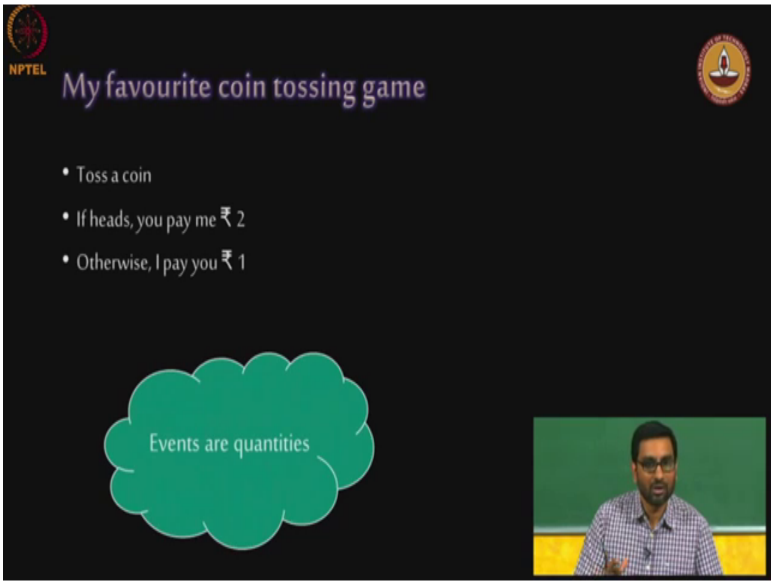
Probability & Computing
Prof. John Augustine
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Module – 02
Discrete Random Variables
Lecture – 09
Segment 1: Basic Definitions

So, the first module we did basics the axioms of probability in the second module we are just going to get one step ahead we are going to try and understand what discrete random variables are and how they help us in the context of computing. So, that is going to be the plan for today to just define what a discrete random variable is and some related notions independence and expectation.

And these are again things that you might have seen or you might be able to relate to, but nevertheless let us try to go through them you know formally and clearly.

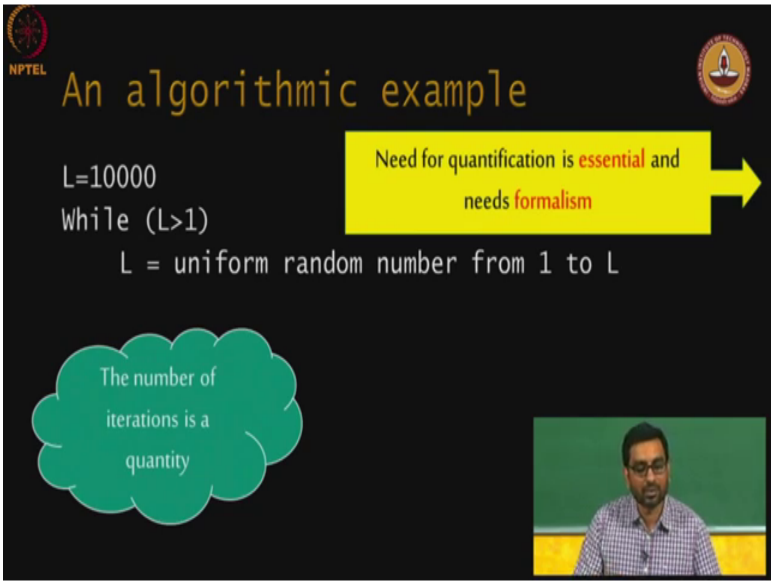
(Refer Slide Time: 00:52)



So, let us start with a with the game I like this game because it is loaded in my favour. So, what you do is you toss a coin and if you land heads, you have to pay me two rupees otherwise I pay you 1 rupee and clearly a very unfair game, but nevertheless let us go with that ok.

And. So, here if you think about it in this game, the outcomes are not just heads and tails.

(Refer Slide Time: 01:21)

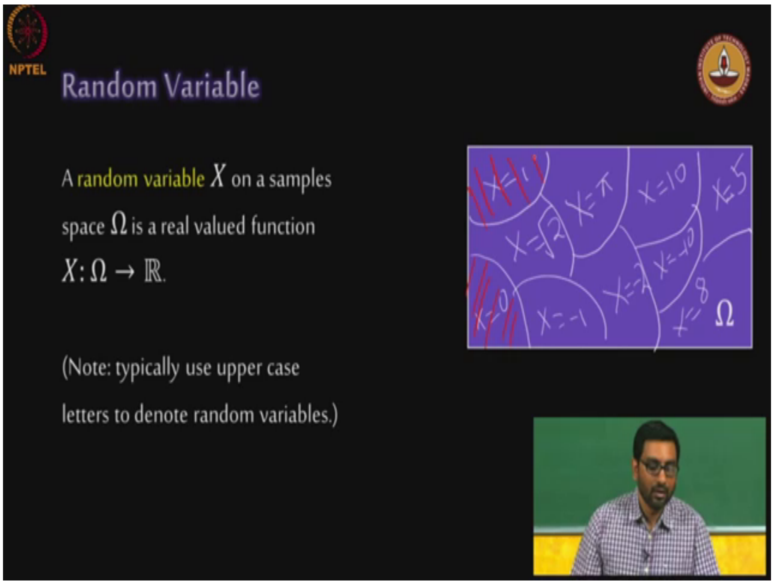


They are quantities let us come to another example where this is closer to computing which is what we ideally ultimately want to understand. It was a simple code snippet if you were you initialize a ran a variable to some 10,000 and put that into a loop and each iteration of the loop we check whether L is still greater than 1 and each time you pick a random number between 1 and L and you assign that value to L.

So, now this is a typical code snippet, you will be interested in understanding how many times this while loop executes and that would tell us something about the running time of the algorithm and clearly that is of interest to us in computing ok. So, this is clearly points us to one thing, it is essential to understand a way to quantify things. Events in the last module we just thought of events as heads or tails picking a card from a sweet and of cards and so on and so forth those are events ah.

But we want to associate the events with some quantities and we need to do that precisely.

(Refer Slide Time: 02:38)



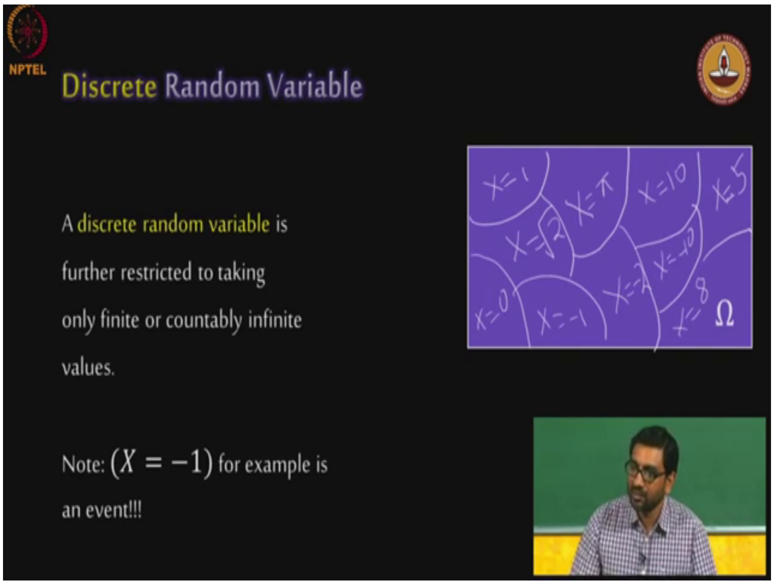
And it is not that difficult the way we do this is the most straightforward way possible, we define something called a random variable for this.

What is a random variable? it is simply a functHion it just takes as input an element in the e sample space and outputs a value a real value for that outcome. So, in other words if you think of this example shown over here, you have the sample space omega and all those outcomes that lead to the that will if you evaluate this using the random variable X ah.

The parts that will evaluate to 0 for example, are this these are all the outcomes that will evaluate to 0 and these are all the outcomes that will evaluate to 1. So, you can see that this is a very natural thing to do just each outcome has some quantity associated with it some real valued quantity associated with it, and we just assign that quantity to that outcome and you get the real you get the random variable ok.

This is all a random variable is and we typically use an uppercase letter to denote random variables and often we are interested in discrete random variables which means that you the number of different values that the variable can take is only finite or countably infinite.

(Refer Slide Time: 03:58)



So, for example, if you go back to this one is this finite yes it just takes two values what about this. So, here we are interested in the number of iterations. So, this is some experiment that is taking place, this code snippet execution of this code snippet is this finite the number of iterations no why not.

Student: I can keep choosing some number right (Refer Time: 04:33).

Yeah. So, you can this can actually be infinite, why because if each time you choose a random number between say 1 and L, if you keep on choosing L you are not going to reduce the value of L. So, if you think about just possibilities, yes it is quite possible that you will be at L for as long as you can think of and that is a essentially the this number of iterations is not finite, but it will be countably infinite ok. So, now, when you think of yes.

Student: (Refer Time: 05:11).

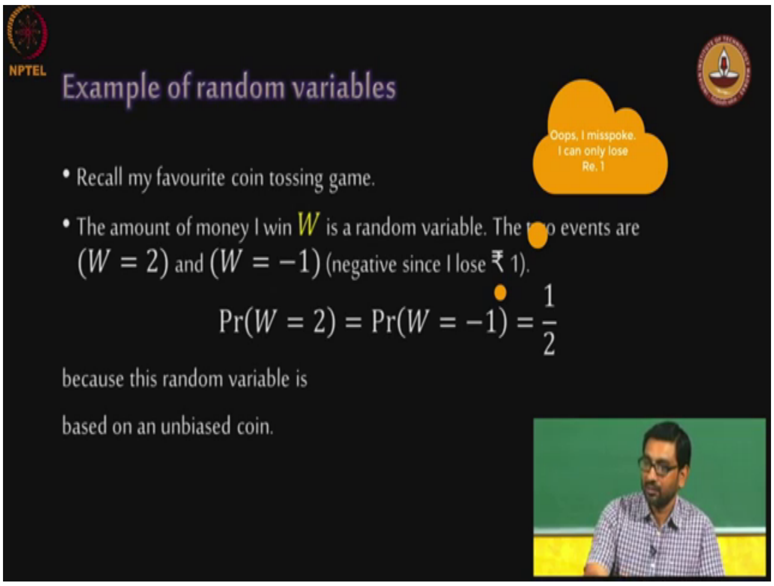
Ah it is a it is in terms of sample space, yes that is still a possibility right it will not be it is not a very likely outcome, but even if you think of uniformly at random it is still with probability well 1 over L you are going to be able to you will be choosing L and thats still a possibility. it is got a positive probability, it is not likely it is not likely especially to keep on choosing L, but it is a possibility yeah.

Student: We keep reducing our sample space (Refer Time: 05:46).

It depends on what you mean by your you precise understanding of your experiments, in this case the experiment is the entire execution of that code snippet ok. So, the sample space is the outcome after the entire code snippet has executed and you are coming out of the while loop and so, then you can ask. So, the sample space really in this case is, did this code snippet execute once or twice or thrice or four times five times and so on those are the outcomes. So, you see there are an infinite number of outcomes and with each of those outcomes we associate the number of iterations as the random variable ok. So, there is an infinite number of outcomes possible keep that.

Back to our notion of random variable it is important to realize one thing. When you take a random variable x and you look at I mean you think of it as having a value say in this case X equal to minus 1. Immediately what you are thinking of is an event y because when you say X equal to minus 1, what you have done is you have isolated your attention to all these outcomes. Those outcomes whose when you take the function X it leads to a value of minus 1 ok. So, going back to our Venn diagram here it is clear that it is any vector ok. So, that is something that is important to keep in mind.

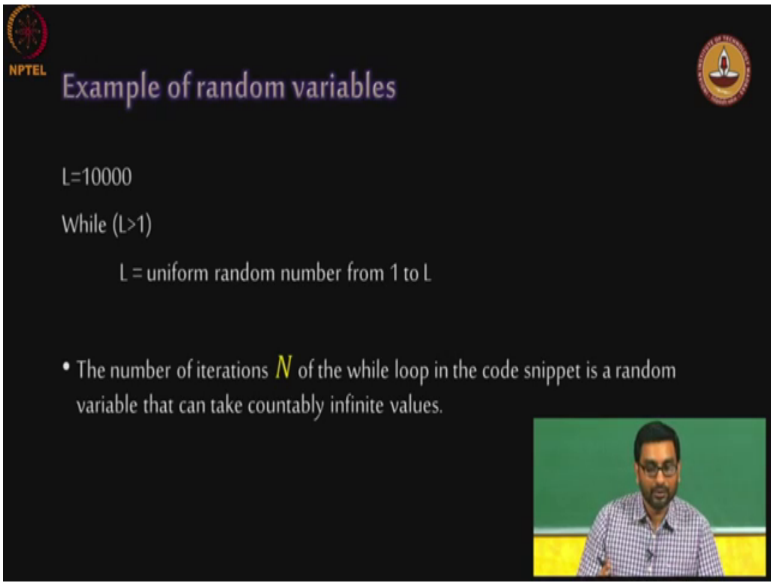
(Refer Slide Time: 07:18)



So, let us look come back to our my favourite things example you probably not your favourite, but um. So, what do we do here we have two outcomes either two in let us look at it just for the fun of it from my perspective, two is what I gain with if the coin comes up heads and I have to give you minus 1. So, it is a loss from me if the heads the coin outcome is said tails right.

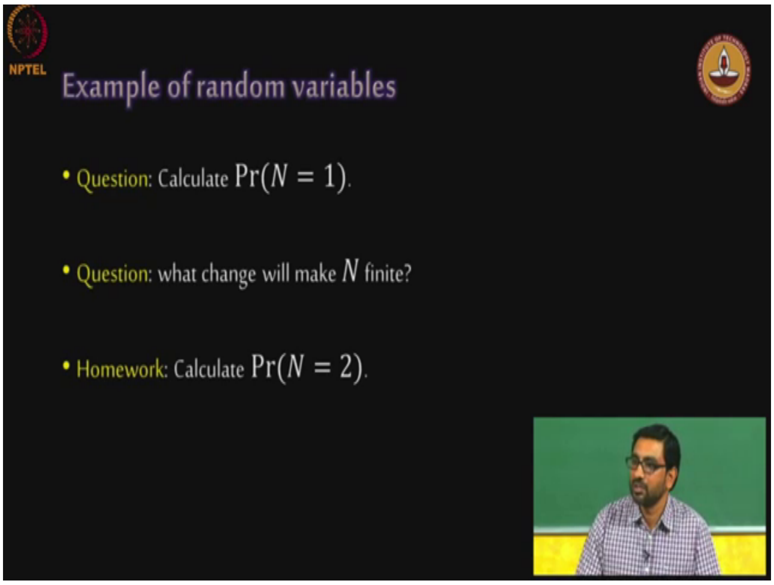
So, now one can think of you know, what do I what am I likely to get you know. So, it is a it is a notion that we need to develop and it is a, but in this case I am equally likely to both gain 2 rupees and lose 2 rupees it is both of these events if you will have equal likelihood ok.

(Refer Slide Time: 08:16)



So, coming back. So, this loop let us look at this loop the number of iterations of the while loop in the code snippet, is again a random variable can take as we saw it can take countably infinite number of values.

(Refer Slide Time: 08:30)



ah What is the probability remember that these are random events N equal to 1, N being let us say here denoting the number of iterations what is the probability that N equal to 1?

Student: (Refer Time: 08:46).

1 over 10,000 yeah in with probability 1 over 10,000 you would choose the a value of L to be 1, in which case it exits the loop and what is this slight change that I can do in order to make this of have finite outcome? Let us go back to this how do I make this sample space finite here?

Student: (Refer Time: 09:09).

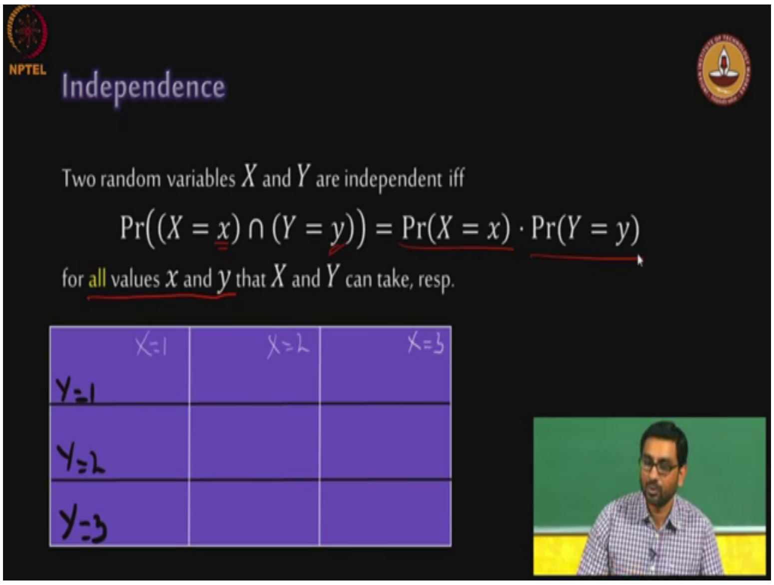
That is one way. So, suppose I just want to play with this line.

Student: (Refer Time: 09:23).

Yes. So, if you choose a uniform random number from 1 to L minus one you are certainly going to be decreasing every iteration. So, the number of iterations is not going to take exceed 10,000. So, one I am going to leave this as a homework what is the probability that N equals 2.

So, in exactly two iterations this loop will exit ok. So, that is a little homework for you.

(Refer Slide Time: 09:49)



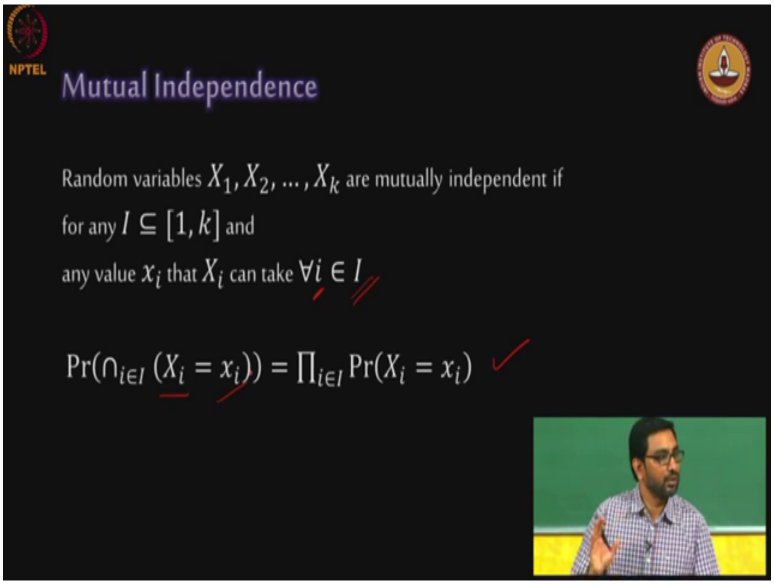
Now that we know what a random variable is, let us think of some is standard things. So, for example, we already know the notion of independence, and if events can be independent random variables are closely tied to events. So, the notion of independence should be extendable here as well.

So, but we have to be a little bit careful about what we mean by two random variables X and Y being independent, they will be independent if and only if now the important thing is this has to hold for all values of x and y, probability that the random variable X takes this little x value, the random variable Y takes this little y value is equal to the probability that random variable X takes the value little x times the probability that random variable Y takes a little value y. This is a very natural definition it is just that you have to realize that this has to hold for all values of x and y ok.

So, think of an example this way. So, you look at the sample space let us say you divided up the sample space. So, now, X equal to1 is this entire column X equal to 2 is this entire column, X equal to 3 is entire column , but Y equal to 1, Y equal to 3 and Y equal to 2 and y equal 3 are all rows let us say.

So, now if I if you take any pair this condition for independence should hold if you think about it ok. So, this is one example where it shows how you can get independence between two random variables.

(Refer Slide Time: 11:37)



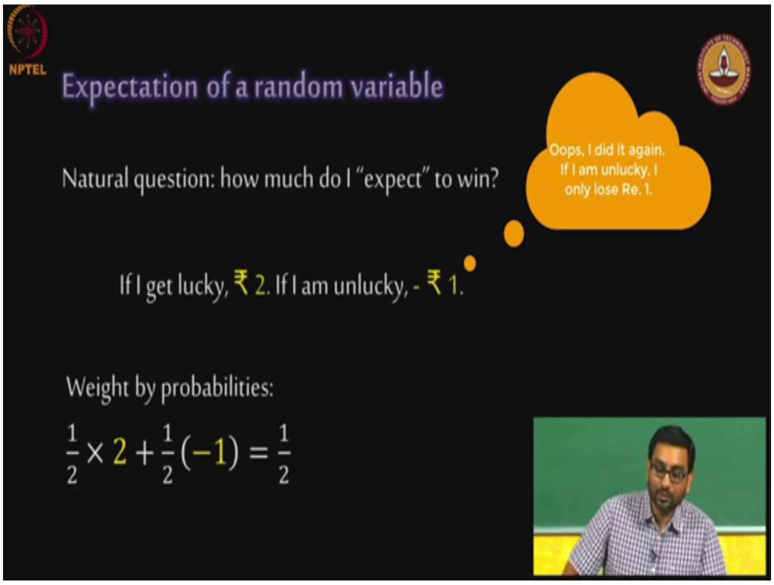
And the notion of mutual independence also extends and we saw the notion of mutual independence as part of our tutorial last class right. So, the notion of independence also extends um.

So, now if you have k random variables, x 1 through x k and for any subset of those k values, there and for any assigned assignable values X i to random variable x i where the iis in this the chosen index i is in the chosen subset of indices this condition has to hold the independence condition has to hold what is that? If you take the intersection of all those events X i equal to x i variable x random variable X i is equal to the value x i.

If you take the intersection of all of those and take the probability of that that should be equal to the product of the individual probabilities and this is again a direct extension of what we already.

Now, comes an important topic. So, what is this we have already alluded to this a random variable is basically measuring some quantity and so, for example.

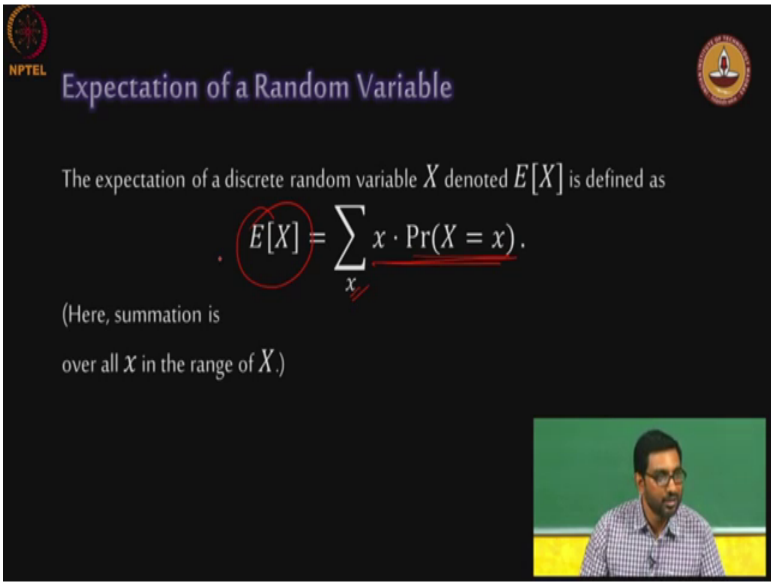
(Refer Slide Time: 12:55)



So, in the in the in the game that we played, a natural question I can ask is how much do I expect to win um. So, it is if I get lucky, I will get 2 rupees if I am unlucky I will lose 2 rupees ok.

And. So, what I can do is I can wait these values these are my either my gain or my loss I can weight them by their individual probabilities, and I get the value half. So, that kind of indicates how likely I mean how much I am likely to get. It makes sense right I am if I am lucky I get to otherwise I lose minus 1 somewhere in the middle is the half value.

(Refer Slide Time: 13:37)

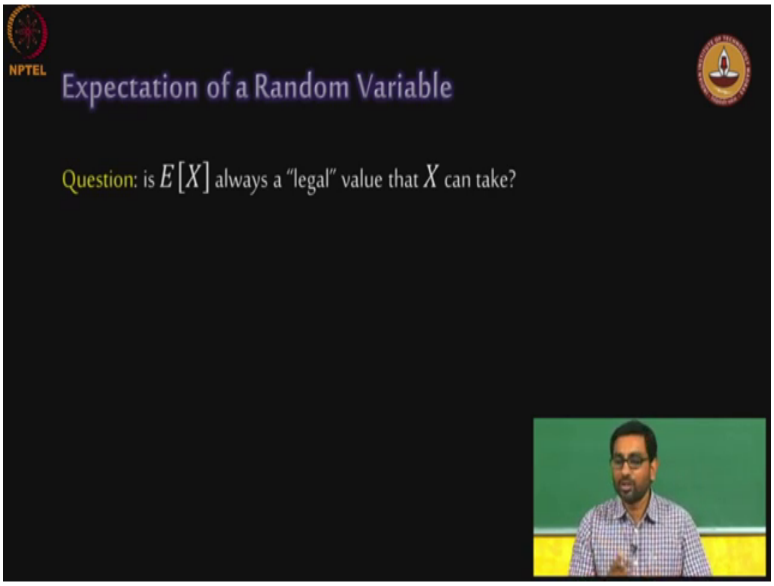


So, in general in you can generalize this notion for any random variable um. So, consider all the possible values that the random variable can take and that is denoted by little x over here, and if I just sum them all up, but weighted by their probability so; obviously, if somewhere it can take a value ah. So, going back to our the loop example, it can yes I remember I said that it can repeatedly keep choosing 10,000 right, but the probability that it will repeatedly keep choosing 10,000 becomes very very small.

So, you have when you are computing the expected value you need to wait the event according order or the value of the power of the possible value of the random variable, with the probability with which such a value will occur right. So, when you do that waiting and then you sum up over all possible values that the random variable can take, what you get is the expectation of that random variable x ok.

So, this is a very natural quantity to study when you think of a random variable. In fact, you usually be the first quantity you will want to understand about the random variable.

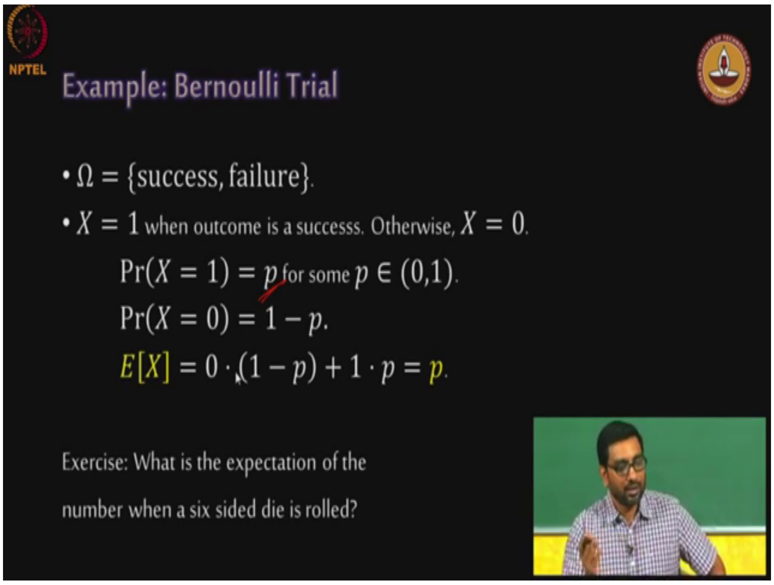
(Refer Slide Time: 14:54)



One quick question will E of X always be a legal value that X can take no. In fact, most of the times it would not.

So, for example, in the game that we played, the expectation we calculated was half and that is not a legal output at all it is just a measure of how much I can I am likely to get ok.

(Refer Slide Time: 15:14)



So, let us let us look at a very simple example. So, here this will be an example for what we are asking both, but also a very important experiment, it is just an experiment where the outcomes are just do typically called success and failure is called a Bernoulli trial ok.

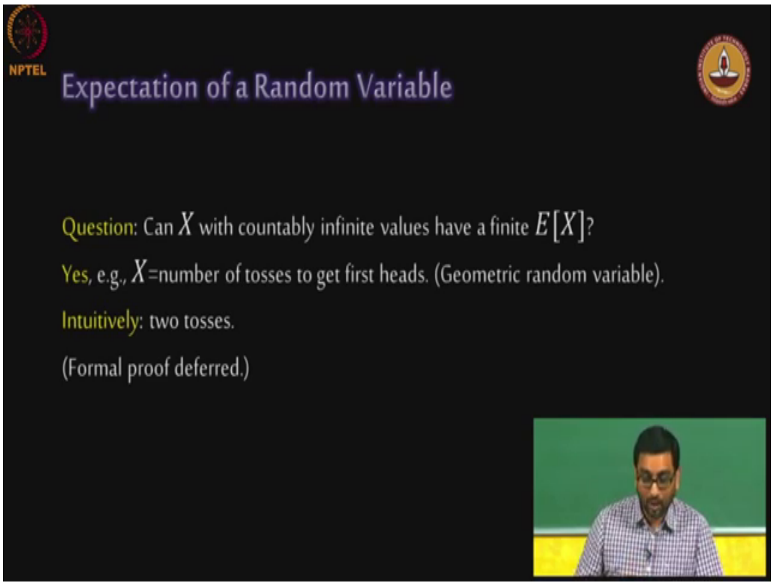
Any two outcome experiment is a Bernoulli trial you can assign one the notion of success the other the notion of failure ok. And a typical thing to do is assign a random variable as well and define a random variable as well, that takes the value 1 corresponding to success and 0 otherwise ok.

So, now you can again ask what is the expectation of x ok. So, a typical thing would be to generalize and say look it is not you know it is often people think of Bernoulli trials as fair coin tosses, but they are more generally the success can happen with some probability p, and failure should be a will occur with probability 1 minus p ok.

So, what is the expectation? it is very simple right you take the individual values. So, zero comes with the probability weighted a weight of 1 minus p and 1 occurs with probability weight of p. So, when you sum them up you get p and this is your expectation of the random variable x ok. Another question you can easily ask is you know what would be. So, now, for a for a die the outcome is very clearly randomly will the number either 1 2 3 up to 6 right.

So, if you work it out you are going to get the expectation to be 3.5 not very difficult to work that out.

(Refer Slide Time: 17:01)



Can x with countably infinite values have a finite expectation. So, we already looked at one countably infinite sample space and random variable cannot have a finite expectation any thoughts on this, what do you guys think?

Student: (Refer Time: 17:19).

Yeah it can. So, one. So, let me introduce one more important random variable. So, here X is the number of tosses just think about a fair point for now number of tosses to get the first heads ok. And if you think about the sample space you can actually get tail states since say for a as an arbitrary long period of time followed by a heads.

So, if you look at the sample space it is countably infinite this is called the geometric random variable the number of times you have to toss before you get the first heads, but intuitively when you think about it in a fair point, if both heads and tails are equal equally likely, within about two rounds you are likely to see you yes it is a very intuitive statement I am making the formal proof will be we will see you look at it in a short while ok. So, this is an example where the when you keep you have an infinite sized sample space countably infinite ah, but nevertheless the expectation is a small.

(Refer Slide Time: 18:28)

So, we come to the end of this segment, we have introduced the notion of random variables and the notion of expectation, independence of random variables and so on. We will study a little bit more about expectation in the next segment.