

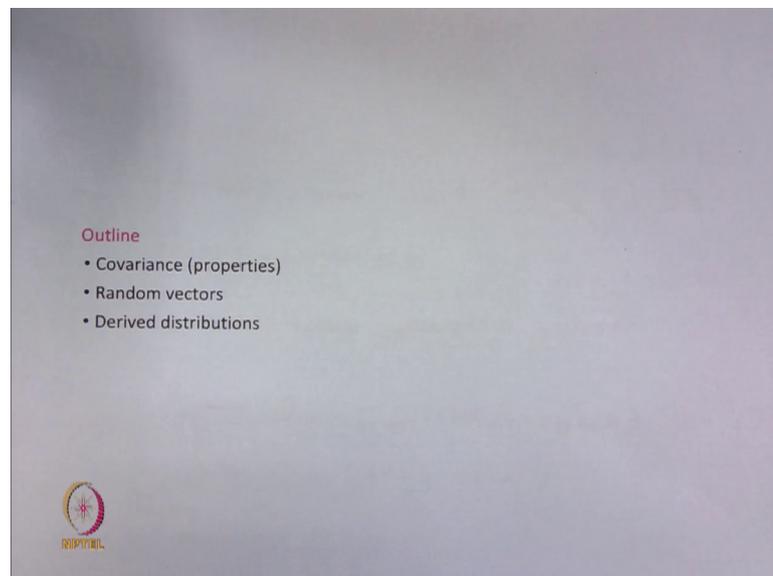
Principles of Digital Communication
Prof. Abhishek Dixit
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture -1
Random Variables & Random Processes: Random Vectors

Good morning, welcome to new lecture on Random Processes. So, in the last lecture what we discussed is; we looked at the multiple random variables. And we cover terms of density functions will looked at joint probability density functions we define what is conditional probability density function. We understood what are statistically independent random variables.

And towards the end of the lecture we also define what is correlation and covariance. Today is going to be the last lecture in hopefully last lecture on random variables. And this is the outline of today's discussion; we will complete the discussion on covariance and look at the properties of covariance.

(Refer Slide Time: 00:59)



We would then cover another important concept that is random vectors. And finally, we will look at derived distributions. So, this is what we have for you today. So, if you recall. So, I am starting from covariance where we left the in the last lecture.

(Refer Slide Time: 01:16)

The image shows a handwritten derivation on a grid background. It starts with the definition of covariance: $Cov(X, Y) = E[XY] - E[X]E[Y]$. Below this, it states that X and Y are statistically independent, which implies $E[XY] = E[X]E[Y]$. Substituting this into the covariance formula, it shows $Cov(X, Y) = E[X]E[Y] - E[X]E[Y] = 0$. The final conclusion is that $Cov(X, Y) = 0$ and that X and Y are uncorrelated. There is a small logo in the bottom left corner of the grid.

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

X & Y are statistically ind.

$$E[XY] = E[X]E[Y]$$
$$Cov(X, Y) = E[XY] - E[X]E[Y]$$
$$= E[X]E[Y] - E[X]E[Y]$$
$$= 0$$

$Cov(X, Y) = 0$ Then X, Y are uncorrelated.

So, if you recall we defined covariance of two random variables X and Y as expected value of XY minus expected value of X into expected value of random variable Y . Now we discussed and derived this property in the last lecture we also said that if X and Y are statistically independent, then expected value of XY is nothing, but it is the product of expected value of X into expected value of Y ok.

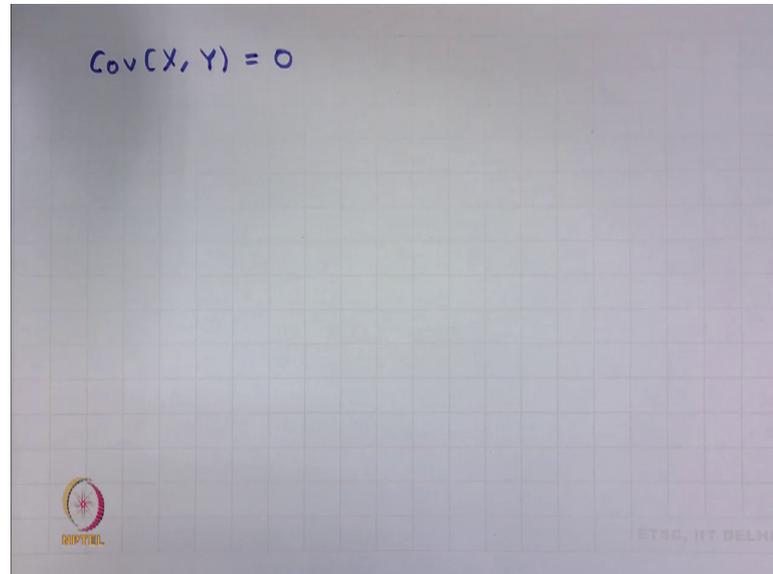
So, let us see what happens if the XY are statistically independent random variance what happens to the covariance of those two random variables and we will see something very interesting happening there. So, if you assume the two random variables to be statistically independent and compute it is covariance. Covariance by definition is this and then we can simply substitute the expected value of XY if the two random variables are statistically independent.

So, substituting this in this expression what we get a something very interesting that the covariance turned out to be 0; that means, if the two random variables are statistical independence their covariance is 0. And if the covariance of the two random variables is 0, then we call them as uncorrelated random variables.

If covariance is 0 then X and Y are also referred to as uncorrelated random variables ok. So, I repeat because this is a very important concepts to understand that if I have two statistically independent random variables then their covariance always 0. And if the

covariance is 0 then you also call those random variables as uncorrelated random variables ok.

(Refer Slide Time: 04:01)


$$\text{Cov}(X, Y) = 0$$

Now, let me ask if the covariance is 0 if covariance of two random variables is 0. Of course they are uncorrelated, but is it guaranteed that they are statistically independent right. So, we have already shown that if they are statistically independent if two random variables is statistical independence that they will be always uncorrelated. But if they are uncorrelated is it guaranteed that they are statistically independent. So, to understand that let me take this example; where I have two statistically independent random variables X and Z.

(Refer Slide Time: 04:37)

The slide contains two probability mass function plots. The first plot for X shows a discrete distribution with values 0 and 1, each with a probability of 1/2. The second plot for Z shows a discrete distribution with values -1 and 1, each with a probability of 1/2, and a mean value of 0 indicated by an arrow. Below the plots, the text reads: "If X and Z are statistically independent, and $U = XZ$, find $Cov(U, X)$." The handwritten derivation follows: $Cov(U, X) = E[UX] - E[U]E[X] = E[UX] = 0$; $E[U] = E[X]E[Z] = 0$; $E[UX] = E[XZX] = E[X^2Z] = E[X^2]E[Z] = 0$. A small logo is visible in the bottom left corner of the slide.

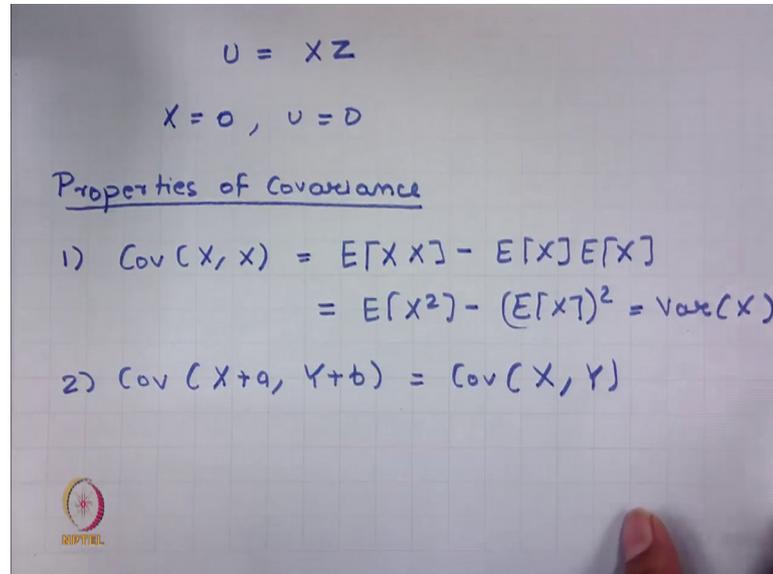
So, the probability mass function of X is given to us the probability mass function of Z is also given to us. And what is also given to us is X and Z are statistically independent. And I define another random variable U which is nothing, but the product of X and Z. Let us find first what is the covariance of U with X. So let us start you have to find covariance of U and X. And by definition it is nothing, but is the expected value of U X minus expected value of U into expected value of X right. So, let us compute first what is expected value of U?

U by definition is X into Z so this is expected value of X into expected value of Z because X and Z are statistical independence and U is nothing, but X times Z. Now if you look at the expected value of Z. So, if you look at Z, where is the centre of gravity? Centre of gravity will be at 0 ok. So, expected value of Z is 0. So, what I will find a with is expected value of U is also 0 ok. So, the covariance of U and X reduces to expected value of U and X because expected value of U is 0. Now what is expected value of U and X.

So, substituting for U which is X Z what I get is expected value of X square Z. And because X Z are statistically independent I can write this as this. And I already know that the expected value of Z is 0 so expected value of U X is also 0 right. Hence the covariance of U and X is 0. So, we can say that U and X are uncorrelated because their covariance is 0. Now our U and X statistically independent that is explore. So, we have

seen that U and X are uncorrelated, but the question that we have to think now is whether they are statistically independent.

(Refer Slide Time: 07:23)



$$U = XZ$$
$$X = 0, U = 0$$
Properties of Covariance
1)
$$\text{Cov}(X, X) = E[XX] - E[X]E[X]$$
$$= E[X^2] - (E[X])^2 = \text{Var}(X)$$

2)
$$\text{Cov}(X+a, Y+b) = \text{Cov}(X, Y)$$

So, I have been given that you is X and Z ok. So, how do I think about whether U and X are statistically independent let us go back to the definition. What I said about to statistical independent random variables, we said that; if I give some information about a random variable X . And if you do not learn anything about Y then the two random variables involved are statistically independent.

So, in this case suppose I give you information that X has taken a value 0. So, look at the PMF of X , X can take two values 0 and 1. So, if I say that X has taken a value 0 you know for sure that you must also be 0 right. So, these two random variables are not statistically independent. If I give you some information about X you know for sure what the value for U would be.

So, this U and X are not a statistical independence; however, they are uncorrelated because covariance turned out to be 0. So, the big statement that we have making is if the two random variables are uncorrelated they need not be statistical independence, but other way round is two. That means, if two random variables are statistically independent they are guaranteed to be uncorrelated ok.

So, let us now continue looking at the properties of covariance's. So, first important property that we have to understand is let me ask what is covariance of X with X ok. So, let us use our definition. So, covariance of X with X is nothing, but expected value of X into X minus expected value of X into expected value of X so this is by definition. So, if you look at this, this is nothing, but expected value of X square minus expected value of X whole square and this is nothing, but it is the variance of X.

Yes, so we learn that covariance of X with X is nothing, but it is the variance of X. Let us see what happens if I take covariance of X plus a and Y plus b where a and b are some constant. So, as we have seen in the case of the variance what is variance let us revise variance. In variance we were interested in square of the distance from the mean. So, we have certain numerical values we calculate the distance of those numerical values from the mean we take the square of the distance and then we found the expected value so that was the variance of X.

And we have said that the variance of X does not depend upon the constant. If you add a constant to a random variable its variance remains unchanged. And covariance is also like variance is we are also interested in finding the expected value of square distance from the mean hence the covariance is also unaffected by the DC constants. So, we can say that covariance of X plus a and Y plus b is nothing, but it is simply the covariance of X and Y. That means, the covariance remains unchanged with the addition of some DC constants. Let us see the third property of covariance.

(Refer Slide Time: 11:07)

$$\begin{aligned} &3) \text{ Covariance is a bilinear func} \\ &\text{Cov}(a_1 X_1 + a_2 X_2, a_3 X_3 + a_4 X_4) \\ &= a_3 a_1 \text{Cov}(X_1, X_3) + a_2 a_4 \text{Cov}(X_2, X_4) \\ &\quad + a_1 a_4 \text{Cov}(X_1, X_4) + a_2 a_3 \text{Cov}(X_2, X_3) \end{aligned}$$

We are making a statement that covariance is a bilinear function. So, what I mean with this is if you take covariance of a 1×1 plus a 2×2 with a 3×3 plus a 4×4 and this because it is a bilinear function you can write this as a 1 times covariance of $X_1 \times X_3$. So, I am taking this with this. So, I will get a $1 a_3$ plus a $2 a_4$ covariance of X_2 with X_4 .

Then a $1 a_4$ covariance of $X_1 \times X_4$, what is left? X_2 with X_3 . So, we have a $2 a_3$ covariance of X_2 with X_3 right. So, covariance is a bilinear function. So, you can use this property sometimes it will be very useful to compute the covariance ok. Let us make use of this property and see why is the so interesting.

(Refer Slide Time: 12:42)

$$\begin{aligned}
 & \text{Var}(a_1 X_1 + a_2 X_2 + \dots + a_m X_m) \\
 &= \text{Cov}(a_1 X_1 + a_2 X_2 + \dots + a_m X_m, \\
 & \quad a_1 X_1 + a_2 X_2 + \dots + a_m X_m) \\
 &= a_1^2 \text{Cov}(X_1, X_1) + a_2^2 \text{Cov}(X_2, X_2) + \dots \\
 & \quad a_m^2 \text{Cov}(X_m, X_m) + a_1 a_2 \text{Cov}(X_1, X_2) \\
 & \quad + a_1 a_3 \text{Cov}(X_1, X_3) + \dots \\
 &= \sum_{i=1}^m a_i^2 \text{Var}(X_i) + \sum_{\substack{i, j=1 \\ i \neq j}}^m a_i a_j \text{Cov}(X_i, X_j)
 \end{aligned}$$

So, let us do let us compute variance of a 1 X 1 plus a 2 X 2 up to a m X m and see what it is? So, by definition this is nothing, but it is the covariance of a 1 X 1 plus a 2 X 2 up to a m X m with itself right. it should be X 1 a 2 X 2 up to a m X m. And using the fact that covariance is a bilinear function I can simplify this as a 1 square into covariance of X 1 with X 1 plus a 2 square into covariance of X 2 with X 2 ok.

You can go on like this a m square into covariance of X m with X m. And then there will be some cross terms for example, a 1 a 2 into covariance of X 1 with X 2. Then we will have a 1 a 3 covariance of X 1 with X 3 and so on so forth right. So, you can write this you can write it compactly this can all be reduced to summation a i square were i goes from 1 to m. And covariance of X 1 with X 1 is nothing, but it is a variance. So, we can write variance of X I ok.

So, we have collected all terms which have the same random variables. So, we can write all such terms compactly in this way then I have some cross terms. So, a i with a j covariance of X i X j where i and j goes from 1 to m, but i is not same as j. Because i same as j has been already collected here. So, I can compact all this terms by using this expression. So, this helps us in computing some interesting quantities for example, let us say if I am given that X 1 X 2 and X m are uncorrelated right.

(Refer Slide Time: 15:36)

$$= \sum_{i=1}^m a_i^2 \text{Var}(X_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^m a_i a_j \text{Cov}(X_i, X_j)$$
$$= \sum_{i=1}^m a_i^2 \text{Var}(X_i)$$

So, I have said the variance of this let me write it again is nothing, but is summation a_i^2 variance of X_i , i going from 1 to m plus $a_i a_j$ covariance of X_i with X_j ; i and j going from 1 to m and i is not same as j . Now if I have been given that these are uncorrelated. So, random variables X_1, X_2 and X_m are uncorrelated if they are uncorrelated their covariance this term will go 0.

So, then you would end up with is just this term ok. So, this is how you can compute the variance of uncorrelated random variables. Of course, if they are statistically independent then also they are uncorrelated. So, you can also use the same expression if the random variables that are given us are statistically independent ok.

(Refer Slide Time: 16:58)

The image shows a handwritten derivation of the correlation coefficient formula on a grid background. At the top, the Greek letter rho (ρ) is labeled as the 'correlation-coefficient'. Below this, the formula for rho is given as the covariance of X and Y divided by the product of the square roots of their variances:
$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}}$$
 To the right of this formula, the text 'Cov(X, Y) = ' is partially visible. Below the main formula, the square roots of the variances are equated to standard deviations:
$$\sqrt{\text{Var}(X)} = \sigma_x \quad \sqrt{\text{Var}(Y)} = \sigma_y$$
 Finally, the correlation coefficient formula is simplified to:
$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}$$
 In the bottom left corner of the grid, there is a small circular logo with the word 'RUPREX' written below it.

Let us now look at another important concept and that is about normalised variance and we call this as correlation coefficient. Correlation coefficient is first let me define this correlation coefficient it is the covariance of X and Y. So, I am defining the correlation coefficient for two random variables X and Y. So, you have to first take the covariance of X and Y and then you have to divide this with square root of variance of X and variance of Y square root of variance of Y.

Now something interesting that you would have noticed concerning the name. So, here we are taking the covariance we are not taking the correlation and still we are saying this as correlation coefficient. Last time also when we said two random variables are uncorrelated we said that their covariance is 0 right. We said two random variables are uncorrelated if their covariance is 0 not their correlation here also we are defining correlation coefficients.

So, it seems we should be in terms of correlation, but it does not we are defining the correlation coefficient in terms of covariance. So, there is some inconsistency in the use of names of the quantities. So, let us see why is this, so interesting quantity ok. So, let me to simplify let me assume that variance of X is sigma X square. So, square root of variance of X is sigma X and square root of variance of Y is sigma Y ok. So, I am defining rho as covariance of X and Y divided by sigma X into sigma Y ok.

(Refer Slide Time: 19:16)

$$\rho = \frac{E[(X - m_x)(Y - m_y)]}{\sigma_x \sigma_y}$$

$m_x = E[X]$
 $m_y = E[Y]$

$$A = \frac{X - m_x}{\sigma_x} \quad B = \frac{Y - m_y}{\sigma_y}$$
$$\rho = E[AB]$$
$$E[(A \pm B)^2] = E[A^2] + E[B^2] \pm 2E[AB]$$


IIT DELHI

Now, let us assume that this rho is let me let me expand this covariance of X and Y as expected value of X minus m_x into Y minus m_y . So, this is by definition this is the covariance of random variable X and Y and then we have in the denominator sigma X and sigma y. Remember I have already introduced this notation of m_x m_y , m_x is nothing, but it is the expected value of random variable X and m_y is the expected value of random variable y.

So, this is the covariance of random variable X and Y and we have divided it with their standard deviation. So, square root of variance is also known as the standard deviation of the random variables. So, let me define A as X minus m_x by sigma x. And let us define B as Y minus m_y by sigma y ok. So, because X is a random variable A is also a random variable and Y is a random variable B is also a random variable.

So, I can write rho E correlation coefficient is nothing, but it is the expected value of A into B. Now you know in general expected value of A plus or minus B whole square is nothing, but it is expected value of a square plus expected value of b square plus or minus 2 times expected value of A B ok. Now let us see what is this expected value of A square.

(Refer Slide Time: 21:23)

$$E[A^2] = E\left[\frac{(X - m_x)^2}{\sigma_x^2}\right] = 1$$
$$E[B^2] = 1$$
$$E[(A \pm B)^2] = 1 + 1 \pm 2\rho$$
$$= 2 \pm 2\rho$$
$$2 \pm 2\rho \geq 0 \quad 1 \pm \rho \geq 0$$
$$1 + \rho \geq 0 \quad 1 - \rho \geq 0$$
$$\rho \geq -1 \quad \rho \leq 1$$

$-1 \leq \rho \leq 1$

So, if I compute the expected value of a square this turns out to be X minus m_x whole square by σ_x square and by definition expected value of X minus m_x whole square is nothing, but it is σ_x square so this quantity is 1. Similarly expected value of B square is 1 ok. So, I can write expected value of A plus or minus B whole square expected value of a square is 1 plus expected value of B square is also 1 plus and minus 2 times expected value of A and B two times expected value of A and B .

And what is expected value of A and B it is defined to be ρ alright. So, we get this as 2 plus minus 2 ρ ok. Now expected value of A plus minus B whole square can never be less than 0 because this quantity is always positive. So, I can say that 2 plus minus 2 ρ is greater than or equals to 0 because this quantity is always non negative. So, 2 plus minus 2 ρ should also be non negative. So, from this I get 1 plus minus ρ is also always non negative. And from this I can say 1 plus ρ is greater than or equals to 0 and 1 minus ρ should also be greater than or equals to 0.

From this I can get ρ is greater than equals to minus 1 from this I get ρ should be less than 1. So, using these two I can say that ρ always lie or correlation coefficient always lie between 1 and minus 1 alright. So, this is this is providing as a some kind of normalised this is some kind of a normalised matrix because it is values only lie between minus 1 and 1.

(Refer Slide Time: 24:08)

$$\begin{aligned} \rho = 0, & \quad X \text{ \& } Y \text{ are uncorrelated} \\ \rho = 1, & \quad Y = aX + b \quad a > 0 \\ \rho = -1 & \quad Y = -aX + b \end{aligned}$$

So, if rho is 0 what can we say about this if rho is 0 then X and Y are uncorrelated. Rho 0 means their covariance of X and Y is zero; that means, X and Y are uncorrelated. Rho 1 means what would be the meaning of rho equals to 1; that means, X and Y are highly correlated and it would be something like this. So, Y will be a linear function of X where a is positive.

So, Y is linear function of X the correlation coefficient between Y and X would be 1. If rho equals to minus 1 would corresponds to the situation when Y is again a linear function of X, but with a negative sign. So, the correlation coefficient between Y and X would be minus 1 in this case. So, we have finished with everything that is important about covariance.

(Refer Slide Time: 25:19)

Random Vectors

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{array}{l} \text{Random Variables} \\ \text{Column Vectors} \end{array}$$
$$X = [x_1 \ x_2 \ \dots \ x_m]^T$$

NPTEL IIT DELHI

Let us now go to the next concept and that is about random vectors. what is a random vector? A random vector is a vector whose elements are random variables. So, for example, this is a vector and elements of this random vector are random variables. Okay so just to introduce a notation here normally if we are not given whether this is a column vector or a row vector the vectors are assumed to be column vectors.

That is normally the notation that is used in most books in digital communication vectors are somehow always assume to be column vectors until and unless you are focusing on special topics in digital communication like error control coding etcetera where the vectors are normally assumed to be row vector. But normally in other parts of digital communication the vectors are assume to be column vectors.

So, this is the same thing that we would follow. So, this is a column vector X is a column vector and normally because it consumes lot of space if you represent a column vector like this. So, vector is also a column vector is also written like this. So, you write it as a row vector and then it take a transpose.

So, that it is a column vector this is more space efficient and hence a column vector is mostly also represented like this we will use this notation because it is easy to see and this is a column vector. So, now we have a random vector instead of a random variable and things are how do we do certain operations on these random vectors.

(Refer Slide Time: 27:20)

$$E[X] = E \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_m] \end{bmatrix}$$
$$X - E[X] = \begin{bmatrix} X_1 - E[X_1] \\ X_2 - E[X_2] \\ \vdots \\ X_m - E[X_m] \end{bmatrix}$$


KTSC, IIT DELHI

So, let me define expectation operation on a random vector X . So, X is now a random vector and not a variable. So, we are not using a different notation for random vectors we are using the same notation, but I will like to stress always that this X is now a random vector. So, if I want to take the expectation of X it is very simple I can take the expectation element wise. So, I can take the expectation element wise. So taking expectation is easy no let us say that I am interested in random vector X minus expectation of X what about this random vector.

How do we think about this? So, it is simple that you subtract this also element wise ok. So, this is how you should understand X minus expected value of X . So, you think about X element wise and subtract it from the expected value of each element. So, this is your vector X minus expected value of X can I add two vectors. So, what is the expected value of X plus Y ?

(Refer Slide Time: 29:02)

$$E[X+Y] = \begin{bmatrix} E[X_1 + Y_1] \\ E[X_2 + Y_2] \\ \vdots \\ E[X_m + Y_m] \end{bmatrix}$$


RIIT DELHI

So, remember these two are vectors X and Y are vectors. So, we can do or take the expected value also element wise. So, I am doing in one step because I hope that you have understood this. So, this is element X_1 X_2 and X_m are elements of vector X . And Y_1 Y_2 and Y_m are the elements of vector Y .

I add these two elements X_1 plus Y_1 X_2 with Y_2 X_m plus Y_m and then I take the expectation element wise. So, this would be the expected value of two random vectors X and Y . Now, let me introduce something complicated that is the variance of X .

(Refer Slide Time: 30:05)

$$\begin{aligned} \text{Var}(X) &= E[XX^T] - E[X]E[X^T] \\ E[X] &= 0 \\ \text{Var}(X) &= E[XX^T] \end{aligned}$$
$$XX^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}_{m \times 1} \begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix}_{1 \times m}$$
$$= \begin{bmatrix} x_1 x_1 & x_1 x_2 & \dots & x_1 x_m \\ x_2 x_1 & x_2 x_2 & \dots & x_2 x_m \\ \vdots & \vdots & \ddots & \vdots \\ x_m x_1 & x_m x_2 & \dots & x_m x_m \end{bmatrix}_{m \times m}$$


RIIT DELHI

So, how do I think about variance? Variance of X we know is given by expected value of X into X transpose minus expected value of X into expected value of X transpose ok. So, we have already seen that variance of X can be given by this expression further more what we assume is this random vector X is a 0 mean random vector. That means, expected value of X is 0 and we make this assumption for two reasons. One reason is that it simplifies this expression.

And second reason is noise is a random process with 0 mean. And thus the random vector obtained would also be with 0 mean. So, if I assume that the random vector is a 0 mean random vector variance of that random vector simplifies to expected value of X into X transpose and remember this definition for variance applies if X is a column vector. And in this course we always assume the vectors to be column vectors ok.

Now let us first look into what is this quantity X into X transpose. So, if I have to find X into X transpose X is $X_1 X_2$ let us assume up to m . So, I have m elements. So, this is m vector and X transpose would be $X_1 X_2 X_m$. So, this is one way m vector and then I have to multiply terms ok. So, then I will get X_1 into $X_1 X_1 X_2 X_1 X_m X_2 X_1 X_2 X_2 X_2 X_m$ rest you can fill it yourself.

Let me write the last row $X_m X_1 X_m X_2 X_m X_m$ and this is m by m matrix ok. Now what we have to do is we have to take the expected value of this matrix and we can do this by pulling up this expectation operator inside this matrix. And we can calculate the expectation term wise ok. So, let us see what we end up with.

(Refer Slide Time: 33:23)

The image shows a handwritten derivation on a grid background. At the top, a covariance matrix is written as a square bracket containing the following elements: the first row has $E[X_1^2]$, $E[X_1 X_2]$, and $\dots E[X_1 X_m]$; the second row has a vertical ellipsis, $E[X_2^2]$, and a vertical ellipsis; the last row has a vertical ellipsis, $E[X_m X_1]$, a horizontal ellipsis, and $E[X_m^2]$. Below this, the formula for covariance is written as $\text{Cov}(X, Y) = \underline{E[XY^T] - E[X]E[Y^T]}$. In the bottom left corner, there is a small circular logo with the text 'RUPAKUL' below it. In the bottom right corner, the text 'ETDC, IIT DELHI' is visible.

So, we will get expected value of X one square expected value of X 2 square and so on. So, forth expected value of X m square and you can fill in all these values this would be expected value of X 1 X 2 this would be expected value of X 1 X m and so on so forth. Rest I believe that you can fill yourself ok. So, now, we have understood clearly I guess there is no more ambiguity what I mean? I taking a variance of a random vector X what about the covariance.

So, covariance of random vector X and Y is also defined in a very similar way it is expected value of X into Y transpose minus expected value of X into expected value of Y transpose it is very similar just you have to change X with a vector Y ok. So, I have finished random vectors they will be very useful when we talk about Gaussian processes or random processes later on.

(Refer Slide Time: 34:44)

Derived distributions

$$Z = X + Y$$

X & Y are statistically indep.
 $f_x(x)$ & $f_y(y)$ are given

pdf of Z ?

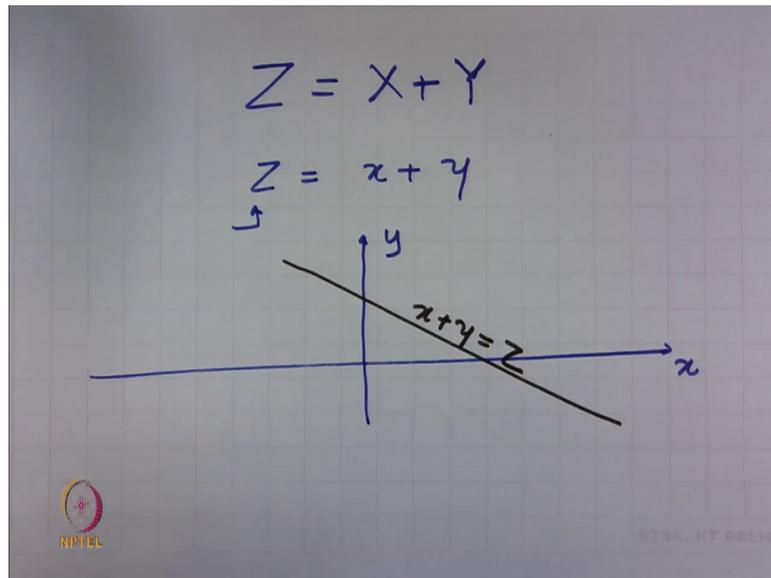
NPTEL

ETSC, IIT DELHI

So, let us now look into another theme and this theme is about derived distributions. So, what is the problem is covered under this topic derived distributions the problem is that suppose you are given a random variable Z . Suppose you have been given also that these X and Y random variables are statistically independent, you have been given the probability density function of these random variables then the problem statement is what is the probability density function of this random variable Z . So, what is the pdf of Z that is the problem? So, problem is to evaluate the probability density function of a random variable.

When you are given the probability density functions of some other random variables; if this random variable is built using other random variables. So, this is the topic of discussion under derived distributions this problem statement can be quite complicated, but here we are considering is simple case in which random variable is just a linear combination of other random variables ok. So, we will be looking into just this problem at hand. So, let us get started.

(Refer Slide Time: 36:55)



So, let us to write again the problem that we have to solve. So, we have to find out the pdf of Z in terms of pdf of X and Y . If X and Y are statistically independent random variables it is also often convenient to think about these random variables in terms of their numerical values. So, I know that numerical values of Z would be a sum of numerical value of X and numerical value of random variable Y also assuming that this Z is some constant.

What I will end up with is an equation of a line which we can draw. So, suppose I have here the numerical values of random variable X and on Y axis. I have a numerical values corresponding to random variable Y I can draw a line X plus Y equals to Z for a given value of Z . We can draw such a line.

(Refer Slide Time: 38:36)

$$\begin{array}{cc} f_Z(z) & F_Z(z) \\ \uparrow \text{pdf} & \text{cdf} \end{array}$$
$$f_Z(z) = \frac{d}{dz} F_Z(z)$$
$$F_Z(z) = P(Z \leq z)$$
$$= P(Z < z)$$

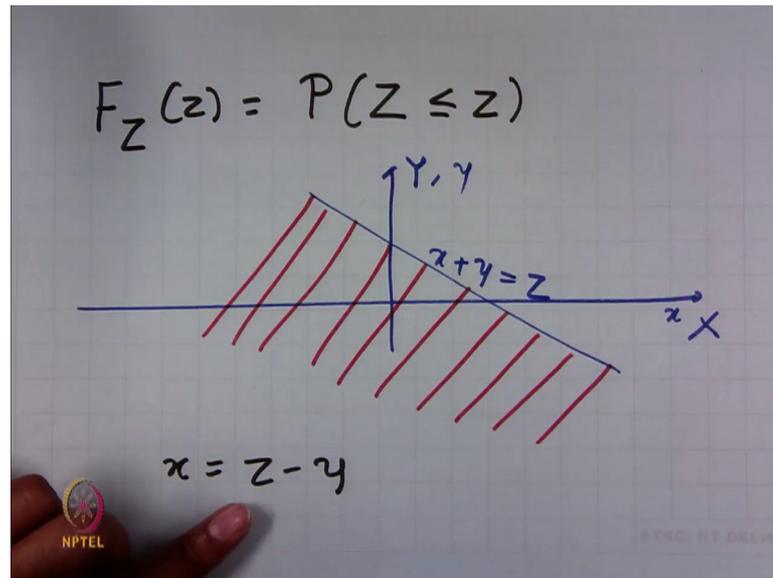
NPTEL

Now, to start thinking about the probability density functions of a random variable. It is often convenient to think it in terms of cumulative distribution function of a random variable. That means, that if I want to evaluate pdf it is often convenient to start thinking in terms of cdf. Because we know that you can easily obtain pdf from cdf by just differentiating it with respect to a variable.

So, in this case you can just differentiate the cdf of Z with respect to Z and you can obtain it is pdf and what is the cdf of a random variable. So, cdf of a random variable is by definition. The probability that a random variable takes a value less than or equals to the numerical value in Gaussian. So, for example, if I want to find the cdf of random variable Z I have the numerical value Z it can be easily obtained by looking at the probabilities with which this random variable Z takes in a value less than or equals to the numerical value Z .

And as we have already said probabilities at a single point is 0 thus this quantity is same as evaluating this quantity. Of course if the probability density function is well behaved meaning that it does not have any impulses and so on so forth. Does whether I find the probability with which random variable Z takes in a value less than or equals to Z or I just evaluate this probability it is one and the same thing ok.

(Refer Slide Time: 40:40)



So, what I am saying is, I am interested in finding the cdf of the Z which you can start thinking by thinking in terms of probabilities for which the random variable Z takes in a value less than or equals to Z . And how to think about this we can think about this in terms of random variable X and Y taking some numerical values because that is what is given to us.

Let me draw that line again. So, we had two random variables we are plotting the numerical values we said that I can draw a line X plus Y equals to Z . If I am interested in this quantity with what probability Z takes in a value less than or equal to Z . Actually I am interested in all points that lie left to this line. And I can think about this probability of the points lying left to the line in terms of X taking certain values and Y taking certain values.

For example, I can rearrange this equation instead of thinking about this I can write this equation. And I can ask this question in a different way I can say if I want to find the probability with which Z takes in a value less than Z instead of thinking about this probability directly can I think about the probabilities with which X takes in a value less than this and that is the idea. So, I can freely choose my Y and for a given Y I put the restriction that X should take in a value less than Z minus Y where Y is the value that I chose.

(Refer Slide Time: 43:04)

$$F_Z(z) = P(Z \leq z)$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_{X,Y}(x,y) dx dy$$
$$x \leq z - y$$
$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

So, let us formally state this if I am interested in finding cdf of Z which is nothing, but this I can think about this in terms of joint probability density function of X and Y . Because a point line to the left of this line will happen when Y and X takes a particular values and thus we have to think it in terms of joint pdf of X and Y . And what I am saying is I can freely choose my Y . So, Y can lie between minus infinity to plus infinity, but for a chosen value of Y my X should be less than or equals to Z minus y .

That means, X can only lie between minus infinity to Z minus Y and if you carry out this integration you can find the cumulative distribution function of random variable Z . Now to simplify this thing I can use the idea that these two random variables are statistical independent and if two random variables are statistically independent the joint pdf is nothing. But it is the product of their marginal pdfs.

(Refer Slide Time: 45:02)

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{z-y} f_X(x) dx dy \\ f_Z(z) &= \frac{d}{dz} F_Z(z) \end{aligned}$$

So, instead of thinking about this quantity I can think it in terms of this quantity and then I can substitute this thing in here. And what I get is. F_Z of Z can be evaluated and using Frobenius theorem, I can rearrange this integration in this way all right.

Now the question is to obtain the probability density function now probability density function of Z can be obtained by differentiating the cdf. So, let us now differentiate this quantity with respect to Z and let us see what we get.

(Refer Slide Time: 46:15)

$$\begin{aligned} f_Z(z) &= \frac{d}{dz} \left[\int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{z-y} f_X(x) dx dy \right] \\ \frac{d}{dz} \int_{-\infty}^{z-y} f_X(x) dx &= \frac{d}{dz} \left[F_X(z-y) - F_X(-\infty) \right] \\ &= \frac{d}{dz} F_X(z-y) = f_X(z-y) \end{aligned}$$

So, pdf of z can be obtained by differentiating integration that we had. What you can appreciate is only this integration is a function of z . All these quantities are independent of Z . So, while differentiating I need to just consider this part. So, let us think about this separately. So, let us see what we get if I differentiate this quantity. So, we get integration of pdf of x is the cdf of x .

So what we will get is this minus F_x of minus infinity. Now this quantity is 0. Why this is 0. Because it is giving me the probability with which x is less than minus infinity and that probability should be 0. So, what I wind up with is simply this term and this quantity is nothing but it is the pdf of x evaluated at z minus y so it is simple.

(Refer Slide Time: 48:16)

$$f_z(z) = \int_{-\infty}^{\infty} f_y(y) f_x(z-y) dy$$

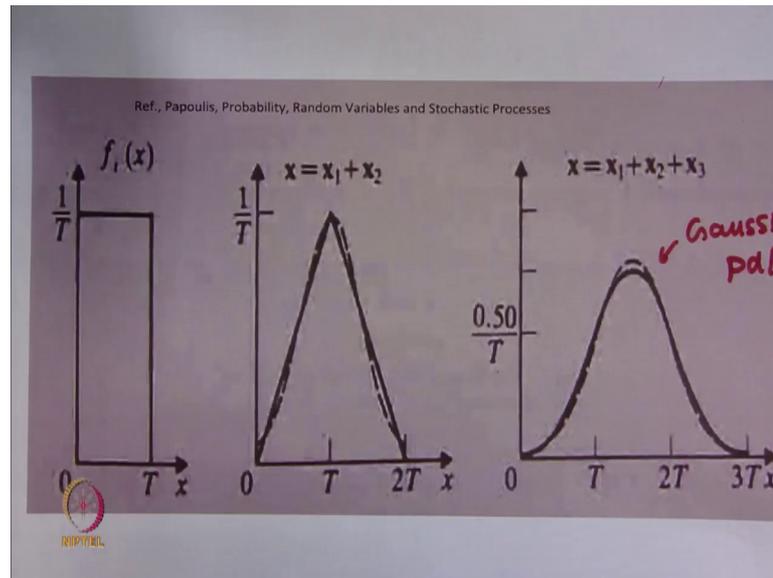
$$f_z(z) = f_y(y) * f_x(x)$$

$$Z = X + Y$$

So, what I get? When I differentiate this integration with respect to Z is simply and what is this if you look at this carefully this is nothing but it is the convolution of the probability density functions of X and Y ok. So, what we have derived is if Z is sum of two random variables X and Y . If these two random variables are statistically independent then I can obtain the probability density function of Z by carrying out.

The convolution of probability density function of Y and probability density function of X . And this is a very important result because it will help you in finding out the probability density function of a random variable in terms of probability density function of other random variables. So, let me show you a picture.

(Refer Slide Time: 49:39)



So, let me assume that I have a random variable which is a uniformly distributed random variable. And I have summed up two such random variables. What I end up with is its density function like this. So, if you take a convolution of a rectangle, you get a triangle. This is from the basic courses in signals and systems. So, I have a uniformly distributed probability density function uniform within a range.

So, you can assume it to be kind of a rectangle density function. And if you take the sum of two such random variables independent random variables, the density of X will be the convolution of the density of X_1 and X_2 . And let us assume that they are identical random variables. So, the resultant density will be a triangle like a triangle. And now if I add one more random variable, the density of that random variable will be the convolution of a rectangle with a triangle, and what you would get is such a shape.

If you look at it closely, this begins to approach a Gaussian pdf, and this is a very important insight. If you take random variables independent of each other, if you take a large sum of them, what you would get is the resultant density function would begin to approach Gaussian. This is not just particular to this example but this holds universally such that there is a theorem which is known as the central limit theorem which says that; if you take a sum of a large number of independent random variables.

The resultant probability density function begins to behave like a Gaussian probability density function or it becomes Gaussian probability density function. So, this is a very

important insight this is a central limit theorem based on which we model noise. Gaussian because noise is nothing, but it is a collection of many independent random phenomenon's.

(Refer Slide Time: 51:50)

$$\begin{aligned}
 Z &= X + Y \\
 f_Z(z) &= \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} * \\
 &\quad \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}} \\
 f_Z(z) &= \frac{1}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} e^{-\frac{(z - (m_x + m_y))^2}{2(\sigma_x^2 + \sigma_y^2)}}
 \end{aligned}$$

So, because we are talking so much about Gaussian let us let us do that previous example assuming that X and Y are Gaussian. So, what is the resultant probability density function if X and Y in Gaussian. So, I know that probability density function of Z would be nothing but it is the convolution of probability density function of X which if it is Gaussian would be like this. So, assuming the variance is sigma x square and mean is m x convolution with the probability density function of y ok.

So, I have taken a convolution of the probability density function of x with probability density function of y. And you can prove this I leave it to you to prove that, this is nothing, but ok. So, what I get is the resultant probability density function of z which is obtained by taking the convolution of the probability density function of X with the probability density function of Y this is the resulted probability density function. So, what you notice thus very important result again.

So, we are covering lot of important stuff today that the resultant probability density function is also Gaussian. So, I have taken a Gaussian random variable, I have taken another Gaussian random variable both where statistically independent of each other I

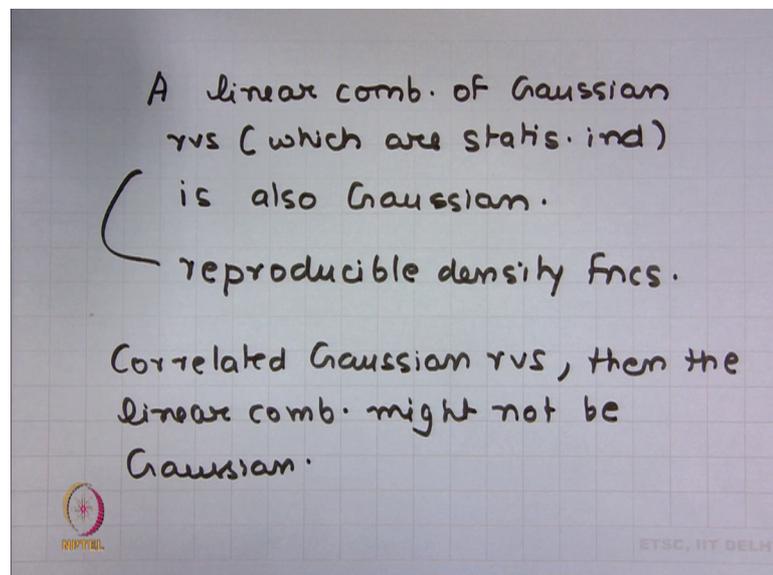
sum them up. And what I end up with is the resultant random variable also as a Gaussian random variable of course with a modified mean so mean has changed to $m_x + m_y$.

So, mean of these two random variables have added and the variance is $\sigma_x^2 + \sigma_y^2$ where the variance of x is this and variance of y is this. So, the mean have changed the means have added and as expected the variances have also added. And the random variable remained Gaussian. And this is also a very central result in digital communication course.

So, let me make a important conclusion about what we have discussed so far is that a linear combination of Gaussian random variables. These random variables note should be statistically independent a linear combination of Gaussian random variables there are random variables which are statistically independent is also Gaussian ok.

So, this is this is very interesting that if you take a statistically independent Gaussian random available you sum them up you make a linear combination out of them and the resultant random variable is a still a Gaussian. The density functions which shows this behaviour are known as reproducible density functions right.

(Refer Slide Time: 55:37)



So, as you can imagine that they are not many density functions which reproducible. So, it is a very special class of density functions which shows this property that you make a linear combination of density functions. And the resultant random variable also exhibit

the same density function. And Gaussian density function behaves in that way. So, important question here is that Gaussian we are making a linear combination of Gaussian random variables, but they are all statistically independent.

If we have correlated Gaussian random variables; that means, they are not statistically independent they are correlated. Then, the linear combination might not be Gaussian ok. This is important it might not be Gaussian ok. So, do not think that if you sum up the Gaussian random variables whatever they are correlated uncorrelated the resultant is all always Gaussian it is not true. If they are statistically independent, yes the resultant is Gaussian if they are correlated then we do not know.

So, we have a special class of Gaussian random variables which are correlated and if still they linear combination gives a Gaussian function and we will learn about such Gaussian random variables in the coming lectures, but in general this is not true. So, now, we have seen two important properties of Gaussian random variables that is a linear combination a Gaussian random variable is also a Gaussian random variable.

And before if you remember we have seen that a linear function of a Gaussian random variable is also a Gaussian random variable. So, these Gaussian densities and Gaussian functions are thus so important and whenever we are in confusion how to model particular phenomena you better choose this to be a Gaussian function ok. So, we have in this lecture have seen several interesting properties of Gaussian functions. And we will see more interesting properties in next lecture.

Thank you.