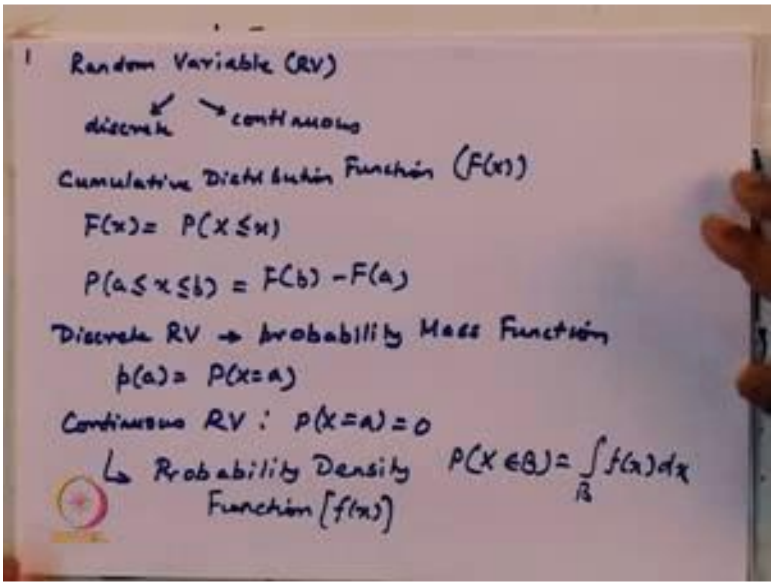
Introduction to Biostatistics
Prof. Shamik Sen
Department of Bioscience and Bioengineering
Indian Institute of Technology, Bombay

Lecture - 20
Expectation, Variance and Covariance

Hello and welcome to today's lecture. So, in the last two lectures we had briefly discussed about random variables and we had introduced upon the concept of expectation.

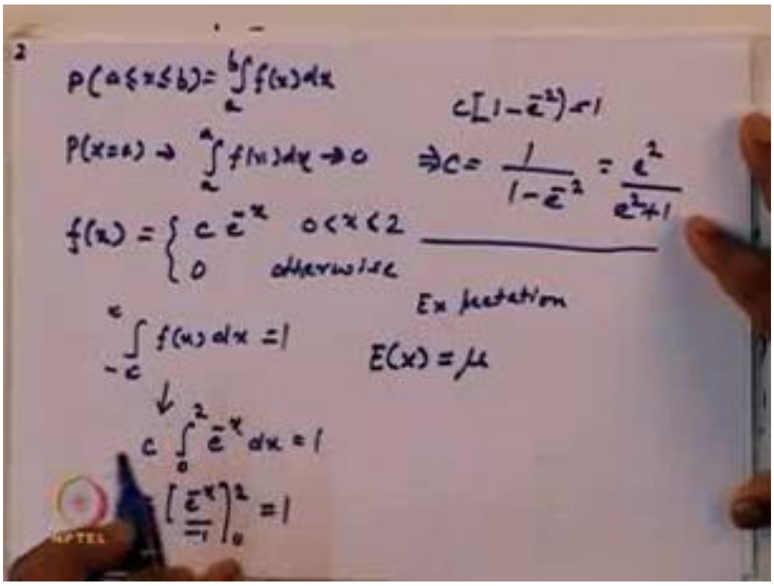
(Refer Slide Time: 00:38)



So, I will begin by briefly recapping whatever we are discussed with respect to random variables. So, a random variable can be discrete or continuous for random variable, you can determine something or define something called a cumulative distribution function. it is typically written as F of x. So, F of x is probability of x less or equal to a value x and. So, using this we can calculate probability in inner regime, let say for example, this is nothing, but F of b minus F of a, for a discrete random variable R V.

So, I will put R V in short you can define a probability mass function which is nothing, but P of a is given by probability of x equal to a, but for a continuous random variable probability of x equal to a is always 0 why is this because for a continuous random variable, you have a concept of a probability density function written as F of x. So, you have probability of x in an interval b is defined as integral of F X D X over that interval.

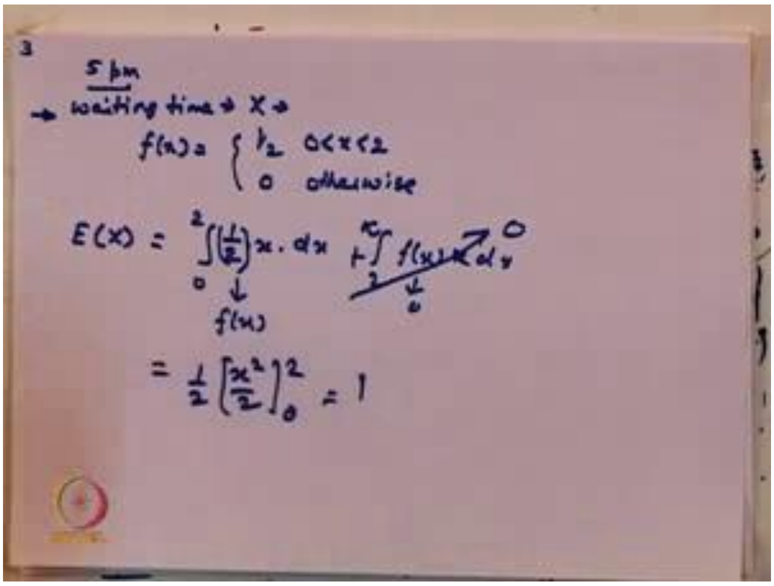
(Refer Slide Time: 02:49)



So, this would mean. So, this would mean that probability of a is integral of a to b f x d x. So, priority of x equal to a what amount to a to a f x d x, and this is nothing, but, because you have the same interval. So, let us take a pretty you know simple example. So, let us define f x as c to the power minus x between 0 2 and equal to 0 otherwise. How do we calculate f of x or how do we calculate the constant. So, you we use the identity that between minus infinity, and plus infinity this integral was return me a value of 1. So, I can use this expression. So, this amount to g c is equal to 1 I can take it out. So, this comes to be. So, this I can simplify as c 0 is 1 minus 2 equal to 1. So, implying c is equal to. So, you have constant as e square by e square plus 1 with in introduce the concept of expectation, and your expectation is nothing, but simply the mean of the population mean of the sample.

So, how can we make use of the concept of example expectation, let us discuss few simple examples. So, imagine you know you are waiting in your room for a letter to arrive and the waiting time. So, x is a waiting time. So, you want to wait.

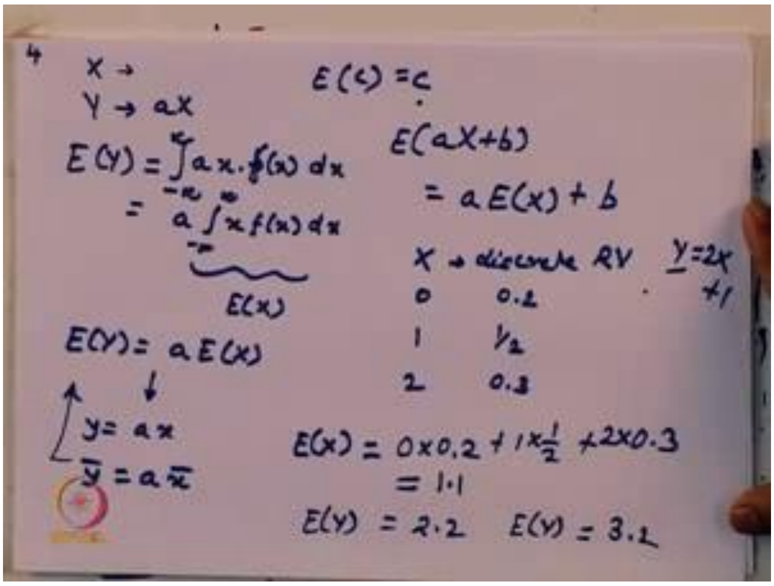
(Refer Slide Time: 05:07)



So, you are expecting the letter to arrive at 5 pm but you have to wait and the waiting time is a random variable x, which is continuous and it is defined as follows it, follows the probability distribution f of x given by. So, the waiting time x is a continuous distribution continuous random variable which follows with the probability distribution as follows. So, you want you ideally expect the, you know the envelope or whatever you waiting for to arrive at 5 pm. Now, how long will you on an average how much do you have to wait that is the question. So, how do you do this this is nothing, but. So, when you say what is an average expected wait time it is nothing, but the expectation x and this I can write as 0 to 2 half into x into d x. So, half is your p x this is your f x. So, x f x d x is nothing, but the expectation. So, either way for x cannot be less than 0. So, you can 0 you can integrate from 0 to 2 and 2 to infinity.

Here, your f of x is nothing, but 0 x d x which is f x is 0. So, essentially this thing is 0. So, you have to integrate from 0 to 2 half d x is equal to half into x s square by 2 0 to 2. So, sorry, so this will be written your value of one, which means that, if this is your distribution for waiting time on an average you will have to wait for one hour to receive your mail. So, let us now do some basic algebra with expectations. So, imagine you have this random variable x, and you have another variable y which is defined as a of x, what is going to be the value of E y.

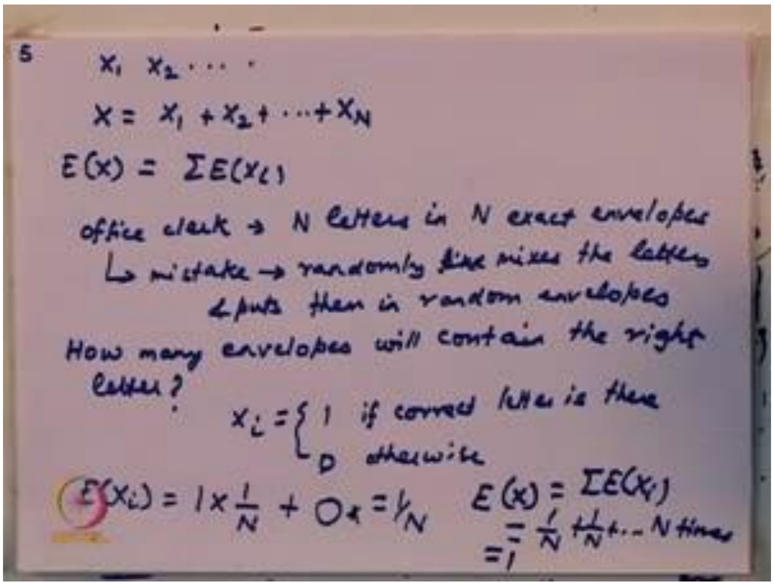
(Refer Slide Time: 07:04)



So, this assumes that the probability for given x the underline probability density function is the same. So, I can write e of y if this was a continuous variable let say E of y is simply integral a into. So, a x into p of x f of x d x. So, this I can take the a out. So, let say I can write minus infinity to infinity x f x d x. Now, what is this? This is nothing, but expectation of x. So, I will have E of y is a E of x. So, this is similar. So, you remember from long back we had done this basic transformation, if you define y is equal to a of x what is y bar is equal to a x bar and this is nothing, but the exact same equation, because expectation of y is simply y bar.

So, this is the same equation. So, we can use the similarity concept and then ask what is e of a constant c e of a constant c is nothing, but the c. So, if you are variable is always constant then on an average you always expect to the constant value. So, we can generalize this and say a of x plus b is equal to a e of x plus b we can extend this. So, let us do a simple example let us say that x is a variable which has discrete random variable. So, let say discrete r v which takes on values of 0 1 and 2 with probabilities 0.2 half and 0.3, so I can calculate the value of expectation of x this as simple equal to 0 in 2.2 plus 1 into half plus 2 into 0.3 on an average I have 0.6 plus 0.5 is 1.1. So, if I define a value y I will let say 2 f x y is equal to 2 f x then E of y is simply going to be 2.2, if I define y is 2 x plus 1 then E of y is going to be 3.2 so on and so forth.

(Refer Slide Time: 09:59)

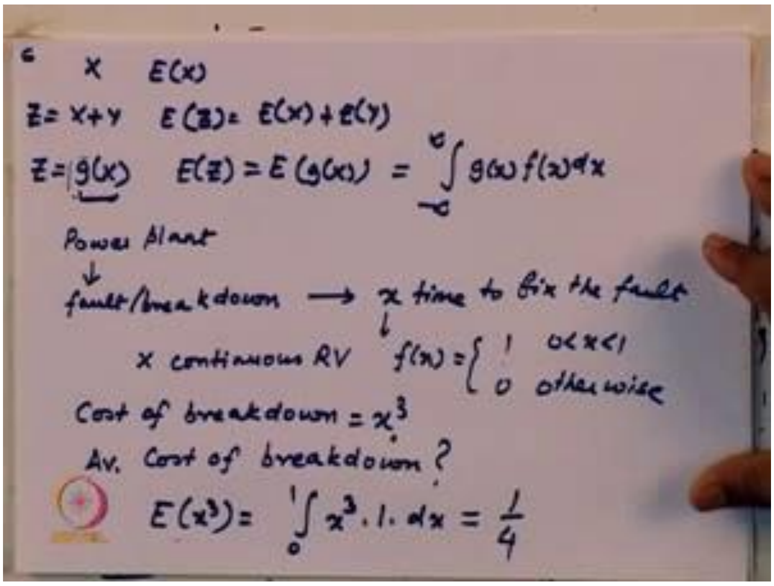


So, now let say you have different random variables X one X 2 so on and so forth. These are independent and I can think of a random variable called X is equal to X 1 plus X 2 plus dot dot dot plus X N. So, what will be my expectation of X expectation of X will come out to be summation E of X i. So, let us take a very simple example see imagine the office clerk has to put N letters in n envelopes N exact envelopes. So, by mistake, so when we say in N letters in N exact envelopes. So, every letter should ideally be in one particular envelope. So, the office clerk makes a mistake and randomly mixes is to mixes the letters and puts them in random envelopes. So, how many how many envelopes will contain the right letter is the question. So, how do we do it?

So, you have N letters and in N exact envelopes to be put let say we define X I, X i is a random variable which can take into values 1 and 0 1 if correct letter is there and 0 otherwise. So, if you have N letters and N envelopes what is the chance what is the chance that you will put the correct letter if you randomly take if you choose one letter from N letters and you put it in one envelope what is the chance the that envelope will have the correct letter there is only one possibility out of total of. So, you can drag out one out of N envelopes in n different ways. So, your expectation of X i is nothing, but one into probability of that is 1 by N X into the probability is 1 by N as we calculated plus for all are the cases it is 0, because the X i is 0 and you know. So, your X value 0 and your probability can be anything. So, all of them is also 1 by N.

So, this is nothing, but 1 by N. So, this is that the first the expected expectation that the first envelope is in the 1st letter is first in low. So, my total expectation then is similarly the summation of E X i N times right 1 by N plus 1 by N, N times what is the value then 1 by N into N is simply equal to 1. So, it means that on an average only 1 letter will be in its correct envelope.

(Refer Slide Time: 14:09)

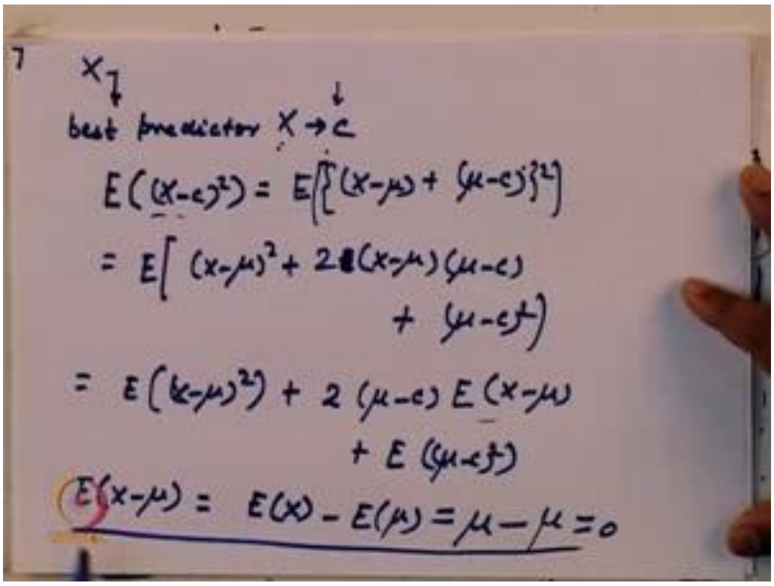


Let us take another (Refer Time: 14:10). So, we can generalize. So, you have X and you can have expectation of X you have X plus Y your expectation of Z. So, let say if you define Z is X plus Y you of Z is simply E of X plus E of Y you can define a problem as a variable Z as g of x in that case E of Z is simply is equal to is E of g of x and what is expectation of g of x is turns out to be integration over minus infinity to infinity g x into whatever the probably distribution you had g x into f x into d x.

So, let us take one simple example of g you know how you can calculate expectation of a function. So, let say in a plant in a power plant, when there is a fault. So, fault or breakdown it takes x amount of time x amount of time to fix the fault. So, and this x, x is a random variable continuous random variable it follows the following probability distribution this is the probability distribution now, you have let say. So, there is a cost associated with one particular with this breakdown the cost of breakdown is x cubed. So, the question is what is the average cost of break down.

What is the average cost of breakdown what will you do. So, essentially you have a function g of x which is x cubed and this is a corresponding probability distribution. So, essentially we are expected we want to find out expectation of x cube this is nothing, but 0 to 1 x cubed into 1 into d x why we had only the limit 0 to 1 because this is the only place where your probability density with p d f has some value. So, this is nothing, but one-fourth. So, this is how you can you know define new functions and if the probability underline probability distribution is the same you can find out the expectation of this new function which is derived from other random variables.

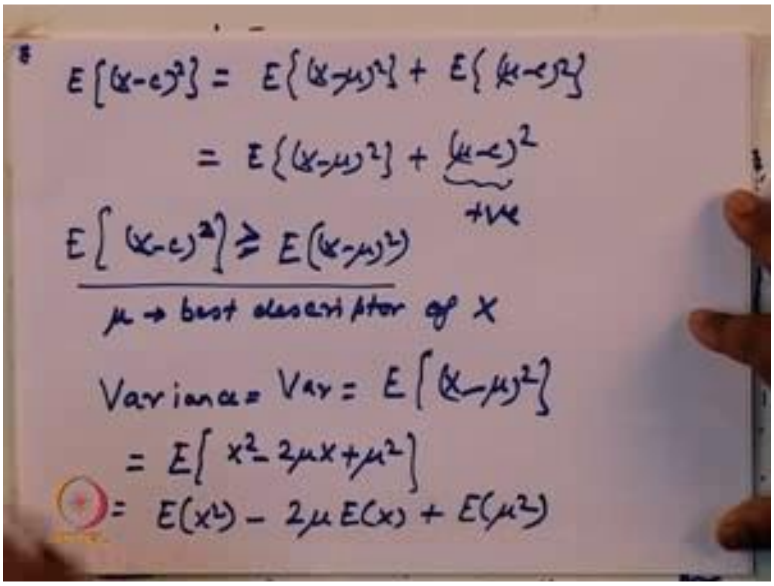
(Refer Slide Time: 17:19)



So, in expectation, you have let say a random variable x we want to know you want to know what is the best descriptor best predictor that x what is the prediction that x will have a value of c what is the value of c which best predicts x. So, in a sense we want to compute the best predictor of x what should we do we should minimize the difference between the value of x and the prediction in other words you want to minimize x minus c whole square. So, the best predictor is what will give you the lowest of expected value for x minus c whole square.

So, I can write down this E of x minus c whole square I can expand that whole square I can write it as expectation of c plus. So, I can expand them laterally expectation of I have the following expression now what is the value of expectation of x minus mu what is the value of expectation of x minus mu expectation of x minus mu is equal to expectation of x minus expectation of mu what is the expectation of x it is mu what is the expectation of mu? Mu is a constant. So, expectation of mu is equal to mu. So, expectation of x minus 0 x minus mu is simply equal to 0.

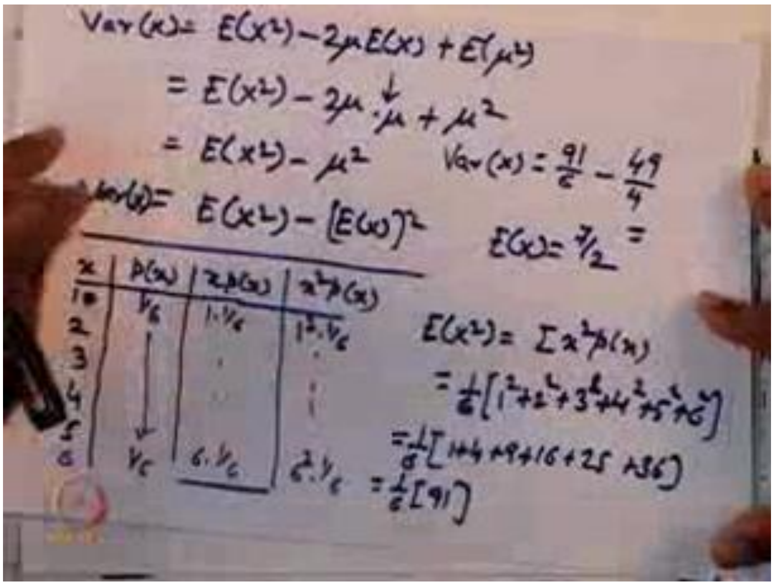
(Refer Slide Time: 19:40)



So, we can write this expression we can modify this expression plus. So, your expectation of x minus c whole square is expectation of x minus mu whole square plus expectation of mu minus c whole square. So, since, I can write this. So, expectation of mu minus c whole square is. So, this is I can further simplify plus. So, expectation this is a constant. So, expectation of this constant is simply that value and mu minus c whole square is a square. So, that is means this is always positive.

So, I can write x minus expectation of x minus c whole square is always greater or equal to x minus mu whole square expectation of x minus. So, this means that my mu is the best description this is why the average is taken such a standard matrix for defining for defining the population of the sample. So, while we have defined this we can also define x variance (Refer Time: 21:22) variance this is defined as expectation of. So, variance is nothing, but expectation of x minus mu whole square. So, I can expand this.

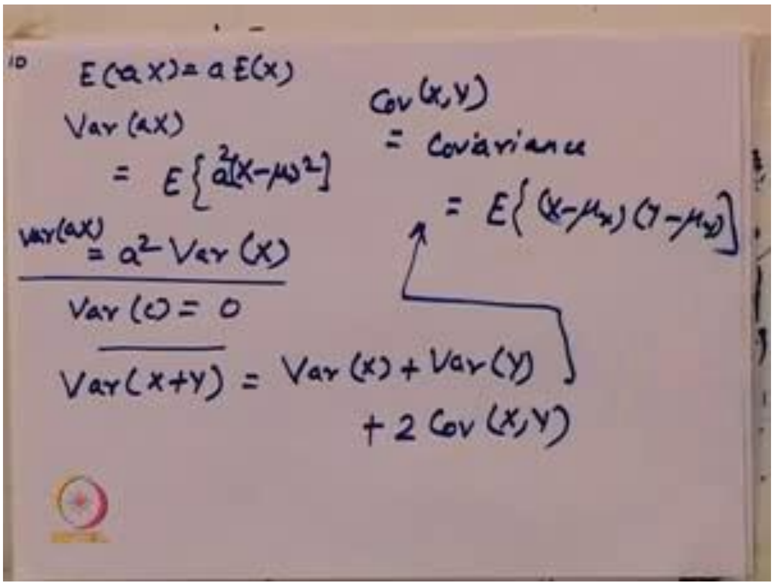
(Refer Slide Time: 22:04)



So, x expectation of x is simply equal to mu and expectation of mu square is simply equal to mu square. So, this gives me. So, you have minus 2 mu square term and a plus mu square term. So, this will give you of minus mu square term. So, variance of x, this is the final expression for variance of x expectation of x square minus expectation of x whole square.

So, let say for the die roll you have x values of g 1 2 3 4 5 and 6, 1 2 3 4 5 and 6, and you have probability all of 1 by 6. So, what is my variance? So, for calculating my expectation of x, so we have x p x and x square p of x. So, x p x is 1 dot 1 by 6 dot dot dot 6 dot of by 6 and we had calculated. So, expectation of x we are already calculated before it turned out to be 7 by 2 let us calculate the value of x square p x it is 1 square dot 1 by 6 like this 6 square dot 1 by 6. So, you have e of x square is summation x square p x plus 4 square plus 5 square plus 6 square is 1 by 6 1 plus 4 plus 9 plus 16 plus 25 plus 36 by 636 to 6 in 25, 91. So, your variance of x will turn out to be 91 by 6 minus x 7 by 2 minus 49 by 4. So, you can simplify this and calculate the final value of variance. So, just we have the following simplifications.

(Refer Slide Time: 24:56)



So, E of a x is equal to a E of x, what about variance of a x? Variance of a x is nothing but you will turn out. So, let us see this expectation of a into x minus mu whole square a square. So, variance of a x is equal to expectation of x mu minus. So, you can take out the a square term variance of a x is going to be a square variance of x variance of a constant is equal to going to be 0. So, you can also write down the expansion for variance of x plus y. So, this is a little bit tricky, but you will see that in the most general case it will come out to be variance of x plus variance of y plus 2 times covariance of x comma y where covariance of x comma y. So, we have cov of x comma y is defined as the covariance is expectation of x minus mu x into y minus mu y.

We will stop here for today, and in the next class we will briefly discuss about covariance and then we will start discussing about 3 important random variables which are relevant to statistics with that I thank you for your attention and look forward to next class.