T is countable, so this only fit converge absolutely we say E of X is finite.

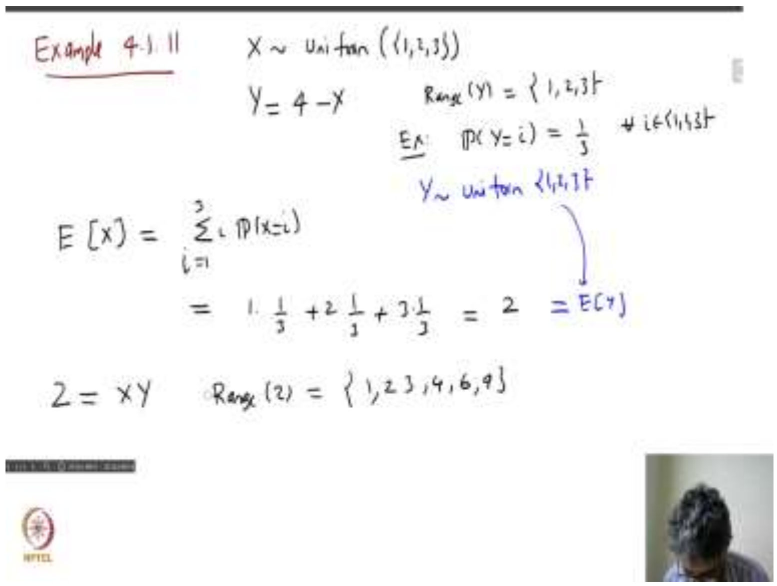
If it diverges to infinity plus or minus infinity we say X has infinite expectation and then if it does not do either if it is diverges somewhere else the series then we say that E of X is not defined. That is what we had done last week, we sort of stabilizes E of X and then we had shown several properties of E of X, one was that if we have E of X is finite is the same as saying that E of the absolute value of X is finite.

So, it is a very important property of E of X and we had also said that we have shown a particular property of sums of random variables, we had said that E of X plus Y we have shown was equal to E of X plus E of Y and this is provided so the caveat here is that provided both E of X and E of Y exist and that is finite. So, let sum of them follow define, we also showed that E of the constant random variable is just the constant and we had not shown this one part which I will give as exercise is the fact that E of alpha times X is equal to alpha times E of X for any alpha real provided E of X is finite. So, this is what we had show last time and we proceed from there and go to the next step.

So, now what we will do now is that we try and understand how far this can go, next example is linear, so this these two facts imply these two facts will imply that E of alpha X plus beta Y is equal to alpha times E of X plus beta times E of Y. So, this is just for any alpha beta, and provide again E of X is finite and E of Y is finite.

So, there is something that one can easily show, so these two things I will leave a exercises I will not try and show them, this is where we were last time? So, the I would like to build on this little bit and this time what E of X actually means for random variable, so the first thing is you know it sums across as linear, so it is a what happens to products? So, let us just see a simple example.

(Refer Slide Time: 05:31)

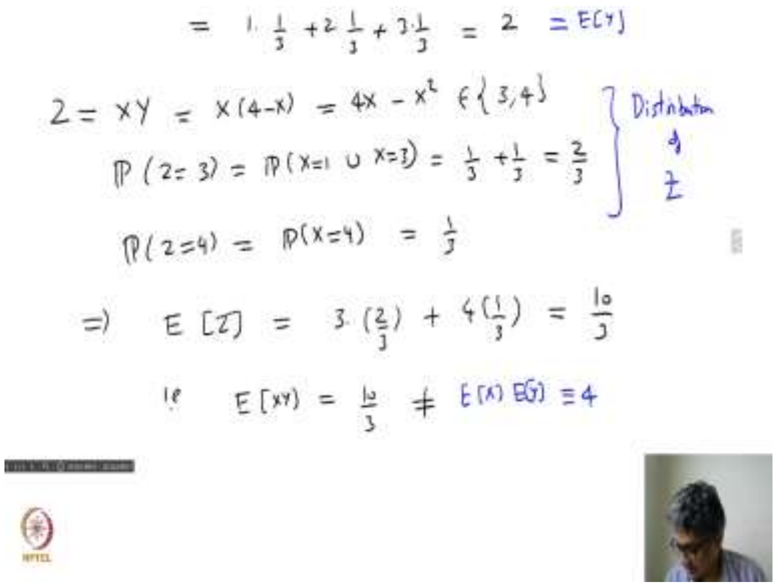


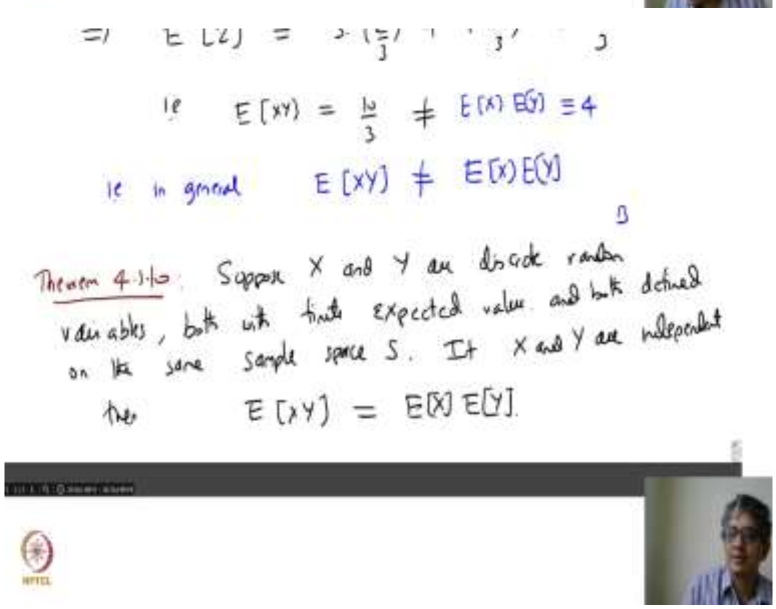

So, say this is example 4.1.11, so the idea is the following that let us say X is uniform 1, 2, 3 and let us say Y is 4 minus X. So, now it is easy to verify that, so let us see so Y takes the range of Y is again 4 minus 1 is 3, it is again 1, 2 and 3 and so also it is a easy exercise to check that probability that Y equal to i is just 1 over 3 for all i. So, Y also is uniform 1, 2, 3.

So, now if you do E of X that is just going to be the sum from i equal to 1 to 3, i times the chance that X is equal to i and that is the same as 1 times 1 1 third plus 2 times 1 third plus 3 times 1 3rd and that gives you 3 plus 2 plus 1, 6 by 3 and that is 2, so since X and Y are the same distribution, this will also imply that E of Y is also equal to, so E of X and E of Y is 2.

Now, again we have seen the expectation sums up, so E of X plus Y is the E of X plus E of Y, but what happens if something like that if I take Z equal to X times Y then what happens? What is X times Y? X times Y is range of Z as is X times Y, so you have take all possible value, so it is 1, so 1 into 1 is 1, 1 is 2, 2 and then 3 and then hitting all possibility (())(08:00) multiply (())(08:01) 4, 6 and 9 this are the value Z can take.

(Refer Slide Time: 8:13)




No, it is not that I mean X is 4 times, so X is this equal to 4, X times 4 minus X sorry and that is the same as 4X minus X square and so this can you can easily check, so the X is 1 this belongs to if X is 1 this belongs to 4 minus 1 is 3, if X is 2 it is 8 minus 4 is X is 2 it is 4 times 2 is 8 minus 4 is just 4 and X is 3, it is 12 minus 9, which is 3 again, so this is all it can take, 3 and 4. What am I making mistake again it is 4 minus 1 is 3.

So, then the chance that Z is equal to 3 is the same as the chance that X is equal to 1 or X is equal to 3. So, that gives you 1 third plus 1 third and that is 2 third. Now, the chance that Z is equal to 4 that happens only when X is 4 and that happens with probability 1, so that is the distribution of

Z. So, now this implies the following that E of Z is going to be equal to 1 third E of Z is going to be 1, Z takes 2 values, value 3 with probability 2 third and value 4 with probability 1 third and that is going to give you 6 plus 4 is 10 by 3.

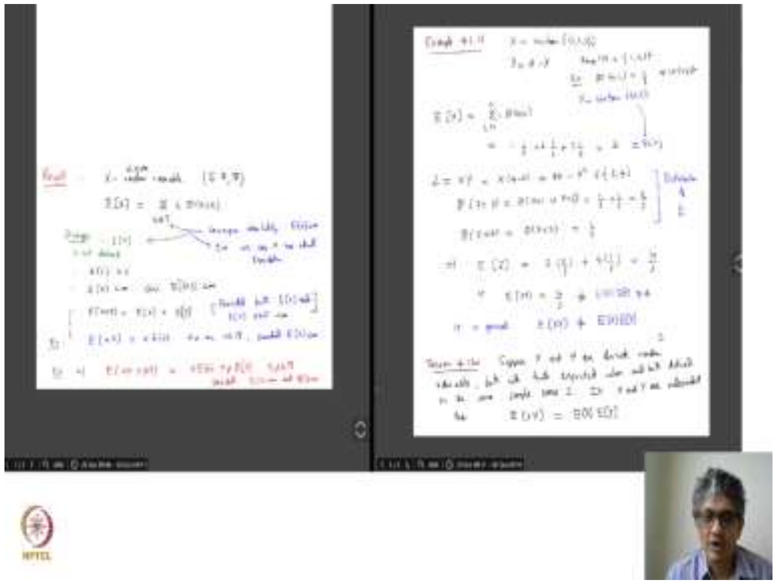
So, now you observe two things now that E of Z is that is what we have shown is that E of X Y Z it is going to be 10 over 3, and that is not the same as let us say E of X time E of Y (())(10:46) both 2. So, that means the expectation does not sort of suitably generalize in general that is in general it is very it is a common mistake people make in beginning of probability that the expectation does not automatically go across the product.

This is also obvious for you if you take two sequence of numbers, you take their means first and multiply them or you take the product and find the mean they did not match, that is how very important aspect of this. So, what is the crucial but this is true in some situations, so I will write the theorem down. It is going to be unique idea and probability that the it is a simple fact moment you see it.

But we it is this thing is not true in general but it is true in the following situation that if the random variables are independent then the things splits up as a product. So, suppose X and Y are discrete random variables both with finite expectation value and let us say both define on the same sample space that is important sample space S. If X and Y are independent then E of X Y is the same as E of X and E of Y.

So, this is a very important aspect in probability. So, the expectation can moves across a product only if the random variable is independent or if random variable, so it is going to equality can also be true otherwise, but in general it is not true like we have seen before. So, we just quickly recap what that the results. We going to split screen again.

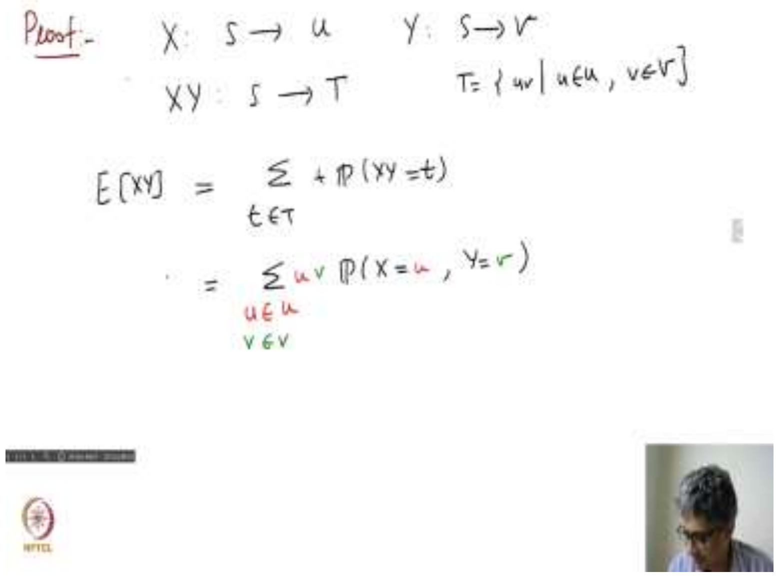
(Refer Slide Time: 13:43)

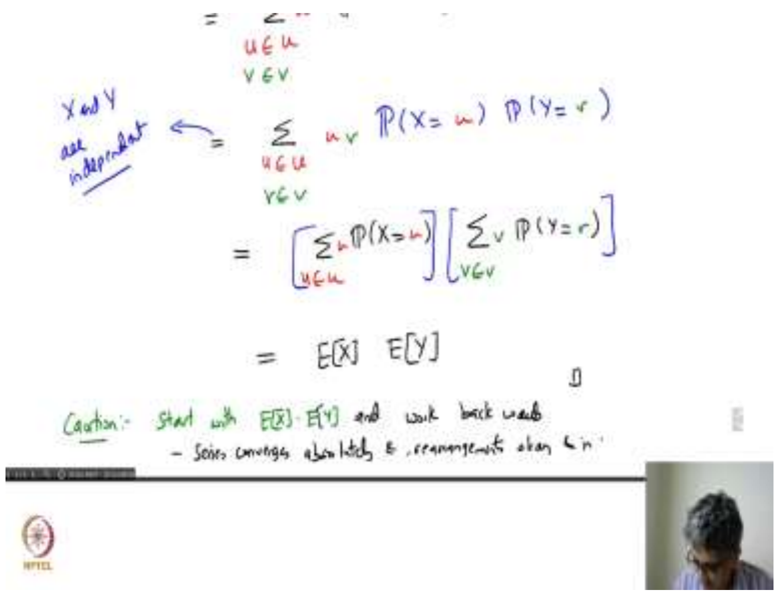


So, here the idea was the following that so we have done this it provide here that E of X is linear and we also immediately check that as a product it does not sort of work of E of X Y is not the same as E of X times E of Y. And that is simple example. And this theorem tells you that if they are independent then the sum will plug in the product though I have just wanted to clarify that it is not an if only if condition you could have the equality be true and the random variable is not be independent we would check that, go along source.

But if general E of X Y is not equal to E of X times E of Y. So, now let me just get back to this idea for the proof this theorem. So, how does I will prove it? The proof is fairly straightforward you have independence crucially, let us do this proof is finished straightforward.

(Refer Slide Time: 14:59)




So, I have X is from let us say sample space S to replace a range is T let us say range is T, let us say Y is from sample space S to let us say its range is u, then XY would take one X times Y I would takes values from S to let us say a sample space let us says let me (())(15:27) let me call this as u let me call this as my v and let us say T where T is the set over u times v such that u is in u and v is in v.

So, now what happens? Now, what is E of XY? E of X times Y is the same as the sum over X, so sorry sum over t in T chance t times the chance that X times Y is equal to t, that is what random variable XY looks like. And now this is the same as the sum over u and u v and v, so let me write this in a different colour, so it will easier, so let us say u and u and let us say v and v it is the chance of u times v and then you will have X is equal to this and Y equal this.

So, I will write that X equal to u and Y and time u here and I will v so v and then X equal to that is hard. Again one has to verify that this going this way you would be little careful because a priori we do not know the series converges, but if you write the proofed on in the backwards way to the clear, so I will just go through it formally at this step.

So, next step is the what you do is you write this as the sum, so now the u and the v split up in the sense the following way, so here use a crucial fact that is u and u, v and v, so you have v and then a u and then here is the fact use the fact that X and Y are independent, so if we use that then you know that this is same as the chance that X is equal to u times the chance that Y is equal to v, that is what you get here.

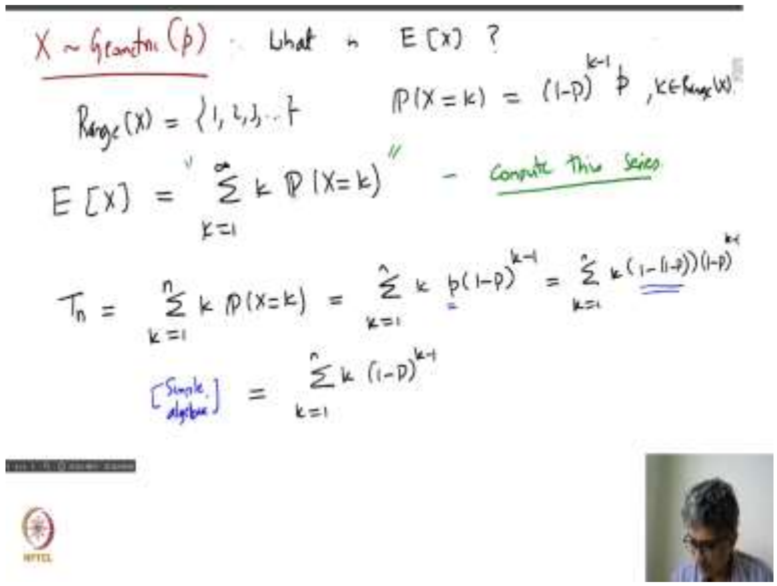
So, now you observe that it is a sum of one term depends only on u, I do not know u and it splits up as plot of two things, so this you think a little bit and make it rigorous is the same as the sum over u and u and sum over v and v. So, sum over u and u and sum and times the chance that X is equal to u and the chance that Y is equal to v.

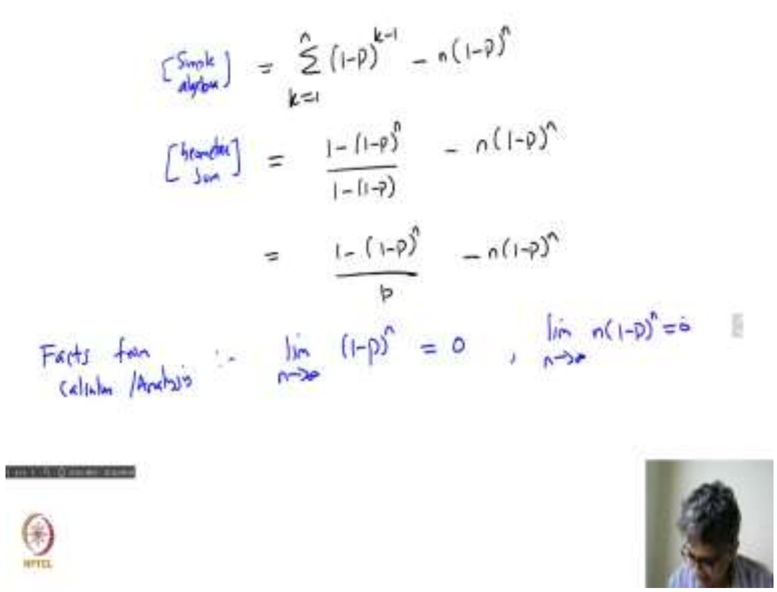
So, here you will have Y and u and here you will have v and v, it splits up the sum splits up above the product that if you think about you just look at and there is a u here, that if you just observe is just the same as E of X times E of Y. So, the just to sure be little bit careful one should not go in this order what you should go reverse order, because you only know that E of X and E of Y exists, so just a word of caution the if you want to be precise about mathematics you start with E of X times E of Y and work backwards for a precise (())(19:21).

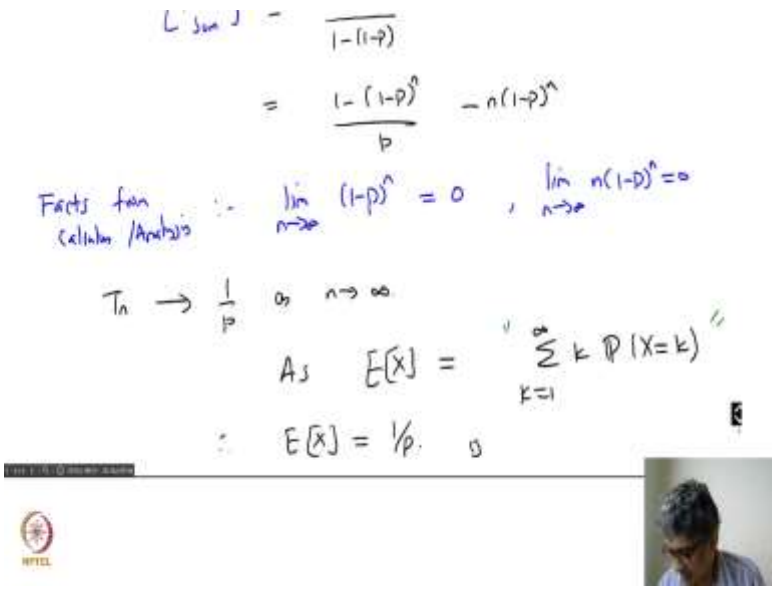
So, you will use what all you use you use that you use series converges absolutely and you will use rearrangements are okay and independence to get to E X one. So, to do it precisely you go backwards you do not go forwards. So, I went forwards sort just to sort of give you idea, but if you will do the precise proof you would start with E of X times E of Y go to this step and now you know they are close absolutely, so you can do anything you want, you can multiply by this up and down what is you got here I use Independence to show this product and same as this joint probability and then you use the fact that the same as this guy and a (())(20:17) That is how you do precisely.

So, that is one important aspect of independence that is the sum has the random variables product random variables the mean splits up as a product. So, couple of things before I serve move on to other aspects of random variable understanding more specific numbers of a random variable to the following. So, two facts I wanted should be illustrate, one is that this is following.

(Refer Slide Time: 20:59)






So, let us do some standard distributions, let us say X is geometric p, so how do you compute the mean of X? So, let us say what mean of X probability? So, what is the means of X going to be if X is geometric? So, we already see if X is uniform we know how to compute it with just finite set, sum is finite so far.

So, geometry let me just walk you through the competition, so what is geometric mean? The range of X is 1, 2, 3 and 4, the distribution of X is the probability that X is equal to k is just you have k minus 1 failures and at the last kth trial you have to successful this is for k minus 1. So, now this implies so what is E of X means? So, E of X traditionally E of X by definition is just going to be the sum over k equal to 1 to infinity k times the chance that X is equal to k. That is what it mean.

So, but we had to make this competition precise, so I will just do the series competition the precise way, so how does one do this? So, first one defines T sub n as the sum, so what am I going to do? I am going to make I am going to compute this series. So, what I do is T sub n is the same as k equal to 1 to n, k times the chance that X is equal to k, that is what one would have to compute first to do the geometry, so what is that going to be? That is just going to be the sum from k equal to 1 to n k times p times 1 minus p to the power k minus 1.

So, this again I do a little trick, I just think of this as the sum from k equal to 1 to n we can write here it is a the sum from k equal to 1 to n k times here I think of this p as 1 minus 1 minus p and then this remain the same as 1 minus (())(23:35). So, all I did was replace this p here by the same quantity, but I do not feel different, but now as I multiply the product out it is an easy exercise by rearrangements all finite sum that can arrange anywhere I want a range and I have a source go pick some and some sense, so I will have k equal to 1 to n k times 1 minus p to the k minus 1.

That is just simple algebra simple. Once you have you are on your way, because now this also is a little bit so this is this is this minus the sum from k equal to 1 to n k times 1 minus p to the power k, I am just multiplying out (())(24:38). Now, what happens to this guy? So, now if you notice it is kind of a alternate (())(24:46) this is k times 1 to the k minus 1 and k times 1 minus p to the k.

So, all you will be left with is the sum from k equal to 1 to n, so each person you have 1 minus p to the power k minus 1 as the only one left. And then the last person here will be left alone. So, if 1 minus p to the power n, n times that is what you get here. So, we have to just move the simple algebra calculus particularly. So, now this is easier the geometric sum we know all do this is just R to the power k minus 1 from 1 to n, the first term is 1, so it is 1 minus 1 minus p to the power n minus 1 the whole thing divided by 1 minus 1 minus 3.

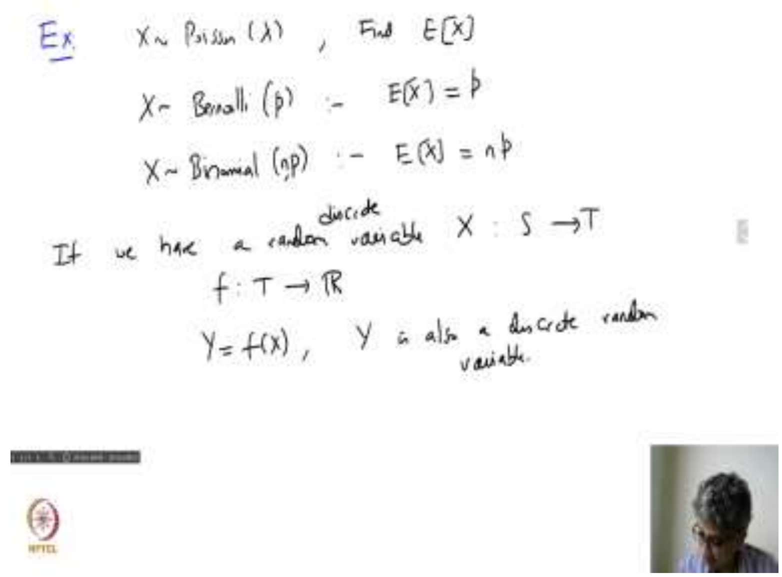
That is what this is. Then minus n times 1 p to the power n, so I will use the fact that I use the geometric (())(25:50) Now, I go to the next step and then this I do a little algebra I have appear the bottom, so I usually read this as 1 minus 1 minus p to the power n minus 1 by p minus n times 1 minus p to the power n.

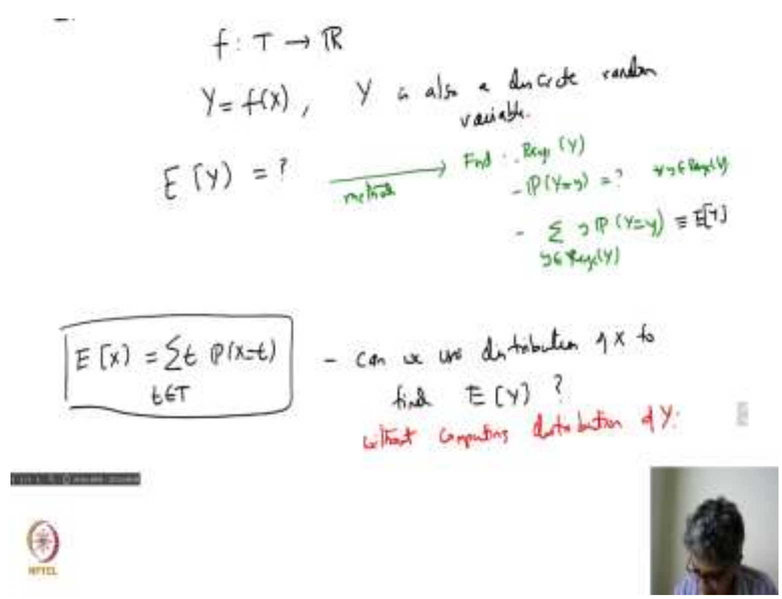
So, now here you have to use two another series facts. So, one fact you can use is this that so the fact in another series you have use or calculus this is the following that limit as n tends to infinity a 1 minus p to the power n minus 1 is equal to 0 as contracting with use, the other fact you need to use is that should I make a mistake in the how many terms here? There are n terms here of course, so the n sorry about that and you also have the fact that limit as n tends to infinity n times 1 minus p to the power n is also 0.

So, these two things I will assume that these two things are true, once you have these two things then what happens is that we see that Tn converges to 1 over p as n goes to infinity, so that means therefore the expected value of X which was the series to the series right here we had the see right here, this is the series this is equal to this, so let me erase this therefore E of X we do this particular series that we had is also equal to 1.

So, as we have X equal to this E of X is essentially limit of Tn’s and that is just 1, so the important competition so do probability properly when we serve be careful you can calculate results that you have to make sure whenever use calculus analysis results, which is the precise about it, you may formally this completely easy to do it requires some (())(28:33) so it is some a way to do all (())(28:38) distributions you can do (())(28:40) you can do in this session, so it is simple exercise to try and check for sure. So, you should try and do exercise when let me try and (())(28:51) down, so it means.

(Refer Slide Time: 28:58)



So, here simple exercise you should try is that if X is Poisson with lambda you find E of X, so at last series competition for all finitely many ones you can easily do this from (())(29:13). So, the X is Bernoulli then you know you can just show that E of X is going to be equal to p, if X is binomial n,p, n comma p and know that E of X is going to be then just check that E of X equal to n times p, there are many ways of doing this one should check that and see.

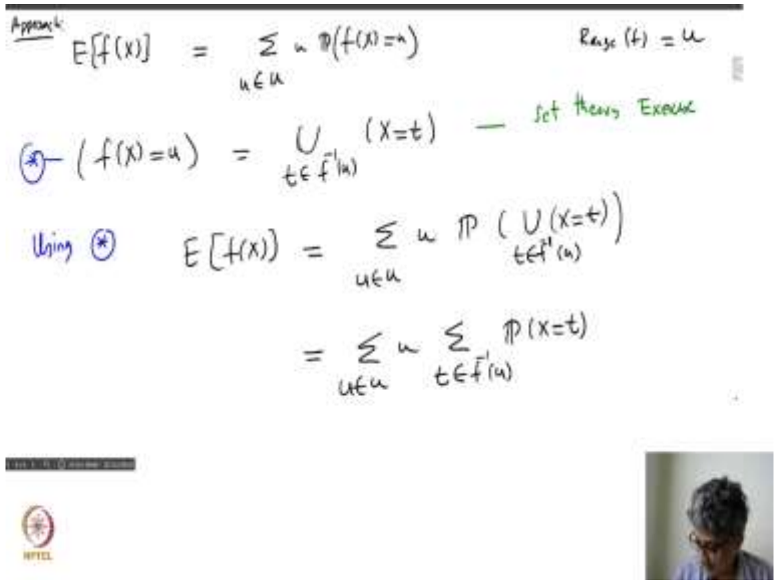
One last aspect of expectations of random variables are such as so now like we already seen if you have random variable X and your function of it then we know how to compute it distribution, invert the function and the try to, similarly if I have random variable X, so if we have random variable X and we have X from let us say S to T and we have function f from let us say T to the real line and let us say X is a discrete random variable.

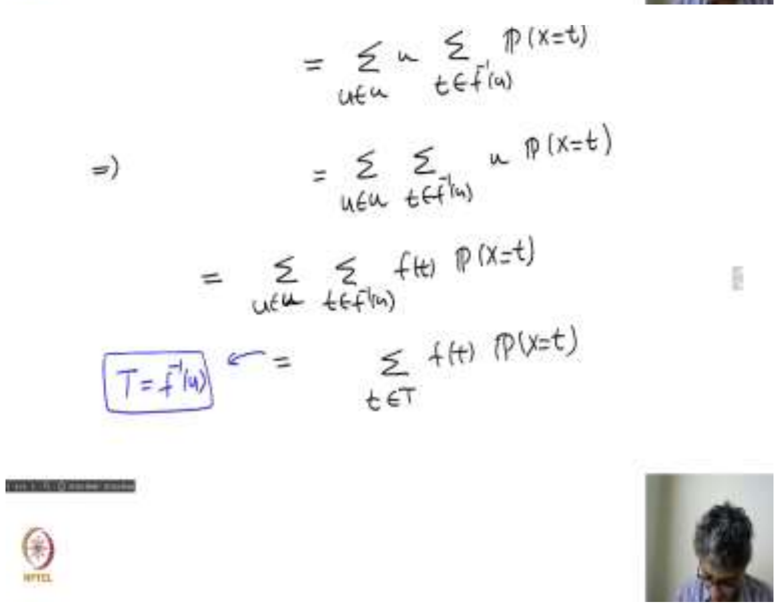
So, E is countable so f will equal to countable values, then you take Y equal to f of X then Y also has Y also is discrete random variable and then now if I have to find E of Y (())(31:16) so one method we know is the following, so one method is a method is following. So, you find a range of Y and then for each Y and Y you do two steps, so one step is you find range of Y that is the first thing then you say that for each Y and Y you try and find the probability Y equal to Y you find for all Y in the range of Y.

Then once you have this step this step and then you do the following then it says even it compute summation over Y in range of Y you do Y times the chance that Y equal to Y. And that will give you the answer of E of Y. This one way of doing this thing. But this may be cumbersome because you already let us try and do it cumbersome do’s again again you want to keep find distribution of Y and then the whole thing, we will just try and see there is an easier way to do this.

So, the so let me try and outliner a sort of an idea before, so the idea is that I have Y’s f of X, so I know E of X is the same as the chance that X is in t, I sum over all possible values in T with E of X, Can I somehow use the distribution of X to find this, can we use a distribution of X to find E of Y? Without going into computing explicitly the distribution of Y? That is the goal, so the idea is that you know that you know one way to compute f of X, E of X to find the distribution E of X and then to do definition, what can we do without them? So, here is one idea, one idea is illustrate the idea and let us see. Here is an idea.

(Refer Slide Time: 33:49)




So, what is let us write the proof down, so what is the what is E of Y? So, E of Y is E of f of X is going to be so this is a approach E of f of X is going to be what that let me discussed is going to be a sum from let us say f is from let us say the range of f, let us call it as u, so it the range of u, so it is going to be the sum I will rewrite it in the following way, sum over u and u, u times the chance that f of X is equal, that is the definition.

Suppose now I do not know the if the things exist, but let us assume everything all sums exist and the all finite, now I come here now observe the following f of X is equal to u this event is the same as the set of all how do I get all sample points when f of X equal to u? What I do is I take the union overall t and T, such that f of X is equal to T, so writing it like this I will rewrite a little differently I just take all t in f inverse of u such that X is equal u, these are both the same events. X equal, they are both the same events.

So, this is a set theory exercise the one or two sort of set theory exercise to establish this fact, that is f of X equal to u is the same as the union over T and f inverse of u such that X is equal to u. Now, let us say I use this fact, so I use this fact let me call it is a star, so using star, so what did you have? We have E of f of X is the same as the sum over u and u that remains the same, I have a you I just replace this probability f of X equal to u as the probability of the union over T and f inverse of u X equal to t, I am not doing anything fancy I am just doing that.

But that is a disjoint union so that we can write it as the sum over the u and u, u time the sum over T in f inverse of u probability X equal to t. Now, this important this implies a very interesting thing, so I have the sum over u and u and sum over t in f inverse of u, so I can now think of my I can flip this sum, if I can do if you allow me, that is the same or I can move the u inside like a think of this as sum over u and u, the sum over t and f inverse of u, I can think of u time the chance that X equal to t.

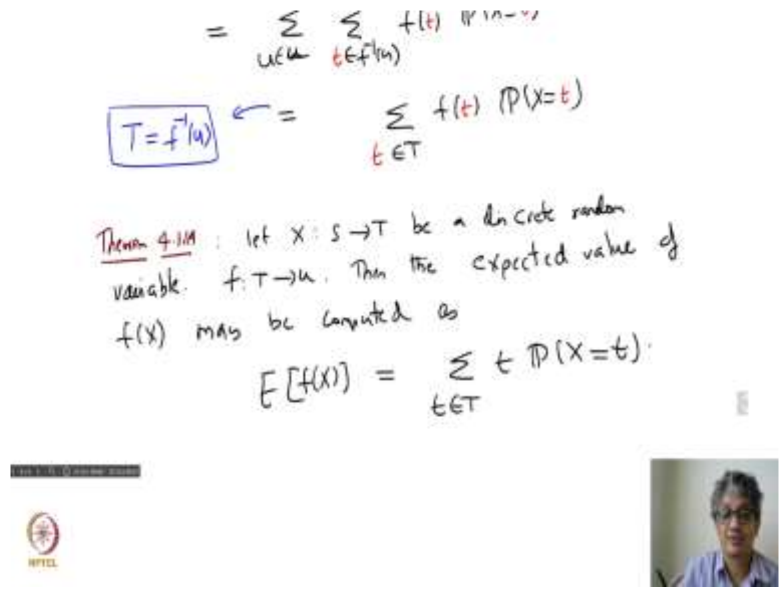
Now, what was u? u was f of t, so I can just think of this as the sum over u and u some over t and f inverse of u, f of t times the chance that X is equal to t. But sum over u and then sum over t f inverse of u that is just one sum, that is just the sum over t and T f of t times the probability X equal to t. So, this step is just the this is the final idea is that all I have to do t is just f inverse of u, that is all I am using.

So, in this step all I am using is the fact is that the set T is just f inverse of u that is all I am using. So, now you see you see what I have done I started off with E of X, I want to note this, but then I came here maybe I will write different colours, so you can be the easier so that I remove all the t's, so I have T and f inverse of u X equal to t, that is what I got this is I got here and this idea I got here.

That is whenever f of X was u that came from some number t over f of t will be equal to, once I have this then I am here again then I come here and the step I replaced the step here I do the fact that this is all mutually exclusive events, so I can just sum of the probabilities of so that is easy to do. Then I come along and I see okay fine I can write I can (())(38:55) sum this way to pull the u inside this so I have this t here I have this t outside so I have t and X equal to t, then I what I do is observe is that u the same as f of t there is no different under the sum condition, let us say (())(39:09)

So, I get f of t here,  so the X equal to t, t in this then once you observe that T is capital T is same as f inverse of u, there is nothing there is no difference. So, all this is just one sub, it is just the sum over t and T f of t, (())(39:32), so what you achieved in this competition the whole way that I do need to go and compute f of X equal to u I can just work with X equal to T to get models. So, here is the theorem that I have shown in this calculation, so let me have this theorem now.

(Refer Slide Time: 40:04)




The theorem is the following that so what we have shown is we actually shown this theorem the theorem is so this theorem 4.1.19, so the theorem is the following that let be the random variable from S to T be discrete random variable you define the function f from T to u then the expected value, 125 of f of X may be computed as E f of X is the same as the chance of t and T t times the probability mass function of X is probability of X, so there is no need of computing distribution of f to compute the distribution, compute expected value of x is just you can use the distribution of X itself to finish the problem.