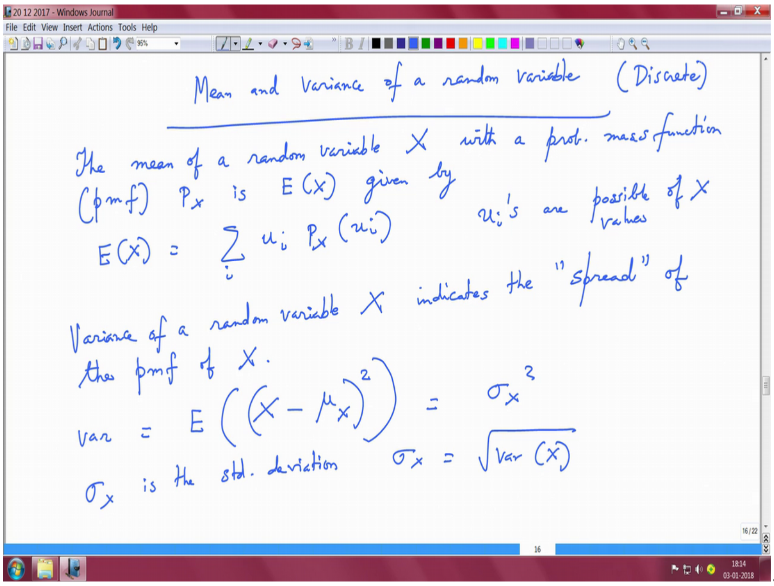
Mathematical Methods and Techniques in Signal Processing – I

Prof. Shayan Srinivasa Garani
Department of Electronic Systems Engineering
Indian Institute of Science, Bangalore

Lecture – 25
Mean and variance of a random variable

Ok. So, since we have understood how to sketch the probability distribution function, how to sketch the probability density function ah, what probability density function and distribution functions mean ah. We would be interested in some statistics of a random variable based upon what we know about the density function. So, often we are interested in statistics; such as mean um, the variance and so on and so forth. So, let us basically define these things right, understand what they, what they mean?

(Refer Slide Time: 00:49)



The mean of a random variable capital X with a probability mass function pmf P of x is denoted as a E of capital X given by. So, the mean is expectation of the random variable X given by this quantity summation over all i u i times the probability mass function of u i; that is these u i's are basically possible values of the random variable X. So, u i's are the possible values of the random variable X, and you are looking at the at the ah statistical mean over all these possible values of the random variable X.

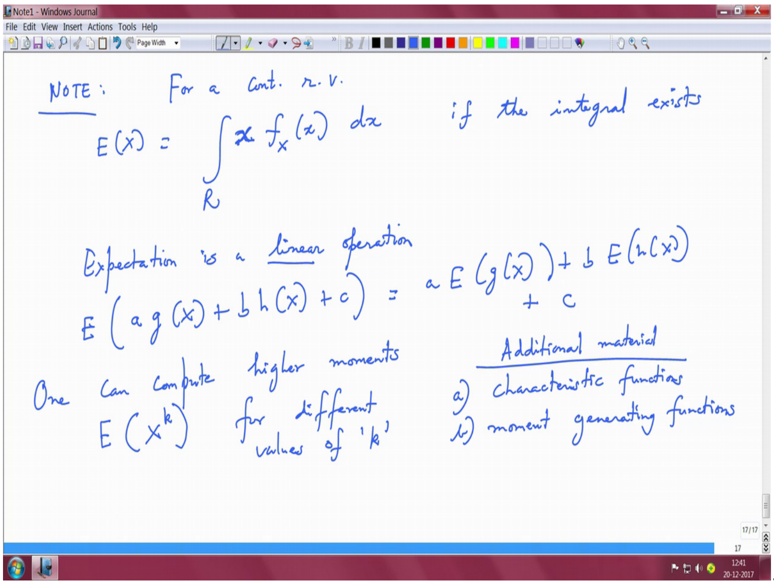
So, this is given by this quantity right, it is its, its weighing the possible value with the probability mass function, associated with this variable that is what this quantity is and then you average it. Now the variance of a random variable X indicates the spread of the probability mass function of the random variable X. So, I mean this is often interesting, because I think people like the mean and the variance , I think mainly from the Gaussian random variable, assume there is a Gaussian distributions are called normal distributions and for Gaussian distributions it is enough if we know, the mean and the variance and you can characterize the probability density function right.

And from that standpoint of few people like to look at it, look at the two statistical quantities which is mean and a variance. So, mean tells you what is the average, statistical average and variance tells you; what is the spread in the probability mass function for the random variable. And the, but this does not mean that these are the only two statistical quantities of interest. Most ah, for most practical purposes you know from data collection that you see from sensors and other things, and if the population follows a certain normal distribution then these two statistical quantities would be useful.

Now what is this variance? Now variance is expectation of the quantity X. So, you have the random variable, you subtract the mean from it and then you compute you square that that quantity and then you take the statistical expectation and that is what you get for variance right. And this quantity is given by sigma x square and sigma x is the standard deviation and it is given by square root of the variance of X ok.

Now for the discrete random variable, you have a mass function and then you can take a statistical average and you can get the expectation in this form through a summation. For continuous random variables the expectation is in the form of an integral. So, since the slide is all about a discrete random variable, I think I have to mention this for clarity that this is for discrete random variables, because we have defined the probability mass function for this.

(Refer Slide Time: 05:48)



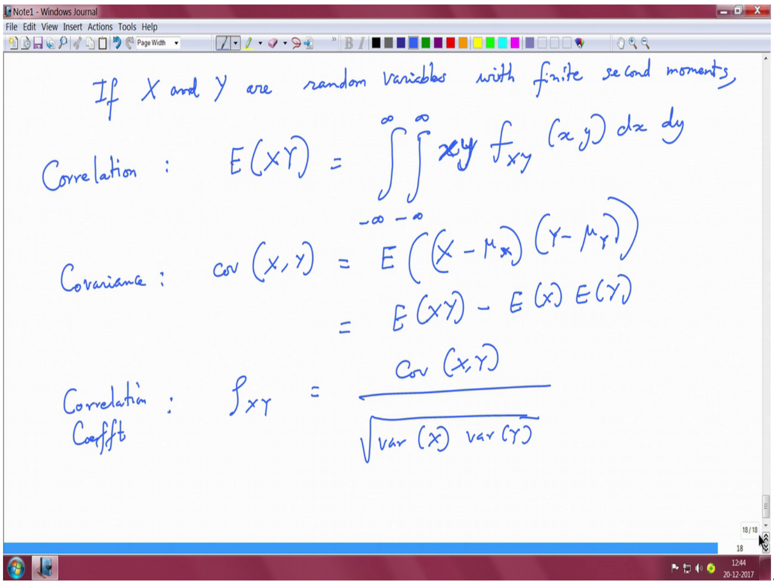
Let us note the following for a continuous random variable, expectation of the random variable X is given by the integral of X times the density function of x integrated over some region R. If the integral exists and this has to be a small x to be careful. So, the expectation of a continuous random variable X is given by this integral x times the density, a small x times the density of the random variable X , and if the integral exists then we can compute the expectation. And one other important property that you need to know is expectation is a linear operation you can prove this, I will leave this as an exercise as a sort of straightforward.

So, if I give you a mixture of, a function of random variable. Let us say we have a times some g of X plus b times h of X plus c this is some mixture and if I have to compute the expectation and I can do this as a times expected value of g of X is b times expected value of h of X plus expectation of a constant is a constant, so you can write it like this. So, this is a an important result, its a linear operation, but its a very straightforward and a trivial result that you can easily prove this. So, as I mentioned to you people are often interested in mean and variance ah, thinking about Gaussian distributions and people like to fit Gaussian distributions, because it is sort of a normal distribution.

Now there are very interesting properties from central limit theorem etcetera. I am not going to delve any of these in this in this is a basic background on probability because it is part of a separate course in itself, but it is not necessary that you have to compute mean and variance and they are enough to characterize the density function right, you can calculate higher moments. So, one can compute higher moments which is basically expectation of X power k for different values of k and then use these statistical quantities as appropriate for your calculations.

So, I will leave this as a additional material for your reading ah, you can look into starting words or any other basic book on probability Sheldon Ross etcetera. So, I would like you to read through the material on characteristic functions and moment generating functions to further your understanding on basic probability. So, and moments are particularly useful when we have to consider um, you know characterizing the pdf from observed data etcetera ok. Now, let us look at some other statistical quantities and define these if X and Y are random variables with finite second moments right.

(Refer Slide Time: 10:15)

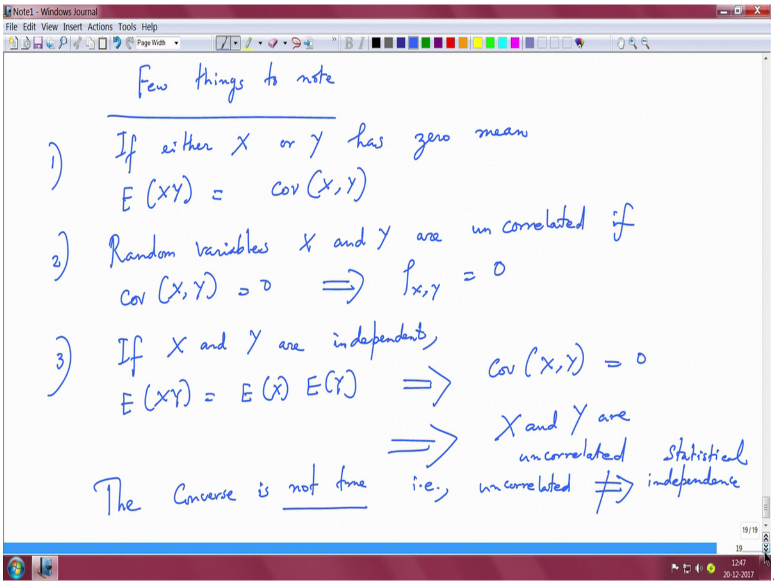


Then we can define the following ; one is the correlation which is given by expected value of XY and this quantity for continuous density functions can be written as follows , this is minus infinity to plus infinity xy times fxy xy dx dy. So, this is small x and small y ; that means, I look at the joint of x and y and I look at the expectation of the joint of x and y over the joint probability density function of x and y .

Covariance is given by, is denoted by cov of X comma Y. So, covariance basically you have to compute, essentially the correlation, but by subtracting the bias from the random variables X and Y. So, this quantity is basically expected value of X minus mu x and Y minus mu y. It is basically subtract the bias which is mu x and mu y from the random variables X and Y and then you compute the correlation for this quantity. And you can evaluate and simplify this as expected value of XY minus expected value of X and expected value of Y; of course, you have to use the property that expectation is linear and arrive at this result.

And along with this set of quantities, one of the important quantities is the correlation coefficient and correlation coefficient between X and Y is given by the covariance of X and Y upon square root of the variance of X times variance of Y. So, these quantities tell you how linearly random variables X and Y are related right. I mean this is this is one of the important metrics and this is a physical meaning for this quantity, and people often compute these quantities for various statistical purposes.

(Refer Slide Time: 13:31)

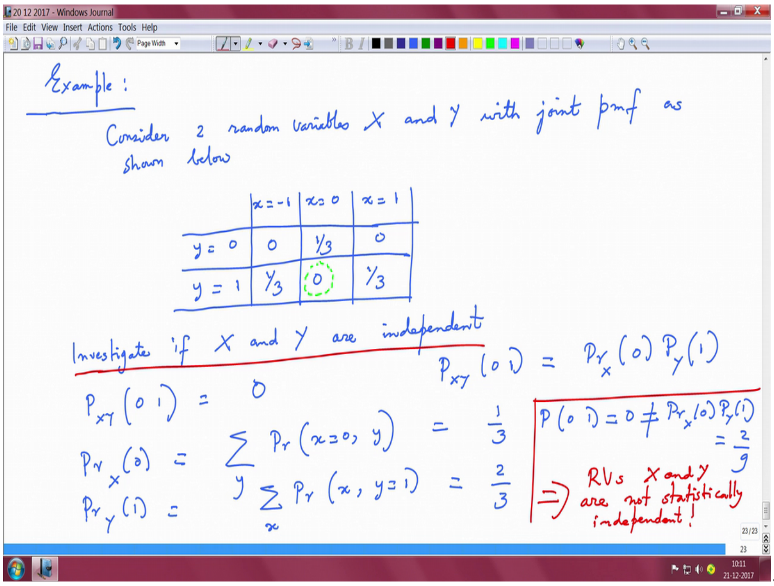


So, a few things to note. First if either X or Y has zero mean and then expected value of X Y is basically the covariance between X and Y. Random variables X and Y are uncorrelated , if the covariance between X and Y equals 0 and which also means that the correlation coefficient between X and Y is 0. This is very important thing when you talk about correlation you are looking at how linearly X and varies with Y right and if they are uncorrelated covariance of X and Y is 0.

There is a another notion of independence and we will see this. So, if, and what is a link between independence and correlation. If X and Y are independent, then expected value of XY is expectation of X times expectation of Y. Now this implies. So, if the expectation of XY is expectation of X times expectation of Y, then covariance between X and Y is 0 which means X and Y are uncorrelated. So, if two random variables are independent, this implies they are uncorrelated, the converse is not true; that is uncorrelated um, this does not imply statistical dependence ok.

We will see an example to clarify this idea and then we will get onto the next concept. So, we were discussing several properties and on probability etcetera, it running through the basics of probability and random processes. So, let us work out an example and investigate, if uncorrelatedness leads to a independence or what can you say about independence and correlatedness of random variables.

(Refer Slide Time: 16:47)



So, let us consider an example, consider two random variables X and Y with joint probability mass function as shown below. So, populate a table what X and Y can take , whose random variable x can take x equals minus 1, it can take x equals 0, it can take x equals 1. y can take y equals 0 and y equals 1. So, and populate the entries for the probability measure on the joint of x and y. So, this is populated as follows. A quick inspection into this table will tell you that this mass sums to 1, because one third plus one third plus one third is 1.

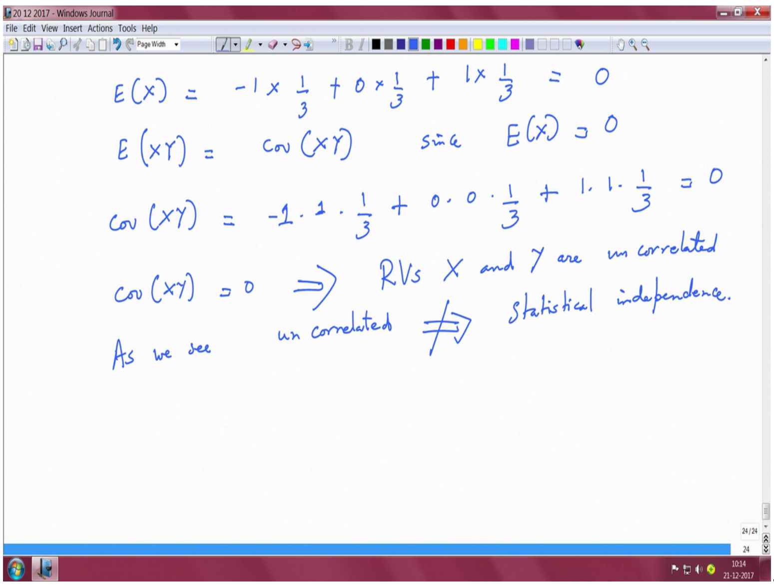
So, therefore, this is indeed um and in all of them are greater than or equal to 0, all the entries are greater than or equal to 0, this is indeed a valid probability mass function. Now let us investigate, if X and Y are independent. Ok let us investigate if the random variables X and Y are independent. Now consider the mass function where X takes 0 and

Y takes 1 right and this is basically this entry here, shown in the green circle. So, probability that X takes 0 and Y takes 1 reading off from the table this is 0 ok.

Now, they are independent if the joint mass function factorizes as this quantity here; that is probability that X takes 0 and probability that Y takes a 1 ok. Now let us compute the probability that X takes 0. Probability that x takes zero is basically this quantity you marginalize over all Y right if I have to write down the steps it is probability that x equals 0 over all y and then if you just sum this over this, second column you get this as 1 by 3. And the probability that y takes one is a summation over all x; that is you marginalize over this row here right.

So, this is probability that over all x y taking the value one and then if you add the entries in this row; that is a second row you get this quantity as two thirds. Now if you examine this equation probability that x takes 0 and y takes 1 is 0, which is not equal to probability that x takes 0 and probability that y takes 1, which is basically two nines here right, this evaluates to two nines and this is not equal to 0. So, therefore, clearly random variables X and Y are not statistically independent. Now let us examine if they are correlated right for which we are bringing these expectations.

(Refer Slide Time: 21:59)



Now, let us look at what the expectation of X is. Now X can take values minus 1 0 and plus 1. Now we have to compute the expectation of X and that evaluates to minus 1 times the probability that X takes a minus 1 which is one third, because again you use the same marginalization trick, then plus probability that X takes 0 is a one third again, and this is times one third plus one times one third and this evaluates to zero.

Then we look at the expectation of XY right. This is basically the covariance of XY , since we just saw that the mean of x is 0. Now let us compute the covariance of XY. This happens to be. Now you have to look at where the masses are masses exist that are a nonzero right. So, this is minus 1 times 1, X takes minus 1 Y takes 1 probability one third plus X takes 0 Y takes 0 , the joint mass is one third then X takes 1 Y takes 1 and the joint probability mass function is one third and this evaluates to zero .

Now, covariance of XY is 0, which implies random variables X and Y are uncorrelated. Now as we see uncorrelated does not imply statistical independence. So, I think this example is enough for us to validate this result right. It comes to by construction we can say that the implication does not follow. So, we just hopefully have digested the material on statistical independence, correlation and some of these ideas. We cut here yeah.

Student: So, it is not independent know (Refer Time: 25:04).

Which one.

Student: (Refer Time: 25:06) the three this, what you showed it is not the previous independence conditions it was not independent right.

Which would they are not statistically.

Student: huh.

Independent.

Student: Not correlative and then it is ah, these are these are non correlated and non I mean not independent, right.

Yeah these are not correlated and not independent yes.

Student: So, (Refer Time: 25:29)

Ok.

Student: Examples should have given there are there are correlated, but they are statistically, not statistically independent huh.

So, let us massage this a little bit.

Student: (Refer Time: 25:47) if it is statistically independent.

Yeah, if it is specifically independent it implies on that has been proved, but the converse from the other way round. If it is uncorrelated that does not mean that they are statistically independent. So, this is an example where they are uncorrelated, but yeah they are uncorrelated, and that does not mean that their focus. There should be a case where I should show that this is basically statistically. No I mean is it uncorrelated, I mean that does not mean anything I cannot say that it is statistically independent that is all.

Student: it does not correlated (Refer Time: 26:24).

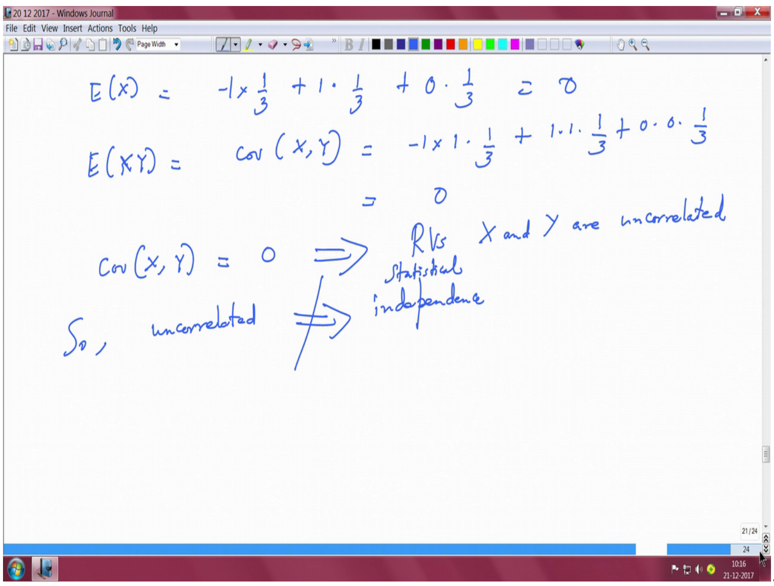
Right it is statistically independent. So, this result was proved here.

Student: See independent implies uncorrelated.

Correct, that is right.

Student: The reverse would mean (Refer Time: 26:47).

(Refer Slide Time: 26:49)



So, the if they are uncorrelated, can you say that they are statistically independent.

Student: (Refer Time: 26:51) yeah.

So, that does not imply that they are statistically independent, because this is what was shown for this mass clear.

Student: So, the;

Question yeah.

Student: (Refer Time: 27:10) that I want to (Refer Time: 27:11).

Huh.

Student: there is slight confusion in the steps. So, last next slide.

Here.

Student: So, essentially when expectation with axis low.

Hum.

Student: Then we can say covariance of XY is equal to expectation of XY.

Correct.

Student: I mean you just saying the LHS and the RHS.

Huh.

Student: But many kind of changes, because of that

Huh ok.

Student: And the next what you computed is actually (Refer Time: 27:38)

So, basically covariance rectifies E XY minus EX EY. So, I calculated EX is 0 here. So, therefore, I do not have to calculate I do not have to worry about EY.

Student: hm.

So, therefore, covariance of XY is basically XY.

Student: So, yeah.

So, you want to interchange the steps slightly.

Student: So, if I say covariance of XY equal to E of XY.

Huh.

Student: It mean something else, I mean the meaning is different and the next step also you computed covariance of XY.

Huh again that shows the mean is essentially 0, covariance of XY is expected value of XY that can happen, if one of the means of the random variables is zero right so.

Student: (Refer Time: 28:17) of the next step also.

Huh.

Student: a covariance of XY the computation will actually.

Ok.

Student: So, that is why I wanted to change the previous step of let it to write (Refer Time: 28:29).

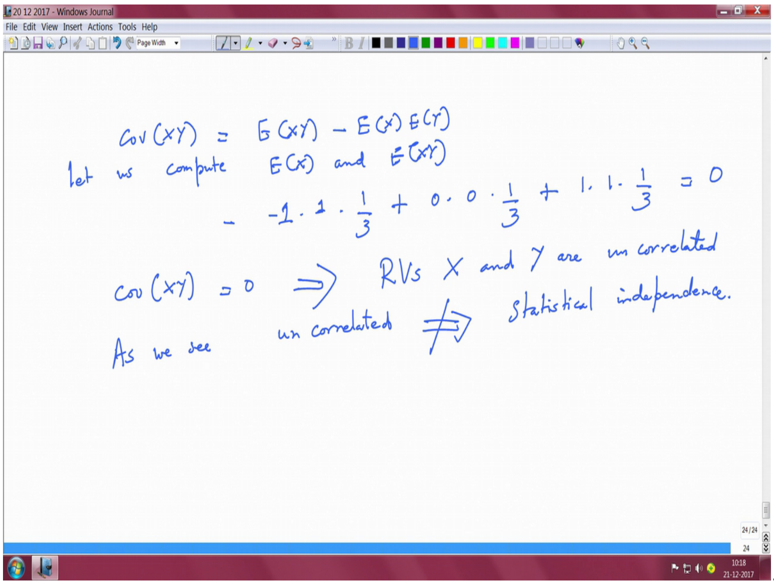
Right, right, let me just do this. Let me just say covariance of XY is EXY minus. Maybe I think I should just write this directly here. I should start with the covariance here.

Student: no I think first one is you have to compute XY (Refer Time: 28:52).

Now, I will put it this way.

Student: then you will compute a XY (Refer Time: 28:54).

(Refer Slide Time: 28:55)



Then I compute EX, expectation of X and expectation of XY, right.

Student: Hum.

Then ok.

Student: (Refer Time: 29:16)

We will follow like that fine. So, probably let us just retake just a little bit of it , again for this. So, it gives you an idea what you will go through the example as you solve the problems. Shall we retake a little bit of this again. Hello.

Yeah, yes sir.

Yeah. So, I think you do not have to get, the other people are getting bored.

We do not know.

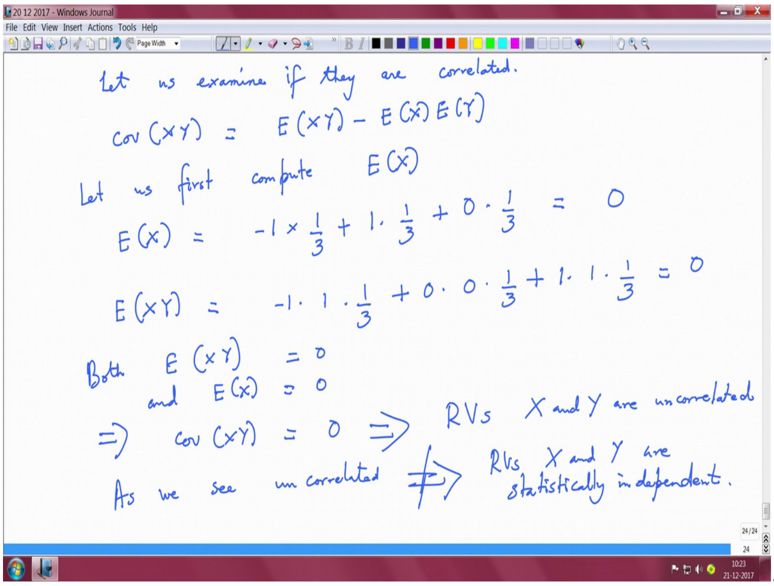
So, I think I finished here. So, any doubts at this point clear, before I am could gets another doubt again on the conclusion.

Student: (Refer Time: 30:07) Pr and some places its P some places its Pr probability (Refer Time: 30:15)

Yeah leave that. Ok, we will get started.

You can continue sir.

(Refer Slide Time: 30:42)



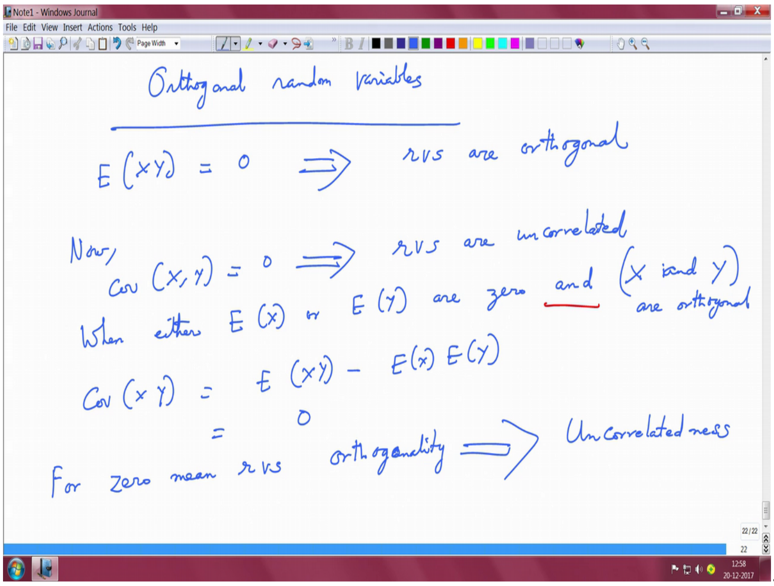
Now we showed that X and Y are not statistically independent. Let us examine , if they are correlated. Now consider the covariance of XY. This can be written as expectation of XY minus expectation of X times expectation of Y. So, we need to compute these quantities. So, let us first compute expectation of X right. So, expectation of X is basically the statistical average for the random variable X. So, x takes values minus 1 0 and 1.

So, minus 1 with the probability 1 by 3 and how can you get this probability by basically marginalizing um. Then it takes 1 with a probability one third, and it takes 0 with the probability one third and this evaluates to 0. So, therefore, the second term in this covariance quantity is 0, because one of the terms is evaluating to 0. Let us compute expectation of XY. So, this is basically ah, this has to be computed over the joint probability mass function of X and Y. So, now, reading off from this table. So, there are only three entries that are nonzero.

So, let us take those quantities. So, X can take minus 1 and Y can take plus 1 and the joint probability mass function is one third. Then X can take 0 Y can take 0. The mass is 1 by 3 and X can take 1 and Y can take 1 and the joint probability mass function is one third. So, this evaluates to zero. So, both of these quantities that is expectation of XY is a 0 and EX happens to be 0. We do not have to evaluate EY, because one of them is heading to 0. Therefore, covariance of XY equals 0. This implies the random variables X and Y are uncorrelated right.

Now, they are uncorrelated this does not mean, so as we see they are uncorrelated that does not imply that random variables X and Y are statistically independent ok. So, this sort of clarifies the ideas behind correlatedness and statistical independence orthogonal random variables.

(Refer Slide Time: 34:21)



If expected value of XY equals 0, this implies random variables are orthogonal. Remember from the basics in linear algebra right, I mean this is basically some measure of an inner product right. If you look at the inner product of two vectors, then if it evaluates inner product is zero then we say the vectors are orthogonal.

In the statistical sense if you look at the expected value of XY X times Y, expected value of X times Y, and that if it evaluates to 0, then we say that the random variables are orthogonal. Now covariance of XY equals 0 implies that the random variables are uncorrelated. Now when either expectation of X or expectation of Y are 0 and X and Y are orthogonal, fixing by orthogonal than expected value of XY is 0. Then covariance of XY which is given by expected value of XY minus expected value of X times expected value of y this evaluates to 0 ; that is if one of X and Y have 0 mean, and expected value of XY, the correlation between X and Y if that evaluates to 0, then basically the covariance is 0.

So, therefore, for 0 mean random variables, orthogonality implies uncorrelatedness. So, if you consider a zero mean. If one of the random variables E 0 mean right, then if the two random variables are orthogonal that implies they are uncorrelated. So, these are some basics of a probability. Of course, this is a entire semester course ah, I am just sort of squeezing this about 1 hour plus lecture. So, that you can get some basic sort of refreshed ah. I advise you to look into the texts that deal with probability theory extensively, and to go through these details carefully. So, that you are sort of familiar if you are not familiar, but these concepts basically help you sail through the rest of the course when whenever we deal with such statistical quantities ok. So, we complete this.