NPTEL

NPTEL ONLINE CERTIFICATION COURSE

Course
On
Analog Communication

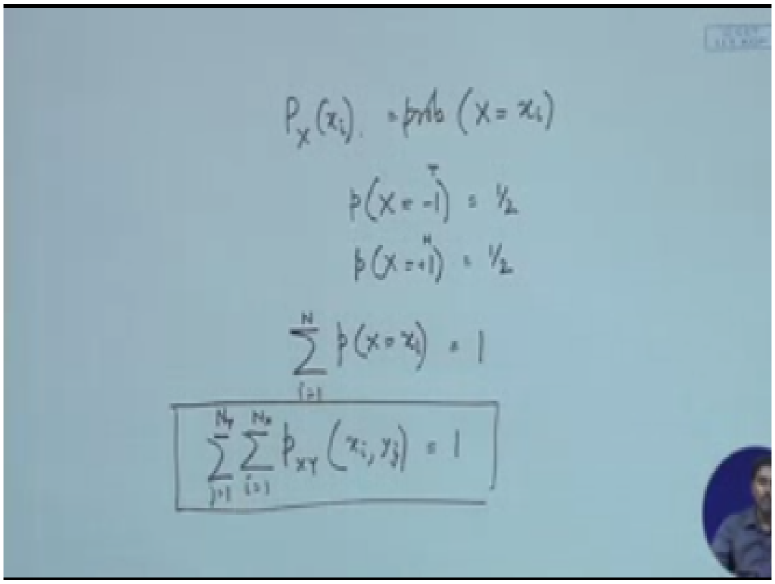
By
Prof. Goutam Das
G S Sanyal School of Telecommunications
Indian Institute of Technology Kharagpur

Lecture 33: Probability Theory (Contd.)

Okay so I think we have in the last class we have discussed already about basic theory of probability right so that is something we started discussing we have already proven Base theorem we have given the definition of random variable that is just like a mapping from events to value in radial axis so that is something we have started and then we define something where the random events are discrete that means it is countable, so whatever events that will be happening that is countable right you toss a coin.

So you have only two outcomes that you can you can map to a real axis maybe put to 0 or 1 or means head and head or minus 1 or 1it is just countable so for that we started defining corresponding probabilities of each events okay.

(Refer Slide Time: 01:22)



So we have given a definition that is called where X is the random variable it is actually saying the probability that that random variable X is equal to a particular event X I or the real axis representation of that event X I what is the probability associated probability of that and we have also told because if we just consider all the events and they are probability if we just sum them that is a certain event because we must be seeing something is happening okay.

So like we say probability of head that is let us say x = - 1 that is probably tail probability of tail is 1/2 and then we say probability of head that is +1 that is equal to1/2 so if we just add these two that means something will be happening so that is a certain event and that addition of probability must be 1we have also given a fundamental theorem of probability theory that is called this summation when it is taken over all the events that can be occurring in that particular experimentation that must be always one.

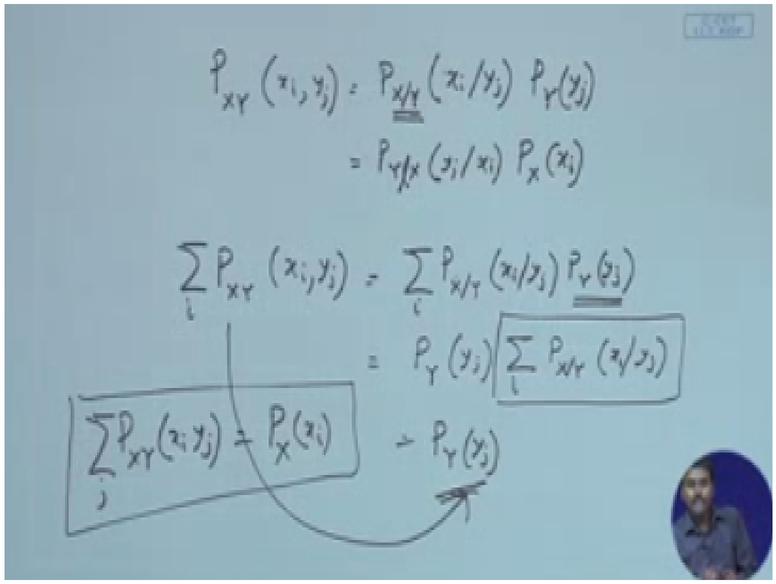
So probability actually sums up to one the way we have defined probability of a random variable okay or associated probability for a particular random variable for a particular value when it takes a particular value. So this is always true that is something we have already discussed okay we have also discussed something called means same representation for joint event where you have two random variable for two events okay given one example of such thing.

So if we have joint event so we can represent it as this x and y okay these two are joint event that means X represent the random variable associated with the experimentation where we are considering that event X or collection of event X and then collection of event Y if you have that so to separate this one and then we say probability of X_Y X_I Y_I that actually means that probability that random variable X takes value X_I from its sample space and simultaneously random variable Y takes value Y I from its sample space okay.

So this is something which can be defined this might not be just same I this might be also Y J because that might have a different variability altogether okay so this is called that joint event and associated joint probability okay and similarly if we just sum this over all the possible combination of X and Y's we must get one, ss if you just sum it will be a joint summation because I have to do combination for all x and y so this is for all I going from 1 to N and all J going from whatever value it goes so let us say this is an X and this is an Y if we just take all of this that must be one okay.

So this is for joint event and accordingly we can keep on augmenting that we can have as many events jointly happening as we wish and accordingly the joint distribution will be formed okay so after defining this with this same definition of notational aspect we started defining the conditional event right.

(Refer Slide Time: 05:15)



So we have also told that the conditional event which is represented as this, so this X given Y it is actually a conditional thing that means there are two random experimentation simultaneously going on which have some random outcomes one is represented by a random variable X which is already mapped to a rail axis and the other one is represented by random experimentation which is a random variable Y which is mapped to again real axis okay.

So X given Y means that I have already observed the event Y which is taking value equal to YJ given that what is the probability that I will be observing that X is happening to be X I that associated probability this is that conditional probability we talk about okay. So whenever we say this is conditional probability so if you just sum this over I okay so what do you expect this is a conditional probability that means y YJ that is already given okay given that all the possible probability values for seeing the event x equals two x i okay.

So if I take all of them that must be again giving me probability of one because given that y J has happened if I am just trying to test what's the Associated probability of each of the possible outcomes of X I and their associated probability if I add all those that must be a certain things because something of those X I will be happening anyway, so that probability becomes one very simple example that we have taken probably we had two giant observation you remember that we were just trying to roll a die and then we were saying whether it is even or odd that is one event and then there was another event that it is less than 4 or greater than 4 right.

So if suppose I say it is already even that is known that is J for me that it can take two outcomes even or odd so I am just saying that given that it is even okay that means only two four and six are possible outcomes so only those cases I will be taking, so I will be rolling the die it is a random experimentation of course but I will be only marking those events where either 2 4 or 6 is happening all other events are nonentity for me because they are not part of this experimentation okay.

So I will just take those things and start counting that frequency definition of probability theory and then what I will say I will try to see a particular XI is happening whether it is less than 4or greater than 4 okay, these two things are there so here if you see less than equal to 4 R 2 so there are 2 favorable outcomes towards this and 1 favorable outcomes which is not towards this so this X_I, I can say or this X_I can say it is less than this or greater than this okay. So if this is my definition of two events and accordingly X is defined so X_I will take two values only let us say 0 and 1 again I am mapping.

So 0 means actually it is less than 4 okay less than or equal to 4 1 means it is actually greater than 4 and then if I just evaluate the calculation or probability calculation what is this suppose let us say 0 given it is even okay or let us say even I represent as one so even an odd these are two events which are mapped to the real axis I say this is 1 this is 0 okay as I have told both the events x and y or both the experimentation x and y and associated random variable must be going to the real axis so this has gone to the real axis I have just denoted 1 & 0 similarly less than equal to 4 and greater than 4these are the two outcomes this must be mapped to a real axis so let us say I put it as 0 and let us say I put it as 1okay.

So whenever I write this that actually means that 1 has occurred in Y so that means this has occurred so it is already even-numbered given that it is even-numbered what's the probability that this has occurred okay that means it is less than for this probability I wish to calculate I have already seen that there are 1 3 5 also out of them these are only my favorable condition towards this Y event that y equal to 1out of them what's the probability that this X will be 0 so there are 2favorable outcome towards that and 1 non favorable they are equally likely because the faces are equally likely to come.

So I can very fairly say that should be 2 by 3 that probability because 2 favorable and overall means possibility are 3 similarly P X Y if I wish to calculate 1 1 that must be 1 / 3 because 1 means this greater than 4there is only one favorable case out of3 so that must be 1 by 3 if you sum them that should give you one okay. So with that simple experiment we can actually again map it towards a conditional probability calculation and we can get back to this theory that is very important that conditional event if the condition upon which it depends on that is already fixed then if you just take a summation over the one which is dependent on if you just take a summation of all the probabilities on that random variable you must be getting one.

Similarly it should be also true if I put it like this fee Y given X suppose it is the same experiment I just now turn the condition okay so given that it is less than equal to four what is the probability that this is even or odd something like that okay so then YJ must come over here and X_I so now it is given X_I I am summing over all J this must be also giving me one that should be always happening you can again test it with that same experiment we have done okay. So you will see that this is happening for all possible cases okay so these two are very important relationship and then we have also understood the base theorem okay.

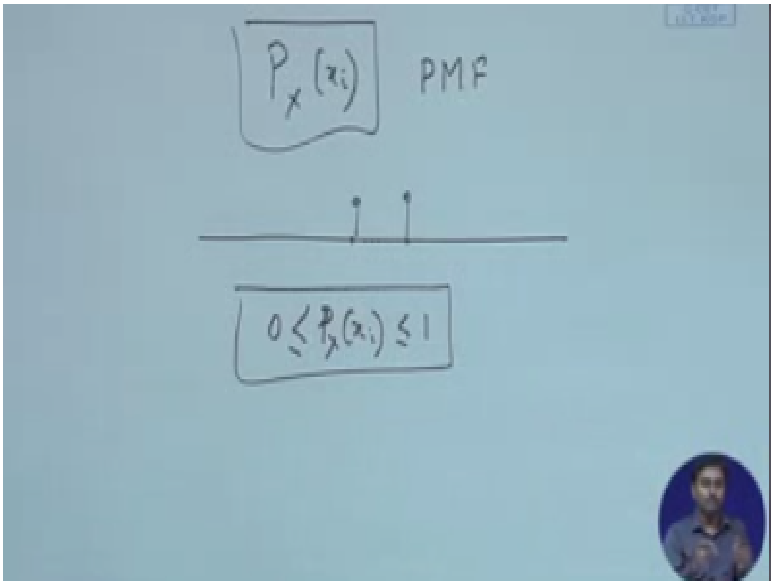
So what is base theorem Base theorem we have called this is the joint event we have already talked about X I and let us say YJ according to base theorem that must be the conditional one x which is the independent thing means on which the whole experiment depends so P X given Y suppose X is conditioned means over Y so that should be X_i given y Ji Py YJ be careful about the notation whenever it is conditional it should be like this that random variable x given random variable y whenever it is not conditional and it is a single random variable so it should be like this whenever it is joined it should be like this okay so that is the notational means representation okay.

So this must be happening and it should be also true for Y given X so base theorem tells me that this is true right so now let us try to evaluate something like this P XY that is the joint one I am taking and I sum it over one of them ok let us say I sum it over I so I can write over I I will be summing P let us give take this definition this X given Y okay, so X given Y so I am just writing the expanding the joint distribution okay our joint probability with the condition 1 so X i given y j py YJ I can write it this way okay.

Now this particular part that does not depend on I so I can take that out so I can really take this py y J out summation i p x given Y X i given YJ now this already we have proven that this should be one just now proven that okay, so if this is 1 this must be py YJ that is a fantastic thing what is happening if you take the joint distribution and you we call this as marginalization if you just sum it one of the random variable okay all possible cases of that will be getting the marginalized distribution recall that.

So this is actually from the joint distribution we can get the individual distribution similarly we can do it for X also so we just have to sum it over J, I will be getting px X I you can you can try this just have to use the other one probably oaky. So this is how you can always any joint distribution given you can get the individual distribution that is possible if we call this distribution okay basically this is called actually our probability mass distribution probably we have not mentioned that so because this is discrete event.

(Refer Slide Time: 15:36)



So whenever we get for a particular random variable P X X I we know all the values of this for every possible values of X I we call this as a PMF or probability mass function okay so and if you just plot it should be like this all those X I they are now we have mapped it to a real axis so there will be some value some discrete value over this particular thing and there will be a Associated probability value which will be plotted over here and we know that the sum of all this probability must be one that has to be and they must be always positive.

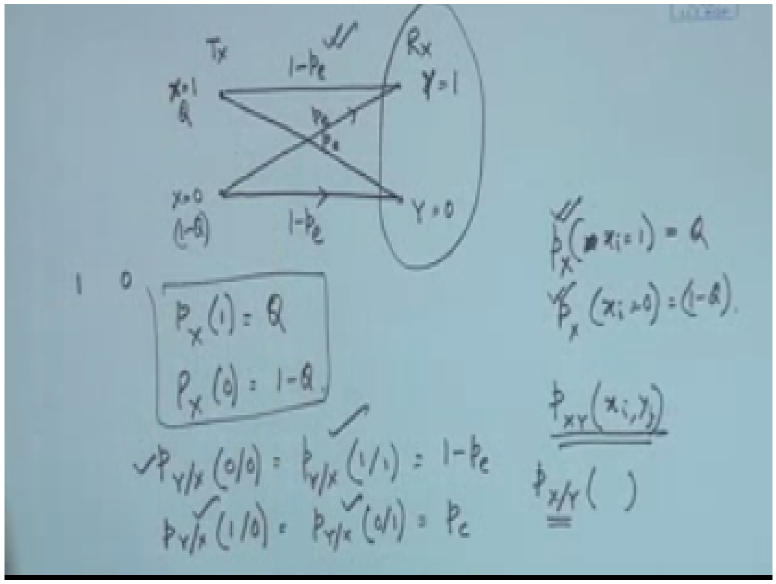
Because it is probability value probability cannot be negative because you will be see the frequency definition of probability says that I will be doing this experiment for let us say infinite number of time out of them how many will be in favor of my X equal to X I okay that can never be negative neither that N can be negative nor that how many means with what frequency I will be getting a favorable outcome that can at least be 0it cannot be less than that I cannot have a negative event happening the negative count of event happening.

So I will be always getting a positive number so definitely the probability value associated probability value whether it is small or big it should be some positive thing and sum of positive should be giving me 1 so therefore I can very fairly say from this that P of X XI must be lying between 0 and 1 because if their sum is 1 individually they must be fraction okay so this is always true for any probability mass function or any probability value okay so these are something which are already known to you okay.

So what we have now learned is we know how to define in a for a discrete random variable how to define Joint Distribution and from there how to marginalize to get individual distribution of each of the random variable that is something we know and The Associated conditional distribution how it is related to joint distribution this is something we have already derived you might be now asking all this Base theorem conditional rule what that has to do with communication.

So let me just give you a simple example which might not be having anything to do with analog communication in particular it is mostly the application is in digital communication but you will see later on when we will be proving something on analog communication there are some requirement of that also but I will just give this example because visually this is very satisfying okay.

(Refer Slide Time: 18:22)



So let us try to take this example that this is my transmitter and this is my receiver so from the transmitter side why I am giving this digital example because you will see that there are only very possible number of symbols that can be happening, so definitely we will have to calculate probability values for a less number of events okay. So that why that is why it is little bit easier to tackle, so let us say from the transmitter side I can only transmit two possible kind of signal okay one is high voltage or let us call it one and one is a low voltage or almost zero voltage which let us call that zero okay.

So it is almost like a binary bit stream which you are trying to transmit and that probably carries information as long as it is properly encoded okay, so it is streams of ones and zeros I am trying to transmit with some symbol okay so let us say at the transmitter there is a because we are transmitting events so there is a random occurrence of these events with let us call that a variable X X can take 1 or X can take 0 so these are the two possible outcome and associated there are some probability okay.

So let us say with Q probability at the transmitter side it is 1 and with 1minus Q probability it is 0 because only2 are there and that probability sum of the probability must be 1 we have already seen that so this is the only possible scenario okay because we only have two events or two sample values okay. So this probability that X_I equals to one okay, so this is Q and probability that X this X I = 0, 1 -Q okay these are called the prior probability what does that means that while transmitting this is the probability with which I generate either0 or 1.

So that is why these are called means without seeing what is happening at the receiver this is what is actually happening at the transmitter, so this is called the prior probability that means before transmission this is this is the statistical balance between the events okay which we are trying to transmit and decode on the other side.

Now we have already talked about channel so will be this 1 or 0 whatever it is it is a voltage level or some specific type of symbol whichever way we represent it when we put it in the channel okay a channel will do its own corruption it will probably add noise it will distort things and all those things will be happening and at the end that the receiver will employ something to detect it okay.

But what will happen because the channel has already corrupted it I might have a wrong detection at the receiver side okay, so there might be erroneous detection at the receiver side so what might happen? Suppose I transmit and let us say this I still say this is y okay random event Y which is the reception okay now because this is already random in between whatever things are happening that is also random so what I will be receiving that must be a random things okay.

So that might be one and that might be also zero so I can only I have only two possibility I already know that whatever he is transmitting that is limited to one or zero so whatever I will be receiving somehow I will be encoding this and I will be detecting only one and zero okay, so this is two things I will be detecting now what might happen because of the channel there might be error.

So that I transmit one and he also receives one this has certain probability so let us call this one - PE where PE is the error probability so immediately I can see that I transmit one and zero is received that must have a probability of P because again these two are the only possible that once I transmit one there are only two possibility that either I will be receiving one or I will be receiving zero.

So there are two possibility and we have already told that sum of probabilities must be one so if error probability that one gets flipped to 0if this is PE somehow I will have to calculate that you will see in your digital communication how to calculate that so if this is P that must be 1- P let us say that same thing happens for the zero also. So the probability that I will be correctly receiving it that zero I transmit and I receive 0 that is 1 - PE and I will be it is wrongly receiving it is okay.

So that is what happens in the whole scenario now of course we will be asking how do I evaluate PE that is a separate domain okay we will have to see how do we evaluate those things but that will not be part of our discussion in analog communication that should be the part of digital communication discussion okay. But whatever that is this is the scenario okay. So now let us try to define whatever we have got so far so if you see over here there are two random events okay, one is this X which is one of the random variable and another one is the Y what is transmitted and what is received okay both of them I have now mapped into real axes it can take values 0 and one other one also can take zero and one okay.

So I know the prior probability which is px 1I already know okay px that that random variable will be taking a value 1 that I already know that I have told Q and P X0 that is 1- Q right now over here whatever I know those are the conditional probability if you see very carefully so I am just mapping whatever we have learned to a physical example okay. So this is the conditional probability that given 1 has been transmitted what is the probability that I will be getting 1 or 0 okay.

So I can immediately write that P Y given X 0 0that means that X that is the given one that if that was 0, 0 was transmitted what is the probability that I will be getting 0 that must be according to our definition 1 - P same thing will be happening according to our definition it is symmetric actually 1 and 0 are both behaving in a symmetric manner due to the channel. So P that Y given X and I will be transmitting 1.

And I will be receiving also 1 these are all 1 - P and just the other thing that P Y given X I will be transmitting 0 and I will be getting 1or py given X I will be transmitting 1 and I will be getting 0 this is P so what I can see that I already know all the possible conditional probabilities right. S these two things I know I know the prior probability now I will apply Base theorem so I know Y given X I want to now know what is the probability that X given Y okay.

So this is something I wish to know and also I wish to know what is the probability that I will be receiving 1or I will be receiving 0 that is actually the receiver probability okay so that is the margin probability coming from the Joint Distribution so if I can define now P XY and this X I Y I or Y J if I can define this I have already talked about marginalizing. So I can marginalize it over X and then I can get py okay.

So there is a possibility that I can now get like prior probability this probability of earlier I would I was defining probability of transmitting something now due to the channel there will be something else which will be happening now I wish to know what is the probability that I will be receiving one and what is the probability that I will be receiving zero so that is something now I have enough tool to calculate that okay and you will be asking why I need the reverse thing that means P now I will be evaluating X given Y with some values, why do I need this?

So in detection what mostly you are interested that at the receiver I detect something okay after doing all those things I just see okay this is 1 or 0 finally I will be taking that now whether it is if I detect that as 1 or 0 I need to understand that okay I have received one but I have a doubt because the channel was creating errors I have a doubt whether one was transmitted or not now I wish to know if I have received one was it one and with what probability was it one actually transmitted and with what probability that should be given by this px given Y.

So now given thing is known to the receiver because receiver will be receiving that receiver does not know what was transmitted he is trying to guess what could have been transmitted, so from this he will be able to backtrack or back calculate this particular information okay and then he will be able to guess probably what kind of error he is expecting right.

So after detecting was the overall error that he is making because once he has this probability value he will be knowing that how many of those things means if he knows the probability of error he will be also knowing if I receive this many out of the definition of frequency definition of probability we can always get to know how many of them will be in error okay.

So this is something he will be knowing and then he can actually try to manipulate that but if he does not know this relationship he will not be able to manipulate that manipulation means probably he has to design a better detector he has to probably use a different channel where the error is less or error is more or some other things he has to put over there okay. So all those things you can do only if he has or the receiver has this understanding okay.

So that is why it is very important that we have this mapping all the conditional probability marginal probability and the overall distribution that is well known and if one is known how to calculate the other one that is pretty much required in any communication. So what we will do now we have so far done even the example done for discrete case we will just now try to see what happens for a continuous case which will be more interesting for us in analog communication specially, thank you.