Communication Engineering
Prof. Surendra Prasad
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture - 27
Review of Probability Theory and Random Processes

We now come to the next part of our course. Next part of our course deals with noise. We will look at the performance of the various communication systems that we have seen or various modulation schemes that we have studied namely various variations of amplitude modulation and the frequency and phase modulation. And try to understand how do they behave in the presence of noise. As you know noise is one of the major limiting factors in the performance of communication systems.

Now, what is this noise that, we talk about, how do we characterize it, we already seen some of that in our introductory marks about the course and the very beginning of the course. We said that noise is generated by every physical component that you might use in the transmitter and the receiver, every electronic component. So, every electronic component therefore, is a source of noise. The transmission medium through which the information passes is a source of noise.

In a very rough sense noise is some random waveform, now when you are dealing with random waveforms. First of all we need to understand, what are so called random waveforms, how do we mathematically characterize them. Unless we have a mathematical modeling characterization for these waveforms, we cannot possibly do an analytical or theoretical study of the performance of communication systems in the presence of noise.

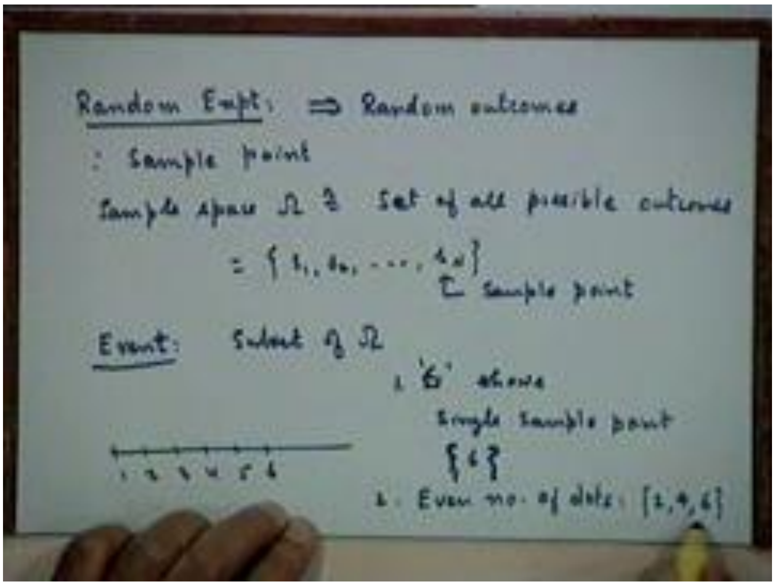
So, that is the next part of our course that, we are going to deal with for at least a few weeks. Now, why we want to characterize noise we need basically I need to deal with random phenomenon and the mathematical discipline with deal with random phenomenon is the discipline of probability theory, random variables and random processes.

In view of finding the class has a varied background; I will go through a very quick brief brush up with probability theory like random variables and then proceed on to random processes as it is relevant for our discussion. So, it will take a while we go though that and we will start the process immediately. So, to start with let me define some basic things, some problem here I think I have to probably might.

Student: ((Refer Time: 03:40))

Because of the transparency, I do not think we are getting a good image; I will do something about it next time, now I will write it does not matter.

(Refer Slide Time: 03:58)



We will start with the concept of the random experiment which is our starting point. A random experiment is any experiment or any phenomenon that we look at whose outcomes are unpredictable and therefore, are random. So, basically a random pro experiment is characterized by random outputs rand by random I mean unpredictable.

An every outcome that you can possibly has a when we perform the experiment in a formal sense; we define it as a sample point. So, sample point is any possible outcome of a random experiment, I am just giving a quick definition of various kinds and then going ahead.

Now, the collection of all possible sample points, that the random experiment can produce that is a collection of all possible outcomes that it can produce we call it a sample space. So, sample space omega is we can say that set of all possible outcomes, so if the outcomes are denoted by s 1 s 2 s sub N. Let us say it has a finite number of outcomes, then this set is aggregation of all possible outcomes constitutes the sample space and each of these individuals points is called a sample point.

I will not give too many examples because most of you familiar with these things, for example, when you toss a coin the two possible outcomes that you can have are heads and tails. So, the number N there is 2 and the sample points consist of H and T the head or occurrence of a head or occurrence of a tail.

Next we defined the concept of an event; an event is either a single sample point or a set of sample points. Basically, in a more general sense a subset of the samples space is an event common, so a subset of the sample space and there can be very various subsets of a sample space is called events.

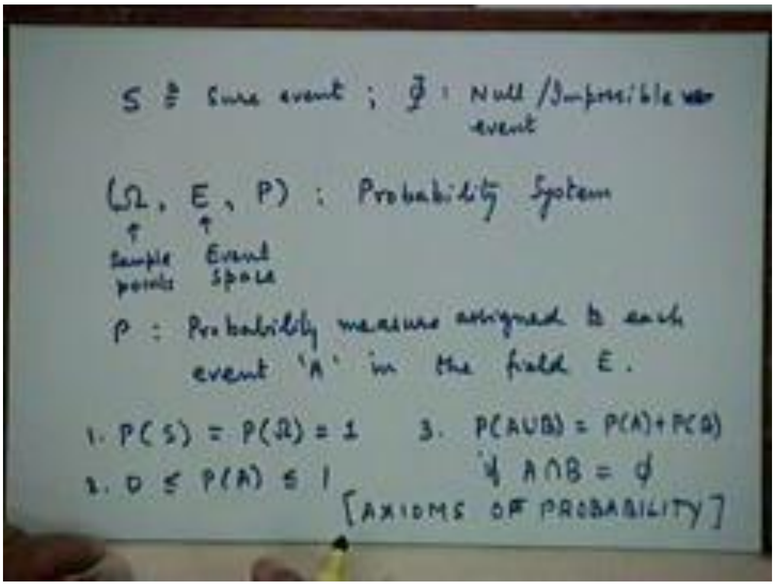
For example, again tossing a coin where would the event could be head or tail occurs, it is this is a now this is a complete sample space either head or tail will occur. But, we define it as an event neither head nor tail occurs when you toss a coin it is an impossible event one of the two will occur.

So, when we talk of various logical combinations of the events of the basic outcomes all this sample points we cont we construct the events. So, if we have N sample points in a sample space the number of possible events that you can talk about is 2 to the power n the power set of s, the power set of from event.

So, there is let me take an example here, suppose we talk about throwing a dice there are 6 possible outcomes. So, this is a one dimensional sample space and various events that I can talk about are let us say this the numeral 6 shows up when you throw a dice, what is the event here, the event here is a single sample point this is a single sample point event.

And typically, an event will denote by this notation this denotes a subset, a subset consisting of a single sample point. We could talk of the event that, the outcome is an even number that you see an even number of dots, so it could be either 2 or 4 or 6, so the event could be that you see an even number of dots when you throw the die. So, the subset that we are talking now talking about now are is this subset, that you see either 2 or 4 or 6. This is the logical combination of the sample points that we have.

(Refer Slide Time: 09:14)



Now, we could have a sure event, a sure event will always occur. For example as I said in the case of throwing of a die, anyone of the 6 numbers shows up it is a sure event, so it is, how it is basically the complete sample space, so sure events are one which will always occur that is why we call it a sure event. Similarly, we define a null event by this notation null or impossible event, it will never occur, both the sure event and the null event are subsets of the sample space.

By having defined a probability having defined a sample space omega and what else we have defined the events are all basically a combination of all possible subsets of the sample space omega. Typically, we call it a field denoted either by E or F or class, so there is one sample space and there is a field, the sample space consisting of consists of sample points and this field is actually an event space. And this field is the set of all possible subsets of omega the power set of omega in some sense.

And having defined these two to complete our formal frame work in which we can discuss probability. We define a probability system by three things, omega the field of events or the sample of the space of events and a probability measure P. So, this triplet omega the corresponding field of events and the probability measure P is called the probability system, in our in which will provide as a frame work for our studies.

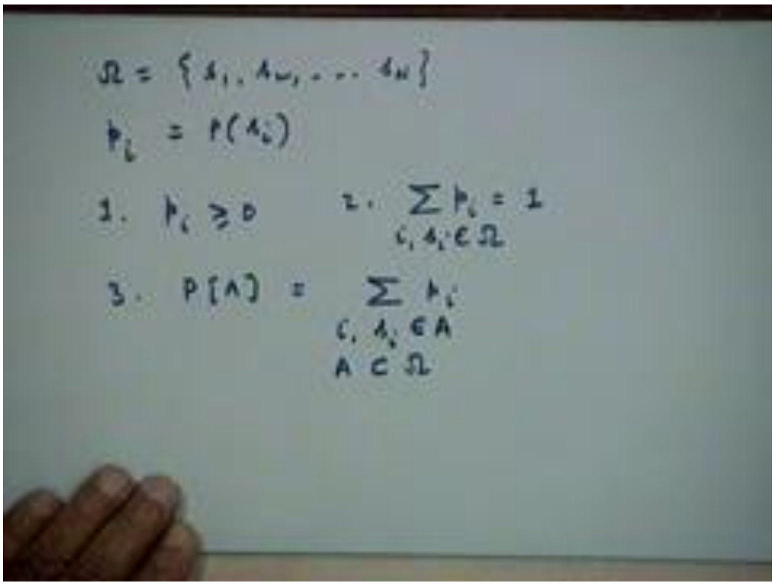
Sometimes this is also called space probability system or probability space. So, it consists of these tripling. Basically this last thing provides a mapping from omega or from E to some real number real positive numbers which will essentially assign a probability measure to the occurrence of each event.

So, p is a probability measure assigned to each event A in the field E and this probability measure that you use to define this probability space should satisfy three axioms, these are very the very important axioms these are called the axioms of probability. The probability measure three that we assigned the probabilities that we assigned to various events must satisfy some properties.

These properties are first of all the probability of s; the sure event is basically the same as the probability of the sample space omega itself, because the sure event is typically the sample space itself. This is equal to 1, that is a first axiom, the second is this has to be for any event A, this probability must be a number which lies between 0 and 1 and three if you consider the union of two events A and B which are disjoint events what are disjoint events A intersection B is null set.

So, if A intersection B is null set, then probability of A union B must be equal to probability of A plus probability of B, if A intersection B is equal to the null set. These three properties which any probability measure must satisfy constitute the basis of probability theory and these are called axioms of probability, it is three basic axioms of probability. Any questions so far.

(Refer Slide Time: 14:53)

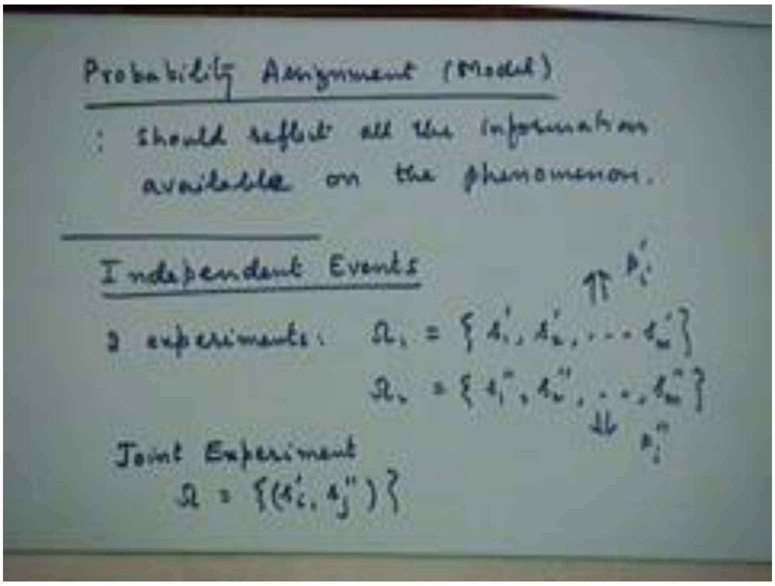


An implication of this is that for example, if you take your sample space omega whose numbers were s 1, s 2, and s N etcetera. Let us we lets denote by p sub I, the probability of s sub I the probability measure of the I th event. Then, the implication of these axioms would be something like this one implication would be where these will be all positive numbers, the probability of each of these sample points or sample is a is greater than or equal to 0.

The second implication would be that, if I add up all these probabilities that is p 1 p 2 p N, this would be equal to 1, where I overall I such that s i belongs to the sample space. And the third implication would be if I consider an event A which is basically consisting of a finite subset of these sample points. Then, the probability of this event will be the some of these some of the probabilities of those sample points of which it is composed.

So, it is simple equal to sigma p I, where I is chosen such that s sub I belongs to the event A and of course A has to be a subset of omega. I have just put in mathematical terms everything that we are discussed, there is nothing really new here, so in a way once this pi's are selected probabilities of all events get determined. In as much as you can identify in every event A the corresponding sample points which constitute the event here. So, once pi's are chosen and the probabilities of all events will get fixed.

(Refer Slide Time: 17:07)



In many cases this probability assignment that you do this probability measure that you choose should follow an appropriate model which is consistent with the physics of the situation. So, the probability assignment usually has some model, but no matter how you choose it must satisfies this axioms of probability.

So, basically it should be related to the physics of the situation and therefore, in a way should reflect all the information that you have, on the phenomenon that you are modeling. Of course you cannot go by the precise physics, because if you could workout the precise physics you did not have to take the course to a random model.

The reason, why we go for a random model is, because it is difficult for us to work on the precise physics that of course, one misunderstands. But, whatever information is available this assignment should be consistent with that information. For example as an example suppose each of this pi’s that I talked about in the previous slide, suppose all these are equal.

Every sample every basic outcome has a same probability; all the outcomes have equal probability. Then suppose, in a particular event a particular event is composed of a subset m of these outcomes, then what can we say about the probability of these events; the probability of that event which will be m upon n.

If because all of them are equally likely, so this is one possible model assumed that all events are equally likely all basic outcomes are equally likely? And then, count the number of basic outcomes which constitute a particular event and this counting process itself determines the probability of that event. So, I will not go further on this. Next we come to the concept of independent events; I am going too fast, because I am assuming that you actually know all these things. But, you will require various brief and quickly brush up.

Suppose, we have two experiments, we perform two experiments and the sample spaces associated with them are omega 1 and omega 2 and let us say the outcomes of if omega 1 outcome of experiment one are s 1 prime to s N prime and omega 2 as s 1 double prime s 2 double prime etcetera. Another say the probabilities associate with these outcomes are p sub I primes and the probabilities associated with these outcomes are p sub I double prime. Then, we define a joint experiment.

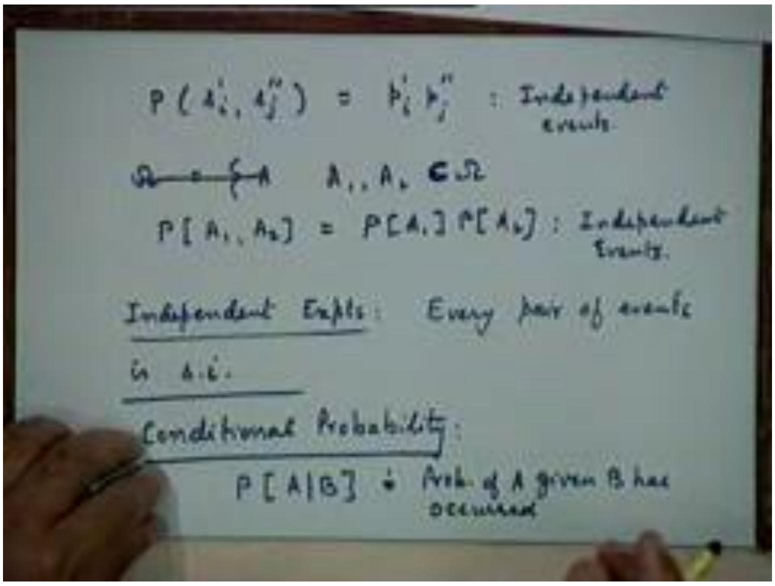
Student: ((Refer Time: 20:38))

These are there could be sometimes the same samples space as a special case, there could be, but we are talking of an intelligential general situation. So, we talk about a joint experiment omega, whose outcomes are the joint commutation the total aggregation of the two sets of outcomes. So, these outcome consisting of such pairs s i prime s j double prime that is experiment one produces s i prime, experiment two produces s j double prime.

So, the basic outcomes now are this pairs, all possible pairs and now the description that we have discussed for the case of a single experiment will carry over to this joint experiment. The only thing is instead of working with simple probabilities we will work by what are called joint probabilities, the joint probability of experiment one producing s i prime and experiment two producing s j prime.

It may so happen that, the occurrence of the outcome from one experiment tells us gives us no information about so where about the outcome that will occur from the second experiment. And that is appearance of anyone of these events has no bearing on the appearance of the events from here, if that happens we say that experiments omega 1 and the two experiments are independent.

(Refer Slide Time: 22:28)



So basically, first of all we have to define independent events you say that two events s i prime and s j double prime are independent. If this joint probability the probability of this joint event is essentially p i prime and p j double prime is the product of these two probabilities, so this is the definition this is independent events. Actually, I have defined in terms of the sample basic outcomes, but these could be replaced with events.

In fact, we can talk about independent events even in the same experiment, we could have in a same experiment omega two events A 1 let us say A 1 and A 2 both belonging to omega both subsets of the same sample space. So, let me write it as both events define in terms of omega and you could have a situation where the joint probability of A 1 and A 2 that is these two events occurring together, this is the product of the corresponding individual probabilities.

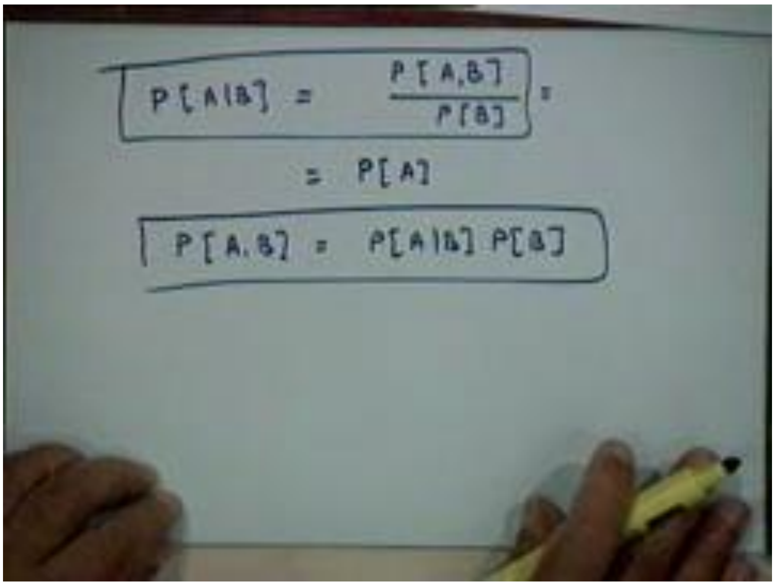
So, these are called independent events, these are not the same thing as a exclusive events there is something called exclusive events for example, if A 1 occurs A 2 will not occur. Incidentally, you can argue out that such exclusively events mutually exclusive events will not be independent; they will have some kind of dependency. But independent events are those such that they are joint probabilities the product of the individual probabilities.

If this happens for every possible pair of events across two experiments, we say that the two experiments are independent statically independent. So, in that case these experiments themselves become independent, so we can also have independent experiments. Basically, implies at every pair of events is statically independent satisfies this property, when we say that the corresponding experiments are statically independent.

So, that is the very quick review of one last concept which I think I must mentioned before I go into random variables, the concept of conditional probability. We can also define probability of a given probability of the event A given that some other event B has occurred. You may have knowledge that a certain event has occurred and now you want to ask yourself event that has occurred where is the probability of another event A; obviously, for this to be meaningful typically A and B should have some relationship with each other.

If we do not have the relationship this well we can still talk about it, but its value will be some fixed value. So, this is called the typical notation probability of A given that B has occurred given that given the B has occurred, so probability of A the event A given that B has occurred.

(Refer Slide Time: 26:52)



It can be shown that one can write this probability in terms of the joint probability of the events A and B upon the probability of event B. If I can argue it out from basic fundamental principles again since it is I will not go into that kind of discussion.

But, it is very easy to say, what is the implication if A and B are independent, what do you say. This will be nothing but probability of A itself, because if A and B are independent, this will be equal to p of A into p of B, P of B will get cancel and we left with p of A. Basically what we are saying is, that if B gives no information about A, then basically this is nothing but probability of A itself. So in general, what do you expect p of A B or A given B to be less than P of A or more than P of A?

Student: ((Refer Time: 27:55))

Think about it

Student: ((Refer Time: 28:01))

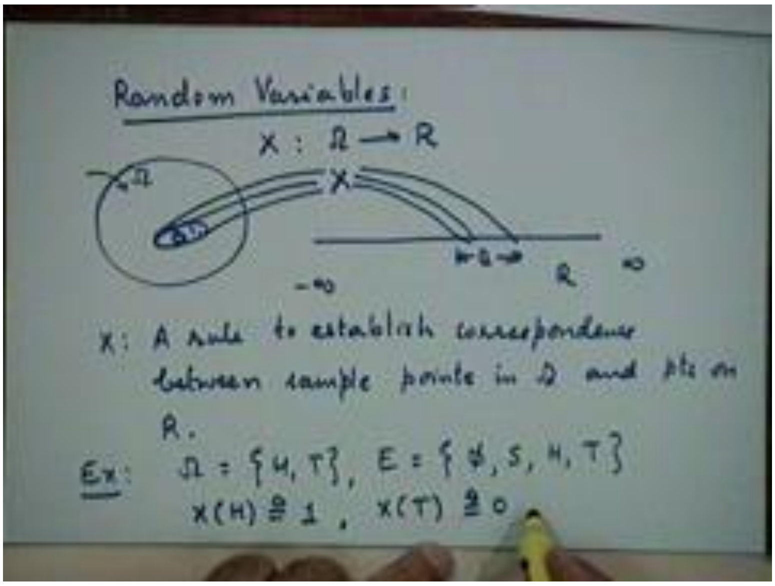
Less than.

Student: ((Refer Time: 28:03))

This is in some sense a lower bound, if there is no basically what we are hopping is that if there is a relationship between A and B we get some more information about the occurrence of A and improve the A prime knowledge about A about the occurrence of A. So, hopefully it should it should improve the probability of A.

So, more about incidentally this relation is also called base formula or base relationship between the conditional probability and the joint probability. Another way of defining the joint probability therefore, is the conditional probability of A given B into probability of B, so these are some very basic formulas that you should know.

(Refer Slide Time: 29:13)



We next turn to random variables, so far when we discussed the probability space or the probability system that I defined. Our basic experiment could produce any kind of events any kind of outcomes which could be descriptive or which could be numerical.

For example your sample space which consists of descriptions likes heads and tails or some other similar non numerical outcomes. And obviously, when we are working at the mathematical model of things is much more convenient to work with things which are numerical in nature rather than non numerical in nature.

And that is the primary motivation, for bringing up the concept of a random variable, a concept if I define this as my sample space, if I denote this circle lets the circle denotes the sample space omega; that means, all the outcomes are points in omega and the events in events which are associated with the omega would be subsets of this. So, suppose this is the subset of this, so this will become an event in that probability space in that defined corresponding for these events this sample space is that, so this is some event let me define it denoted by B.

Now, basically a random variable is a mapping if I say a random variable X, it is a name that I give to a mapping which maps events in this sample space to corresponding subsets of the real line. So, it is a mapping from the original sample space omega to the real line, so X is a mapping from omega to the real line R.

So, all the points which constitute this event B, they will typically get mapped to through this random variable X to a subset of the real line, so this event B will now be designated something like this on the real line. So, when I say the event B has occurred in the original sample space, what we could equivalently say is that the random variable X has taken some values in this range on the real line.

So basically, in these sense events here, events in the original sample space get mapped to subsets of the real line. Instead of being subsets of omega, here omega becomes equal to omega gets mapped to the real line R the sample space omega gets map to get complete real line R from minus infinity to infinity and subsets of R become equivalent events that.

We can talk about the mapping is given a name X and X we call is a random variable, but remember, there is nothing random about this mapping, this mapping is unique a point to point mapping. If this event occurs, this will map to this interval, if another event occurs it will map to some very fixed corresponding interval. So, there is nothing random about the mapping itself.

But in as much as this events which occur, are random and therefore, the range of values that will see associated with this are random, they become random and therefore, we call this a random variable. But there is nothing absolutely nothing random about this mapping itself.

So, this is a rule random variable is a rule or a mapping to establish correspondence between samples points in omega and points on the real line. Actually therefore, in a strict sense it is not really a variable at all, it is a functional mapping, it is a function whose domain is this and whose range is this.

So, the random variable is actually a function it is really a misnomer to call it a random variable and also there is nothing random about it. The reason why we call it random is because the things that we are mapping are random maybe that is basically the point.

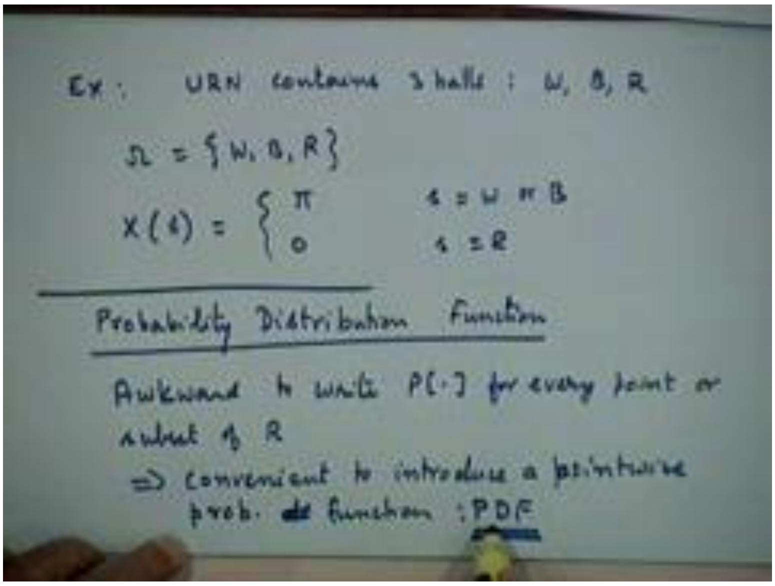
Let me illustrate this, so that there is no confusion about it take a very small example take a simplest possible example of tossing a coin this outcome are non numerical. So, your sample space omega here is H and T considerably, what is the corresponding event space here, what is E here, what is it will consist of the null set, the impossible event, the sure event the H and the T. The sure event is basically H union T and the null event is basically H intersection T intersection of H and T neither H nor T are occurs.

So, basically these are all four possible events you can talk about in this case and suppose, I now define a mapping like this, so X is a mapping which maps the outcome H to the value 1. So, this is the mapping that I am defining arbitrarily and it is X of X with the argument T is equal to 0, so its mapped H to 1 and T to 0.

Now therefore, you can see that is a function, it is a function whose domain is this and its range is the real line. So, it is X now can take only two possible values in this particular case in the 1 and 0, while I can still talk about events as subsets of the real line. For example, I can talk about the event that X the value of X is less than minus 5 although I will never have minus 5 here.

But, I can talk about his event, so it will be an example of a null event impossible event X less than or equal to minus 5, X less than or equal to 2 becomes a sure event. Because, the two values that I have assigned to H and T will be contain in this interval X less than or equal to 2 will be a sure event and so on and so forth.

(Refer Slide Time: 36:52)



Let me take one more example suppose an URN contains three balls a white one a blue one and a red one and let us say we choose a ball at random that is experiment. So, your sample space here is you will see either this or this or this or this are the three samples that you can have. So, let me define the random variable X to be like this lets say it will take the value pi, if s is equal to W or B it will take the value 0, if s is equal to R, this is another possible mapping.

So, once again X becomes the binary value random variable, so I can defined any kind of mapping and that becomes the random variable, for the different definitions of the mapping. I will have different random variables from same sample space and from the same experiment. I can define many different random variables on the same omega, all it takes is defined a different mapping will be there.

Can we define a mapping which we called the random variable? So, is the concept of the random variable, we now define the probability measure which helps it makes it very easy for us to work with the new domain that we have created the new range in which the random variable works namely the real line.

So, when you talk about probabilities now, instead of talking about probabilities the original sample space omega is much more meaning full to talk about probabilities in the new sample space which is a real line. And a probability measure which makes it very easy for us to specify the probabilities on the real line is a probability distribution function.

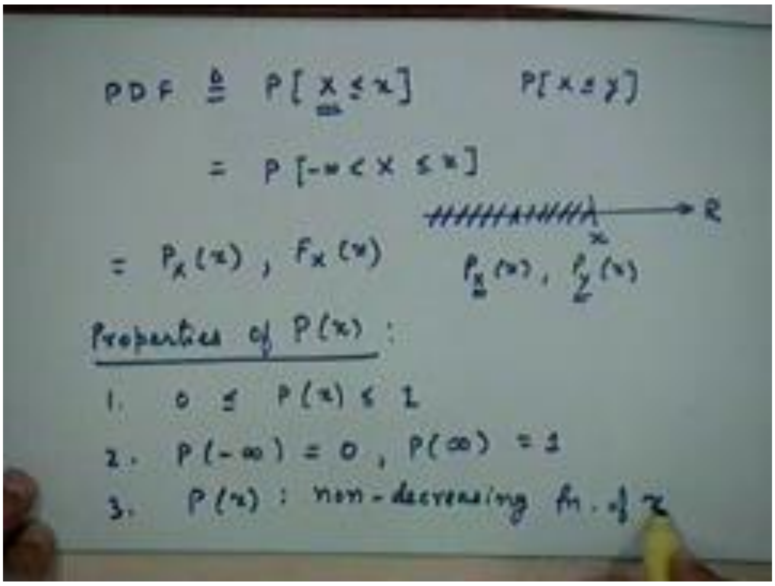
Before, I go into this please understand that not all in not in all situations you need to create such a mapping. In many situations this mapping is implicit in the experiment itself, you are measuring something numerically and the outcome is random. So, you already have some kind of a mapping which is implicit in the experimentation process which you do not have to explicitly define there. So, that something that you must understand that is par part of this formulation part of this frame working in which we are working.

So, let me comeback to the probability distribution function, since it is rather awkward to write the probability of every single event that is every possible point on the real line, that is what we need to do, otherwise is not it if we want to define the probability of various event from the real line now. We will have to define the probability of every possible interval; every possible point on the real line and that will be that will become, so awkward, so difficult to do.

For every point or subset of R, it is convenient to introduce a point wise function which is called the probability distribution function. So, it is convenient to introduce a point wise probability function which we call the probability distribution function, I will denote it by P D F which stands for probability distribution function.

It is a point wise probability function in the sense that it is defined by every point in the real line in a particular way. But, it is defined in such a manner that given this probability function, I can write the probability of any point is of probability of the random variable X, taking any particular value. Once I note this function I can so this awkwardness goes it becomes a very simple frame work.

(Refer Slide Time: 41:53)



The formal definition is probability distribution function is defined as the probability of the random variable X, taking a value less than or equal to some numerical value small x. Please remember in this notation, this capital X is the name of the random variable and the small x that I am using is defining a numerical range or which you take this random variable this mapping is checking the values.

I could as well I have said X taking the value less than or equal to y it won't make any difference because this small x and this small y is just some numerical value that I have selected. So, this is equal to the probability that X lies in the interval from minus infinity to small x which is essentially same that this is your real line and this is a point X, we are taking about the probability that X lies in this interval.

This is many times denoted by P sub X of X some books you will find F sub X of X a typical notation for the probability distribution function. In this notation the subscript X indicates the name of the random variable and small x denotes the numerical value which about which we are talking which specifies your range from minus infinity to that value.

It is probably sounding very elementary to some of you, because some of you have done this course nevertheless that is quickly brushed through. So, what we this is, what is called the probability distribution function sometimes we just denote this X the subscript x when it is clear which random variables we talking about.

You may be talking about two different random variables X and Y, let us say in that case please remember this two distribution functions although I am using the same notation P. The fact that their subscripts are different will be two different distinct functions altogether they are not the same functions.

So, the functional differentiation now comes through the subscript, because this is the probability distribution function associated with X, this is the distributed function associated with Y both are being considered as function functions of some variable X. So, small x is the actually the variable. The variable which takes on the numerical values, capital X and capital Y are the names of this the random variable.

Now, the distribution function will have some properties which are derived from which are result of the basic actions of probability. So, there are some basic properties associated with the distribution function, here at the moment denoting dropping the subscript X capital X.

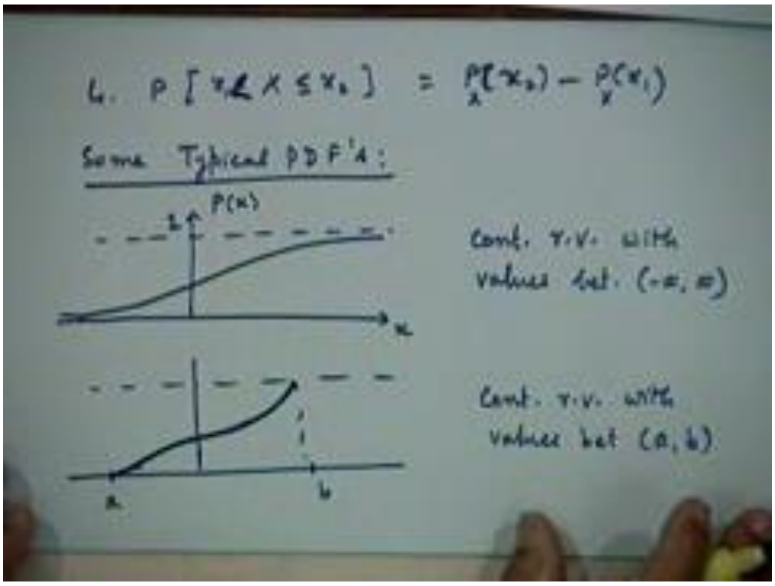
Because, I am not dealing with different variables, one is that for every value of small x, the value of this probability the value of this will be between 0 and 1. Basically, what does it denote, this denote some event on the real line and we already seen that every event will have a positive probability, according to the axioms of probability. This really follows from the axioms one of the axioms of the probability.

Then, what can we say about this, this corresponds to the impossible event that X is less than or equal to minus infinity, it is not possible according to the way we defined this. So, this is the impossible event its probability must be 0. Similarly, the probability associate the distribution function value associated with X equal to infinity could be equal to 1 corresponding to the sure event and third P of X is always a non-decreasing function which is quiet obvious from the definition.

Because, if I consider two values one less than this another value less than this then. Obviously, this event, this event will be the union of this event and this event and the probability of this therefore would be the probability of this plus the probability of this. It can never be less than the probability of this.

Even if this is a null event, this can happen this is a null event for in that case also the probability of this event will be at both at least this much. So, it can never decrease as small x increases, so it is a non-decreasing function of X.

(Refer Slide Time: 47:04)



The last the probability the X lies in some finite interval X 1 should be strictly less than and X 2 is equal to this is I think quiet; obviously, so I will not prove it you can prove it yourself. This is the value of the distribution function at X 2 minus the value of the distribution function at X 1

Student: ((Refer Time: 47:29))

No, it makes the difference, because by definition the point X 2 is included in.

Student: ((Refer Time: 47:40))

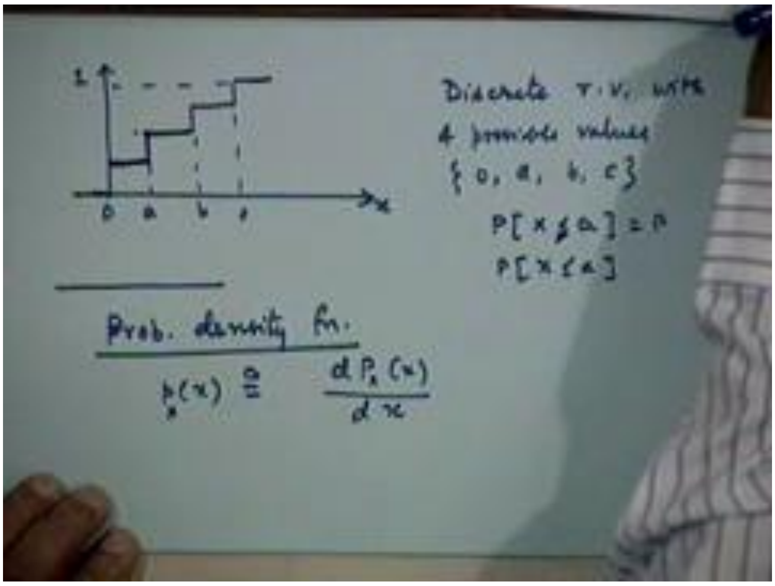
Here, we are doing a general treatment this is supposed to be valid for both continuous as well as discrete and no way else.

Let me consider, some typical distribution functions, if I take into account this properties when we just discussed based on this various kinds of distribution functions. But, one this is sure that every distribution function P of X, if this is I am I am plotting it against the variable small x.

It will start from 0 on the extreme left hand side, somewhere on the extreme left hand I mean somewhere it could be extreme left hand side or beyond that and cannot exceed one, it is the mapping which lie bet lies between 0 and 1, that is one thing. Second is it is a non-decreasing function, so typical curve could be like that it cannot be less than 0 or it cannot be more than 1, it has to be non-decreasing.

Another situation could be like this, here I am going from minus infinity and this actually denotes the continuous random variable, whose values lie between minus infinity and plus infinity. I could have a continuous random variable whose values lie in a finite interval let us say from a to b, so it is 0 and goes up to b. It can never decrease of course, I hope this is nor this is non-decreasing function once again I could also have so this is a continuous random variable with the values between only finite interval value a to b.

(Refer Slide Time: 50:27)



We can have a third kind of situation where P of X is something like this which is 0 up from up to here become some finite value here, some other finite value here and so on and so forth, but cannot exceed one; this is the case of a discrete random variable.

Because basically, what it means is that it can take these values which are denoted by a b c etcetera here. This is an example of a discrete random variable with four possible values namely the value 0, the value a, the value b and the value c, so does this make sense the last example.

Basically for example, if this is the only four values it can take and I ask the question what is the probability that X lies between X lies between X is less than or equal to a. Let us say, put strictly x less than or equal X less than a it will be looking, but the probability that X is equal to 0.

So, it will be constant between 0 to a, because between 0 to a, cannot take any values, but if I say probability X is less than or equal to a, then which includes this probability also, because probability that X is equal to a, is this much. So that, in the discrete case this continuity means that the point x equal to a has a finite probability.

In a continuous random variable if I ask the question, what is the probability that probability X equal to a, the answer will be 0, what is the probability that X lies in a finite interval. I can compute from the distribution function. For a continuous random variable will be P X minus P X, it will be equal to 0 and so on and so forth. Now, I can go on, but let me just some close on a relative concept of a probability density function. It is more convenient to work with is denoted by small p of X and is defined as a derivative of the distribution function.

Let me, ((Refer Time: 47:04)) just I think I will start from here next time, but what I it will be nice, if you can now quickly review as much of probability theory as you can from any book which you have access to t is given in Lattice book the review you could read Lattice book or you could read Hikes book the older edition. And if you can do that then it will be easy for me to it will be more meaningful then we complete this review very quickly in this class.

Thank you very much.