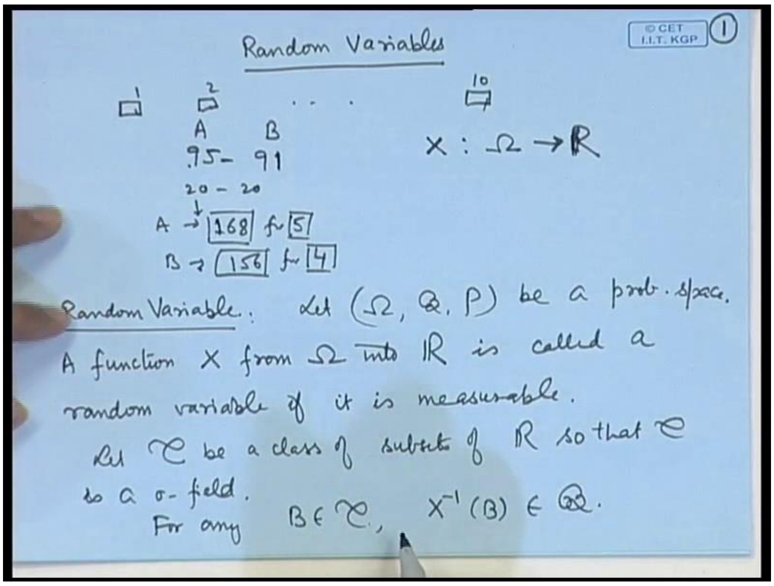
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 13
Random Variables

So far we were discussing the loss of probability. In the loss of the probability we have a random experiment, as a consequence of that we have a sample space, we consider a class of subsets of the sample space which we call our event is space are the events. And then we define a probability function on that.

Now we consider various type of problems for example, calculating the probability of occurrence of a certain number in throwing of a die, probability of occurrence of certain card in a drying, probability of various kind of events. However, in most of the practical situations we may not be interested in the full physical description of the sample space or the events, rather we may be interested in certain numerical characteristic of the event consider. Suppose I have 10 instruments and they are operating for a certain amount of time. Now after working for a certain amount of time we may like to know that how many of them are actually working in a proper way and how many of them are not working properly. Now if there are 10 instruments, it may happen that 7 of them or working properly and 3 of them are not working properly.

(Refer Slide Time: 02:06)



At this stage we may not be interested in knowing the position. Suppose we are saying 1 instrument, 2 instruments and so on 10th instrument; 1, 2 up to 10. We are not very particular whether instrument number 1 as failed or 2 has failed or 10 has failed; that means, which three of them a failed may not be of interest rather we are interested in the total number of the instruments which have failed and total number which are working.

You look at say a game of basket ball. Now in a game of basket ball after the completion of the play the score line will read something like say 95 91; team A may score 95 and team B may score 95. The outcome of the match is that even e has beaten the team. And to a particular observer this particular phenomena are this basket ball match has the final outcome that the team A won the game, are the finally scores are 95 and 91. Whereas, for audience they may be interested they may be looking at which player is scored how many baskets? What was the percentage of the position of the team A over the team B over the ball etcetera?

So, similarly if you consider a game of cricket, then suppose we are looking at a game of 20-20. Then team A is score say 168 runs for say loss of 5 wickets in 20 overs, team B may score say 156 runs for the loss of say 4 wickets at the end of 20 overs. The outcome of the match is that the team A won over team B, or if you look at the totally scores then these are the totally scores. If we look at the total number of wickets then these are the total number of wickets lost. We may not be interested in seeing that out of 20 overs there are 120 balls; at which ball how many runs we are a scored, how many balls we are dot balls. Though these are important data but they may be useless for a certain person who is interested only in knowing what is the total number of score, total wickets taking are a particular player scored how many runs or how many wickets he has taken.

Therefore, for example if we are looking at say weather. Now the weather is a complex phenomena, so if we look at say monsoon season then during a full monsoon season the amount of rain varies day to day from region to region from time to time. That means, in a particular day also a particular portion of the day may be dry, particular portion may be cloudy, and particular portion may be rainy.

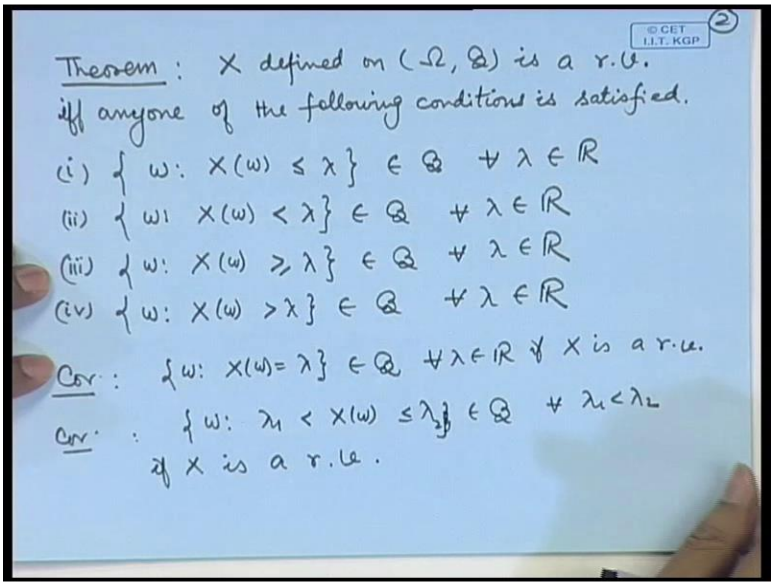
Whereas, for the weather scientist individual phenomena may not be of interest, he may like to know that how much total rains are there in different zones of the country during the entire monsoon season. If we are looking at say crop of a certain variety and we look at the total yield say wheat. Then we will be interested in knowing that how much total crop of wheat was produced in India during 2008 to 2009. We may not be interested in knowing which plot of, which former gave how much yield or how many facilities where their because of which something happen. The total amount may be of interest to us.

In effect what I am trying to convey is that in order to effectively study a certain random phenomena or a certain random experiment it is required to quantify the phenomena. That means in for every outcome we associate a number, a real number and then we can a study the probability distribution of that, that brings us to the concept of Random Variables. So, what is the random variable? Then roughly speaking random variable if we give a notation X then it is a function from omega into T. That means, for every outcome we associate a real number. For example, runs a scored by player, the basket said by a player in a basket ball game, the yield of a certain crop, the amount of rain fall in a certain region etcetera. So, all of this will correspond to certain event in the sample space and we are associating certain real number there. So, in particular any random variable is a real valued realization of a sample space.

So, it is a function from the omega into R. However to be more precise, because we have defined the probability function over a sigma field of subsets of omega, it is important that when we want to talk about the probabilities of associated with those numbers then the inverse images of those subsets must be in the set b. So, this is called the condition of measurability. And the formal definition will then become, the formal definition of random variable is; so let omega B P be a probabilities space. A function X from omega into R, R is a real line is called a random variable if it is measurable.

So, it is plane there condition of measurability; let C be a class of subsets of R so that C is a sigma field. Then for any B belonging to C X inverse B must belong to a script B. This is the condition for measurability.

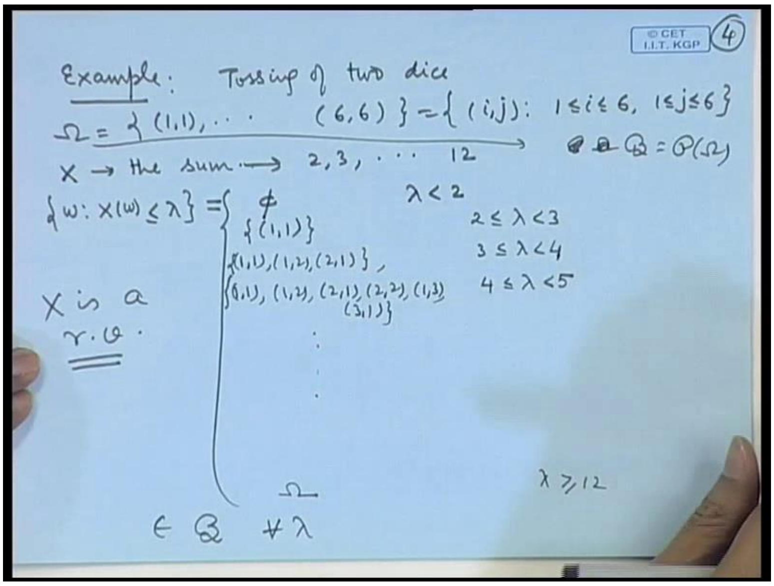
(Refer Slide Time: 09:20)



And equivalent condition here is that we can state in the form of a theorem. So, X defined on omega B is a random variable if and only if any one of the following conditions is satisfied. The set of omega such that X omega is less than or equal to lambda belongs to B for all lambda. The set of all those sample points such that X omega is strictly less than lambda belongs to B. The set of all those points such that X omega is greater than or equal to lambda belongs to B for all lambda belonging to R. The set of all those omega such that X omega is strictly greater than lambda belongs to B. So, if any of these conditions is satisfied then X is a measurable function, and therefore it is a random variable.

As a corollary to this the set of all those omegas such that X omega is equal to lambda then this is also for all lambda belonging to R if X is a random variable. Similarly, if I consider the set of all those omegas such that if X is a random variable. So, these are some natural consequences. Let me just give example here to show that how we will you check the condition of measurability.

(Refer Slide Time: 11:52)



Let us consider say tossing of a coin. Let us consider the random experiment of tossing of a coin. So, the sample is space is here H, T; we define the random variable X has the number of heads. So, naturally here X H is equal to one and X T is equal to 0. So, if you look at the set omega such that X omega is less than or equal to lambda say. Now if lambda is negative then there is no point for which X omega is satisfying that condition less than or equal to lambda. That means this is an empty set.

If I take lambda is equal to 0 then X T is equal to 0; that means there is a point here which is satisfying the condition that X omega is less than or equal to 0. You can observe that this condition remains true for lambda up to 1. That means, it is strictly less than 1, because a next value that random variable takes is actually equal to 1. So, when it eventually takes 1 then there are two points that is H and T which satisfy the condition that X omega is less than or equal to 1 or any point above that. So, this becomes full omega for lambda greater than 1.

So, if my sigma field here is consisting of say phi, H, T and omega then you can see that this set is always belonging to B. Hence, this X is a random variable. Let us take another example; say tossing of say two dice. Here sample is space is say 1, 1 and so on up to 6, 6. All the combinations of numbers i, j for 1 less than or equal to i less than or equal to 6, 1 less than or equal to j less than or equal to 6 all the combination of natural numbers between this. So, if I define the random variable X as say the sum observed and we want to write down the set omega say X omega less than or equal to lambda. Now if you look at this sum, sum will vary from 1 to 12. So, if I have lambda to be less than 2 then this set is going to be phi. If I take lambda to be between 2 and say 3 then I have the point 1 1 here.

If I have lambda between 3 and 4, greater than or equal to 4 but a strictly less than 4 then we will have 1 1, 1 2, 2 1; if I have lambda greater than or equal to 5 but a strictly less than 5 then this will become equal to 1 1, 1 2, 2 1, 2 2, 1 3, 3 1. Like that if I say lambda is greater than or equal to 12 then this will be full omega. So, if I consider my sigma field here as power set of omega then this will belong to B for all lambda. And therefore, X is a random variable.

The purpose of considering this measurability condition is that when I want to talk about the probabilities of these events in terms of random variable then they should be well defined in the original sample space.