NPTEL

NPTEL ONLINE CERTIFICATION COURSE

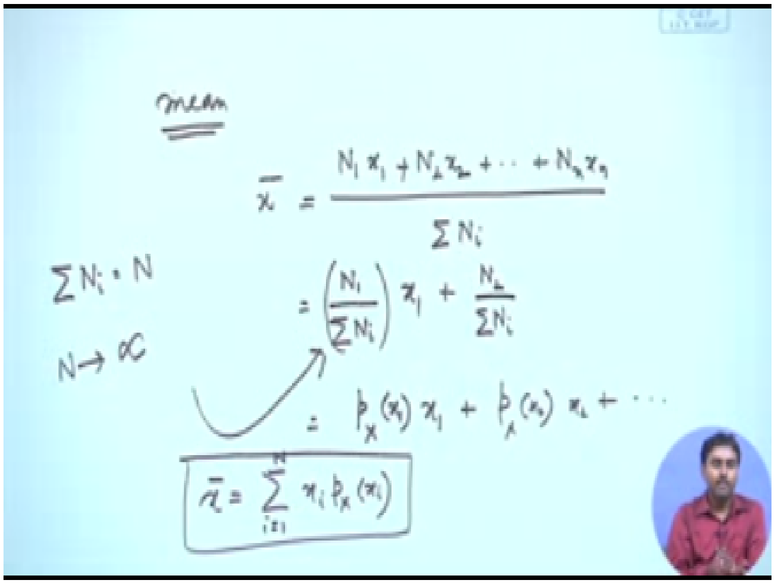
Course
on
Analog communication

by
Prof. Goutam Das
G S Sanyal School of Telecommunication
Indian Institute of Technology Kharagpur

Lecture 36: Probability Theory (Contd.)

Okay so for Gaussian random variable we have already started discussing about means, standard deviation, so let us try to characterize those things now okay. What do you mean? And grow there are related to corresponding PDF and CDF, so that is something that which we will target in this class. So let us try to see the basic and classic definition of mean.

(Refer Slide Time: 00:49)



It is just the average, how do you do averaging, we actually suppose I go to the class and I say that I want to get average height of all the students, so what I do? You actually ask every student their height, you note them down, and you add them together and divided by the number of students' right. That is how we calculate average, so this well knows thing okay. Now suppose we want to do this in a little different way, we say we want average height but we want to just do this with just the integer number.

Okay so it is 5 feet, 6 feet, 7 feet, something like that okay, we will have no other things. So basically what we will say that if somebody is having height from 4 to 5 will report that as 4 okay, if somebody whoever are falling from 5 to 6 will report that as 5 and so on. If I just do that then it will be not a single observation for each one of them okay. so will have to then see, how many of them are falling under that category and then according to that we have take.

So what will happen? It is like this will probably after doing that we wish to calculate that average, so then we are saying, okay how many of them are between 4 and 5, we will take that number. Let say that number is N_1 okay, so and each of them we are evaluating that is to me is 4 or I can take that as 4.5 also whichever I take that N_1 and N_2 whatever the value, so let say that is x_1 I will do that. And then I will also calculate how many are from 5 to 6, so I will take the number I will multiply by the number that it represents okay.

It might be 5.5 okay, something like this and I will do for N number of such things, so let say Nn number last one I get x_n and then I will divide the whole thing with the number of students, so it is nothing sum of this Ni. So in h category they were N1, N2 or N3, N I number so I will divide by this. So that gives me a sense of average, if I do it this way okay. Otherwise it is also true, this was a specially created experimentation but if I wish to calculate suppose I take a coin and I say if it is head then it should be +1 or if it is tails it is -1.

We have already told random variable means it has to be map whatever the events are there that has to be marked to a real axis, so we have marked it – 1 and +1 and now I do this experiment for let say 5 lakhs time okay and then I count, how many heads are occurred? I multiply all of them with +1, and how many tails are occurred? I multiply all of them with -1 okay. So that as many numbers into +1, as many numbers into -1 and divide by the total number. That is how evaluate things, so it is always true.

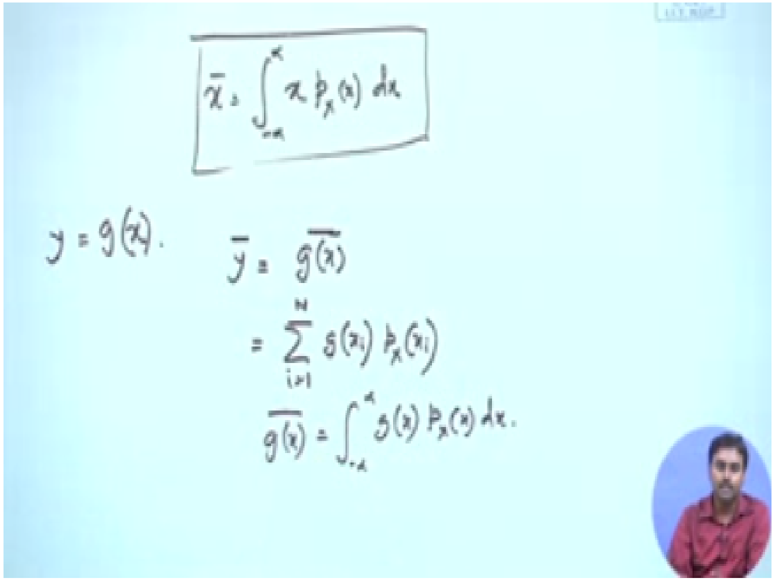
So for the dye also it will similar thing there will be 6 such part right. Now let us try to see little bit more carefully, so this $N_1/\sum N_i$ that is factor which coming into x_1. We will always be getting these kind of factor, $N_2//\sum N_i$ okay. So the $/\sum N_i$ is actually the amount of time I have done that experiment, repeated that experiment, so that is N. my frequency definition of probability was that N_i take it towards ∞. So once this is going towards to ∞, what happens to this? This actually tells the probability that this x_1 will be happening.

That is what we have defined it okay, so for head and tail if we just take that from 5 lakhs to 5 core or even more then almost half the time it will be head, almost half the time it will be tail. So how many times it has occurred/ as many times the experiment has be run, so this ratio is always in frequency term gives me the probability. So this must be as/$\sum N_i$ this must be the probability of associated random value next one.

So I can write this as p_x of x_1, that x_1 will be happening x x_1 + this I can write p_x of x_2 x x_2 similarly it goes on okay. So immediately I can see the x bar is nothing but in a discrete case as many times I will be doing this experiment right, so that is the famous definition of mean. So mean is nothing but your random variables are there, all possible things you take and you multiply with associated probability, you add it up for all the possible outcomes whatever you get that is mean.

So from classic definition of mean we could see that, it is definition of mean, it is associated probability okay. if we just extent it from ∑ to ∫ that is how we go to continuous form, so that immediately that ∑ will become that ∫ and inside it will be associated PDF.

(Refer Slide Time: 07:17)



So I can write x bar if it is a now the continuous variable, so that should be the ∑ will become ∫ in the entire real axis where this can happen, x x px x dx right, so that is for the continuous, so we get a associated mean, so this how mean been calculated and that is why probably PDF is so important because mean is something which is called the average. So whenever there is a randomness happening you tend to get the average, try to guess the average what is the average of it, so on a average what should I expect from this random experimentation.

So whenever you try to guess the average you need to have this PDF associated probability density function of that particular or for the case of discrete variable you need to have that PMF, so either PDF or PMF will be searching because you know that through which we will be able calculate the associated mean, one of measurable parameter okay. So now let say if I have a function of a random variable okay, so what do I mean by that?

Suppose this x is the random variable and what I get is the function which marks that random variable to some other axis. So it is just the functional mapping, which ever you wish to do that okay. so suppose there is something called temperature variation okay, so something varies from the temperature square okay. Now temperature in a day is a random variable, you take that and your that targeted variable is varying with temperature square.

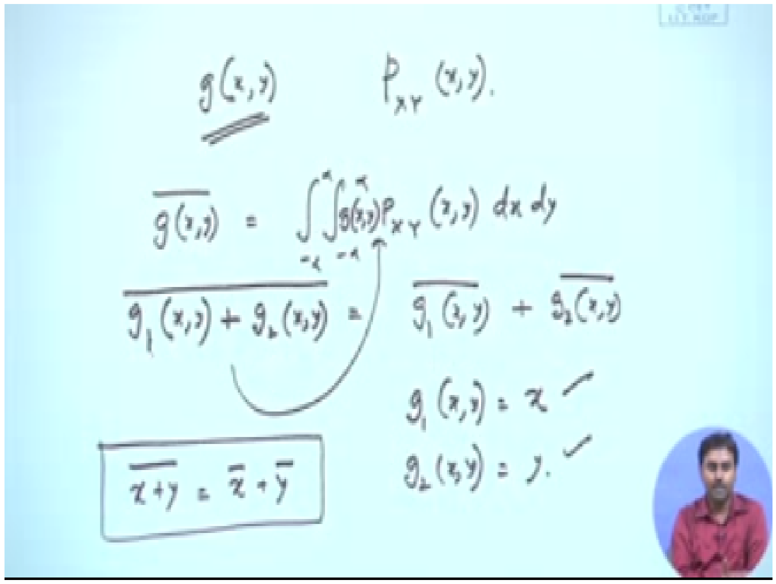
So it is mapping that random variable to that particular temperature square variable what you have to find oaky, so that kind of thing. So that will be our definitely another random variable which completely depends on input random variable okay, so that if I just say y, I might have to finally, I might not observe the temperature but I might observe that outcome which is the temperature square probably okay, so that is how I can measure things.

So I might have to get my y bar, whatever I am getting the random variable, so this y is also random variable which is just an eventually function of random variable. So I need to evaluate the mean of that, so what happens, if I wish to calculate this y bar this g x bar, so whenever there is a x happening right, to calculate the average I was multiplying x with it associated probability. Now this x will change to g x, the associated PDF is still on x, the value of x that is the, if I just correlate this probability to frequency term that should be the case.

So I can always write this as ∑ if it goes up to n, so g x_i x it is associated probability, so this is something that I can write or in a continuous one it will just be the integration. So g x must be -∞ right. so similar definition it is instead of x it has become gx, so this is more generic, if gx = x then I get the earlier definition that is the because it is a direct mapping, it maps to itself, 1 to 1 mapping and maps to itself only if the x = gx.

If any other thing is still have the some kind of mapping and I will take that associated probability and I will still get the average of that okay. So this is the functional average we talked about, now let say suppose I have.

(Refer Slide Time: 11:43)



A joint thing okay, where there are two random variable x and y and they have a joint PDF which is defined as Pxy x, y and I also have a joint function; I want to evaluate that associated mean okay. That is nothing but same thing I have to multiply for each x,y with it is associated PDF and integrate over the entire means domain of x and y, so I wish to calculate g (x,y) that must be -∞, to+∞ this Pxy and this is should be multiplied with d xy right. That is the average of g xy that we will be able to calculate, if we just do it that way okay.

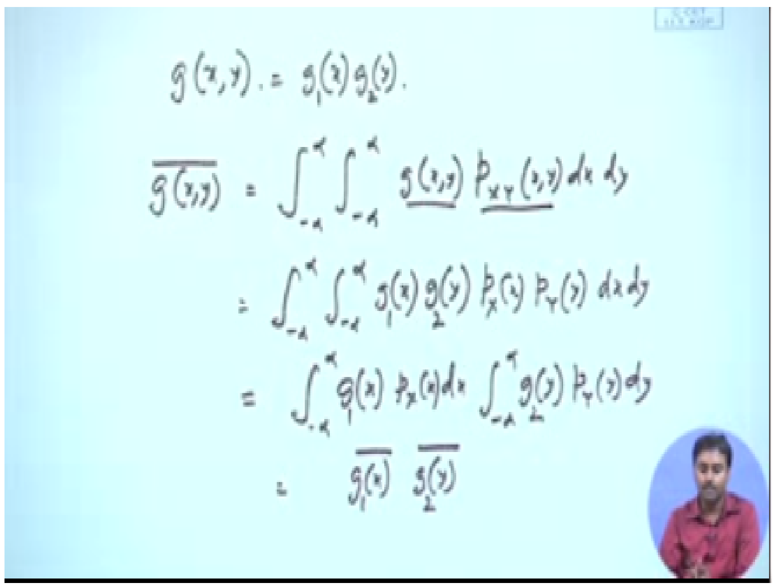
So if this is the case I can also now start writing if I have g1 x,y that is 1 mapping and another g2 x,y another mapping, random variable remains the same. I wish to evaluate the average of this one, so this is very clearly it can be seen that, then two integration will be separated out and I can immediately that it is nothing but g1 x, y bar + g2 x,y bar right this is very true because now you once replace the whole thing over here then two integration will be separated out.

So any functional mapping if two addition functional mapping I wish to get the mean of that, I can separate out as if I calculate the mean of the 1^{st} function and functional mean of the 2^{nd} function, so this is always true that is pretty obvious from the integration definition or true PDF the way we define mean right. So now I can just say that this g1 x, y is nothing but x and this g2 x,y is nothing but y. I can have this functional mapping right.

So it is always possible, it is a function of x and y but the probably y = 0 always, this is also function of x and y but x = 0. So this can happen this is a valid function, if I just put it over there, so what will happen x + y bar is always x + y, so this is always true. For this is to happen you do not need anything remember no independence among x y and all other things are required. Mean is always additive, if we take 2 random variable whether they are independent co related whatever it is, you don't care you can separately calculate the mean and you can always add that.

So this something which is very important, if you always remember that for mean you do not have to think whether those two random variable are at all dependent on or independent or any other relation, you do not have to care about that. You can always say okay if I can individually get the mean of these two I can always add them up that should be the overall mean, so this is always true okay. Now let see if this g x, y is nothing but.

(Refer Slide Time: 15:42)

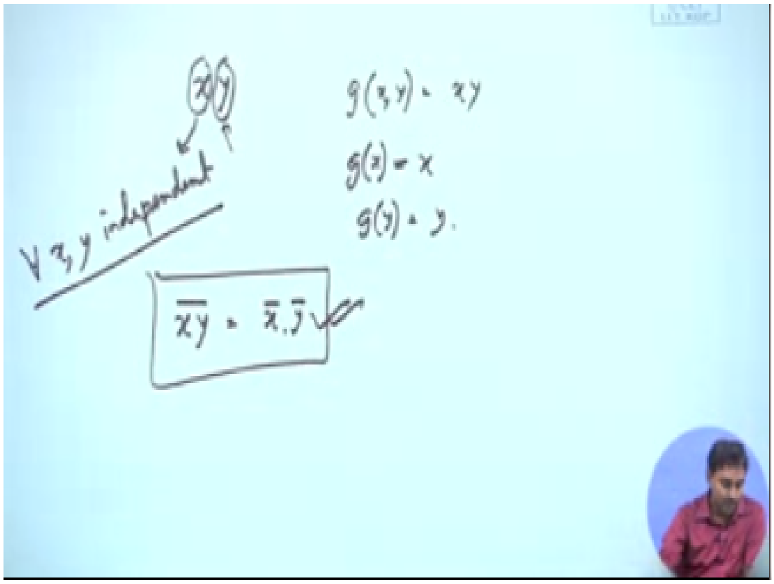


A product of 2 functions where they are separately just function of x or y okay, if this is the functional mapping and this is the nature of these two function then what how do we evaluate the overall mean value. So let say I want to still get g x, y or this particular function the mean of that right, so I can again write so this is g (x, y), P (xy), x, y dx, dy. Now I know that this can be separated right. Now the question is if this can be separated?

That is the condition where we will be able to separate these two if x and y are independent, immediately I have proven that it can be written as Px x and Py y, now the ∫ can be separated out, so I can write this a gx + P x x dx ∫ - and +∞ gy x Py y, so this is nothing but gx. So this will always happen, remember not necessarily this grand function as to be same mapping g. so in that case it will be just g1, g2 whatever it is, so the mapping can be different.

The one we have taken probably that is the specific function we all know that for Gaussian that is what happens but you do not have to that kind of similar mapping. Any way whatever the mapping is you need to be or it should be possible that you separate them out okay, so and you can put them as multiplication where 1 part is just the function of one of the variable other part is just the function of other variable. As long as this is possible and if random variable that is x and y, those are independent then you know that overall function mean is just the multiplication of that mean. So if now that function I take x and y it is separable exactly.

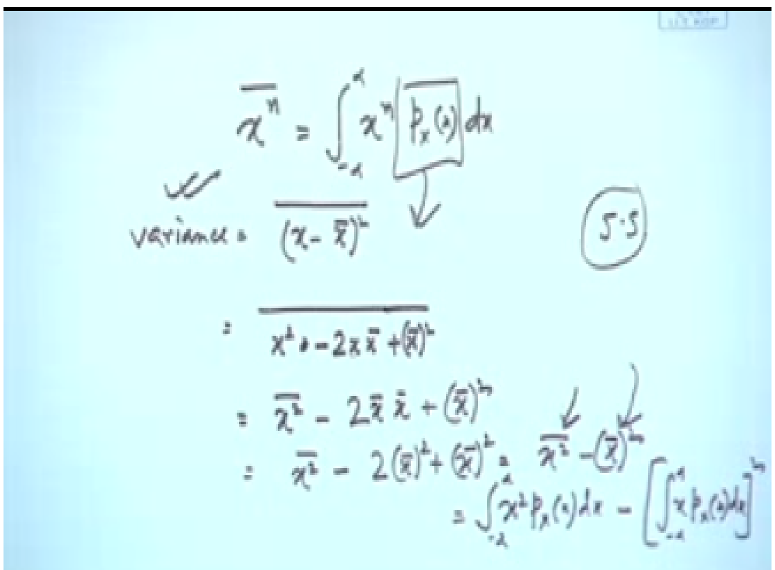
(Refer Slide Time: 18:31)



Because you have just this multiplication this, this is just the function of x, this is just the function of y that is happened, so by g(xy) is nothing but xy and gx is also the typical example where both are reexamining as same function. This is x gy and y okay, as long as this is happening I can write this when if x y are independent. So this is the strong condition if two of the random variable are independent then they are mean of multiplication it is nothing but multiplication of mean. This is a very strong condition for xy, so this always happens okay.

Now let us talk about some of the higher moments, so mean is we called that, most of you will be knowing already that is the 1st moment actually, so higher moments are defined as this.

(Refer Slide Time: 19:54)



This is by definition x^n okay, so higher moment why mean is called the 1st moment, it is just because if you just put 1 you get the equation of mean, so that is why mean or average value is the 1st moment and then if you just put it has 2 it has the specific mean, so that is the 2nd moment and from 2nd moment you can actually get the variants or standard deviation that we will talk about now.

So when we talk about variant, where we actually mean by this, this as a big significance in any statistical process, what it is saying? It is actually whenever we write this we are taking a average okay, so bar means average. This x is the random variable right we are picking different a values what we do we first calculate the mean and here what we are trying to do for every value of x we are actually taking how far it is from the mean right.

And square that because if they are on both sides of mean than they will cancel each other, if I just add them together, so we square and add them add all of them, divide by as many random samples are there divide by that N I will get this average here. So this is actually saying, it has the physical significance, the saying how far it is deviating from the mean value. The square of those deviations so basically you are taking all the samples if this value is very low that means immediately you can say most of the values are lying very close to the mean value.

If this is very high that means it is flatter we are talking about the Gaussian with σ^2 low or high that is the case. If it is having a higher value that means the most of the values are actually lying away from mean. I might I suppose the height of a particular class or section that is reported from it is all over the place, suppose let say from 4 to 7 feet, everything you observe okay. then you can say that may be the mean value is still 5.5 or .5.6.

But there is a huge deviation, this might happen where as if I just do it for the age in a particular class what will happen they will all running very close to the average age because generally in a class we just give admission if their ages are almost similar right. so if you do that sp for age what will happen, you will immediately see that their ages are not deviating from the mean value. Whereas height or any other weight if you start observing that you will see a lot of deviation.

So in that case just mean does not characterized everything, for a particular random experiment if the mean suppose for the 2 random experimentation the mean might still remain same but it might deviate quite small from the mean value or it might deviate heavily from the mean value. So these two samples are quite different, if you just observe those two samples are those two random outcomes we will see there, they have a huge difference, probably mean will be still the same.

So a typical example if I give, if you take age of a particular of class 6 okay, if you take the random samples from 6th standard or you go and take the age value of competitive exam and there is a coaching center over there and you are trying to take that age value, you will see a lot of deviation okay. So where as here if you just take those age it will be almost similar okay. So that is what happens, you can have similar mean value reported but you can have a large deviation.

So basically that characterized the randomness of the associated process, so that is why just mean will not characterizes anything, you have to go to this variance, it say something about but variance also it is not enough, we will see that all those higher moments that as to be characterized, then only the complete random process can be characterized it. The best thing to characterized random variable is this one. This provides all the information because once you know the distribution all is average things came first.

People started seeing okay if something is random can we measure some or guess something random variability, so they started with mean initially then they could see that okay if there is wider variability and smaller variability still the mean is same, they went to the second order. So then they could see that some more things more thing that could extract but we could see that you go to the higher order.

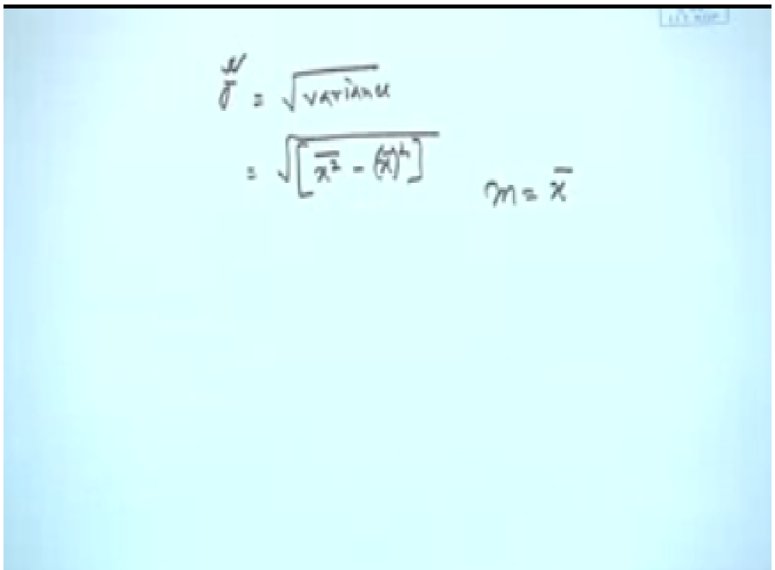
But finally people could understand it is just that the PDF which characterized everything because if you once you know this you can actually generate all those moments. So if some how you can get the PDF you have everything. So all the statistics related to it is already in your grip, that is why PDF is the relationship between and all those moments right. any way so what we were discussing that if I have the PDF, I can go to the 1^{st} ,2^{nd}, and all those things.

So this was my definition of variance let see how it is related to those 1 and 2ne order moments. So if just expand this I can write this is just the average value, so this can be x^2 – 2x x bar + x bar^2 remember here is the mean value that is not the random variable, so taking bar of that will return you the same value right because every time it will be same thing. So whether you take the average or any other thing, you will be getting the same value.

So now this averaging suppose it has the associated PDF it is just an integration right, so ∫ with +,- that will be distributed, so I can write this as well as x^2 – 2x x bar + x bar^2. So finally what do I get - x^2 – 2x x bar + x bar2. So once you have evaluated your mean and then through this process you have evaluated your second moment, so just the different of these two are your variance. Calculation of variance which is of measurable quantity people used to do, they use to always take every value from mean how far it deviates.

This thing can now easily be calculated if you know the probability of them okay so that was our target that we wanted to all these measurable parameter that people will used to do how do we correlate that with our this moment definition right. so this we can write as $x^2 P_x^x$ dx so that is all immediately we get our variance and standard deviation is nothing but the √ of that variance because just you do not want them to cancel it out, you have squared it. And then you have taken the average of square but if I want to see the actual deviation so I need to again do a square root of that. So that is characterized by σ.

(Refer Slide Time: 30:03)



Often that is √ of variance or $\sqrt{x^2} - x$ right so this is how we will get the σ, this was the σ which we are characterizing for Gaussian. So you remember there were two parameters for the Gaussian distribution, one was the mean which was x bar, so that m was actually x bar and you can verify that take that Gaussian distribution you plug in x into P_X^X integrate it you will always getting this f back.

Again do that variance calculation you will get back this σ. So this is what happens, in the next class what we will do, so far we defined all these things in the next class we are try to do is we will try to see some more relationship between these two particular bar that variance means, how they are related for a particular random experiment when they are independent, how do means we have already seen whether see it is always addition of mean.

But for standard deviation what happens or variance what happens can they be evaluated this separately so that something we will try to characterize and then we will probably go into slowly towards our new definition that is correlation okay that is something that we will try to do thank you.