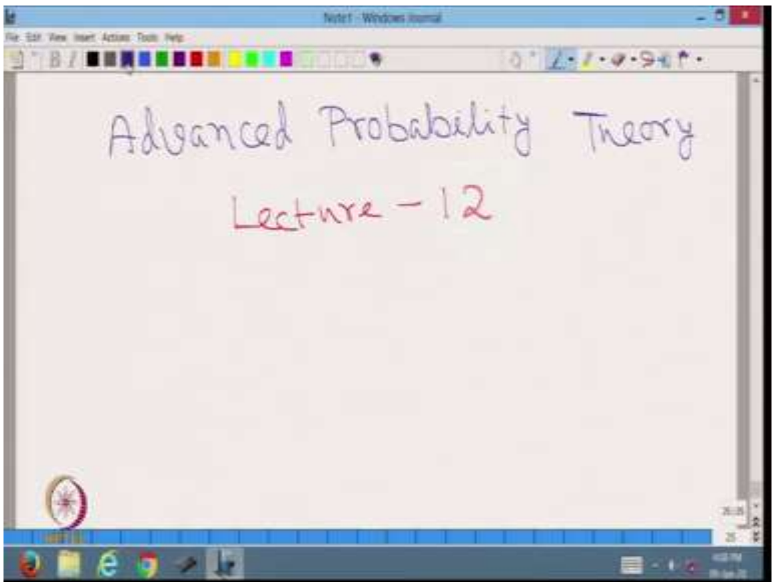
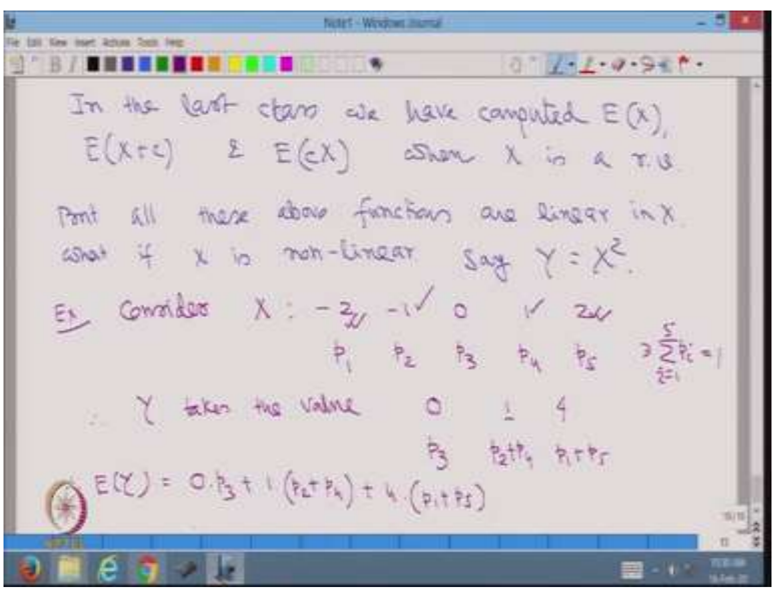
Advanced Probability Theory
Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi
Lecture 12

(Refer Slide Time: 00:25)



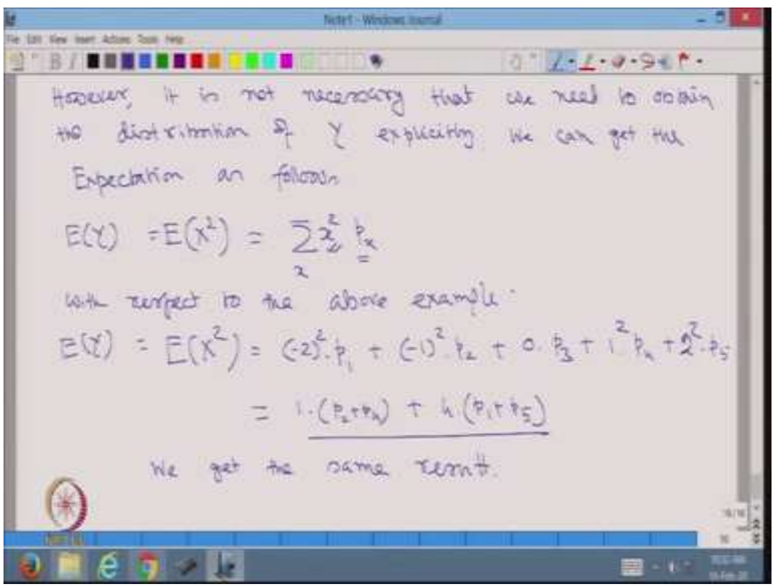


Welcome, students to the MOOCS lecture series on Advanced Probability Theory, this is lecture number 12. In the last class, we have computed expectation of X, expectation of X plus c and expectation of c X when X is a random variable. But all these above functions are linear in X. What if x is non-linear?

Say, Y is equal to X square. Example, consider X which takes the value minus 2, minus 1, 0, 1 and 2. Suppose the probability is at p1, p2, p3, p4 and p5, such that sigma pi, i is equal to 1 to 5 is equal to 1. Therefore, Y takes the value 0, 1 and 4, Y will take the value 0 when X is equal to 0. Therefore, the probability is p3. Y will take the value 1 when X is minus 1 or X is plus 1.

Therefore, this probability is p2 plus p4 and y is tilted the value 4 if X is minus 2 or plus 2 and that probability is p1 plus p5. Therefore, expectation of Y is equal to 0 times p3 plus 1 times p2 plus p4 plus 4 times p1 plus p5.

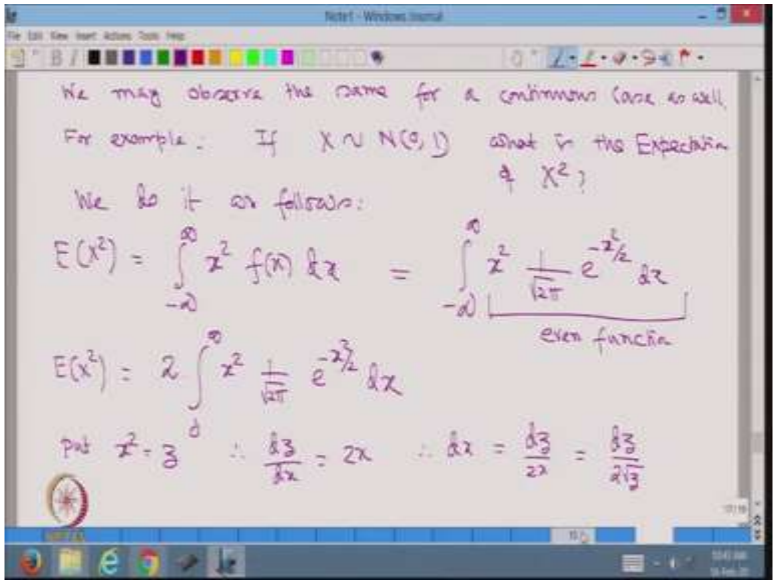
(Refer Slide Time: 03:19)

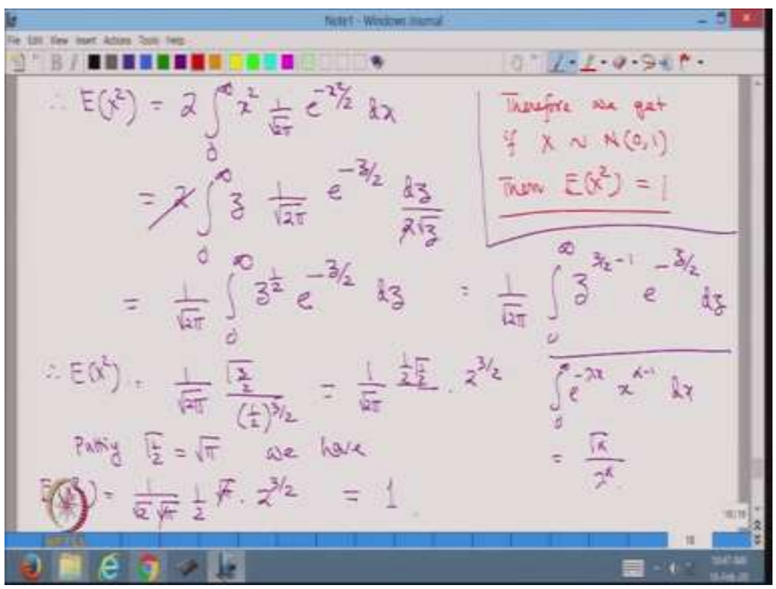


However, it is not necessary that we need to obtain the distribution of Y explicitly. We can do it, we can get the expectation as follows. Expectation of Y is equal to expectation of X square is equal to sigma X square into px X that is I am changing the value from X to X square and multiplying by the probability of X and I am summing over X.

So with respect to the example, above example, expected value of Y is equal to expected value of X square is equal to minus 2 square times p1 plus minus 1 square times p2 plus 0 times p3 plus 1 square times p4 plus 2 square times p5. And we can see that this is equal to 1 times p2 plus p4 plus 4 times p1 plus p5, that is, we get the same result.

(Refer Slide Time: 05:41)





We may observe the same for a continuous case as well. For example, if X is normal 0, 1, what is the expectation of X square? So we do it as follows. Expectation of X square is equal to minus infinity to infinity X square into fx, dx which is equal to minus infinity to infinity X square 1 over root over 2 Pi e to the power minus X square by 2 dx.

Now note that this is an even function, therefore expectation of X square is equal to 2 into integration 0 to infinity, X square 1 over root over 2 Pi e to the power minus X square by 2 dx. Put X square is equal to z, therefore dz dx is equal to 2x. Therefore, dx is equal to dz upon 2x is equal to dz upon 2 root over z.

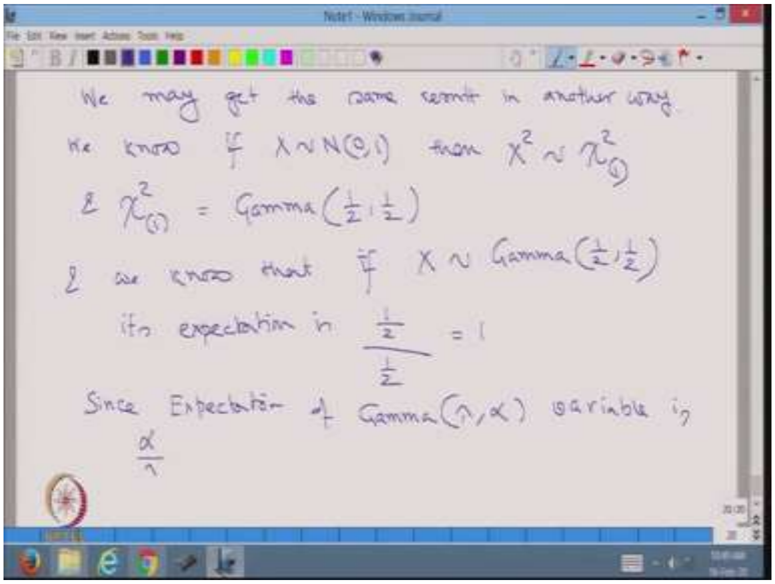
Therefore, expectation of X square is equal to 2 into 0 to infinity, X square 1 over root over 2 Pi e to the power minus X square by 2 dx is same as 2 into integration 0 to infinity, put z in place of X square and we get z 1 over root over 2 Pi e to the power minus z by 2 dz 2 root z is equal to this 2 cancels this 2.

Therefore, we have 1 over root over 2 Pi integration to infinity, z to the power half e to the power minus z by 2 dz is equal to 1 over root over 2 Pi 0 to infinity z to the power 3 by 2 minus 1 e to the power minus z by 2 dz. Now this is a gamma integral and we know that integration e to the power minus lambda x, x to the power alpha minus 1, 0 to infinity dx is equal to gamma alpha upon lambda to the power alpha.

Therefore, this expectation of X square is equal to 1 over root over 2 Pi gamma 3 by 2 divided by half to the power 3 by 2 is equal to 1 over root over 2 Pi gamma 3 by 2 we write it as half gamma half multiplied by 2 to the power 3 by 2.

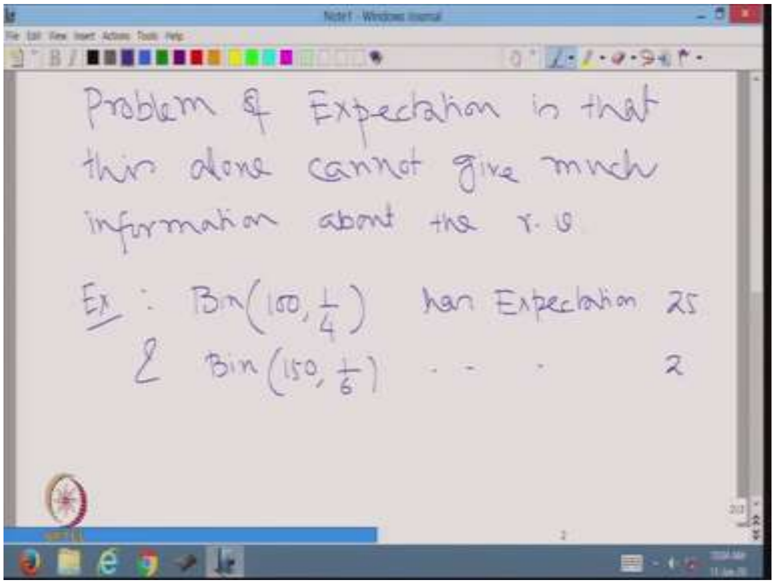
Putting gamma half is equal to root Pi, we have expectation of X square is equal to 1 over root over 2 root over Pi half root over Pi into 2 to the power 3 by 2 which is equal to all the term gets cancelled is equal to 1. Therefore, we get, if X is normal 0, 1 then expectation of X square is equal to 1.

(Refer Slide Time: 11:22)



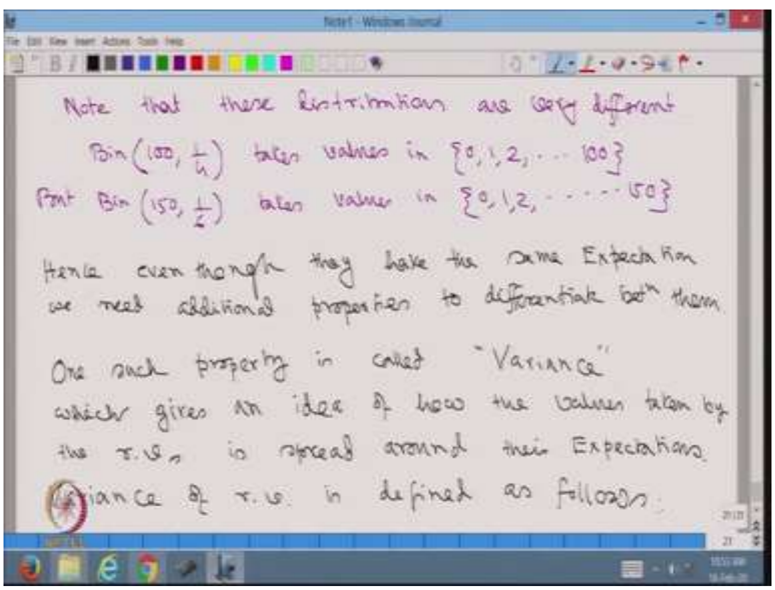
We may get the same result in another way. We know if X is normal 0, 1, then X square is distributed as chi square with 1 degrees of freedom and chi square with 1 is equal to gamma half comma half. And we know that if X is distributed as gamma half comma half, its expectation is half over half is equal to 1, since expectation of gamma lambda comma alpha variable is alpha over lambda that we have seen already.

(Refer Slide Time: 13:02)



Now problem of expectation is that this alone cannot give much information about the random variable. For example, binomial 100 comma 1 by 4 has expectation np which is 25 and binomial 150 comma 1 by 6 also has expectation 25.

(Refer Slide Time: 14:15)

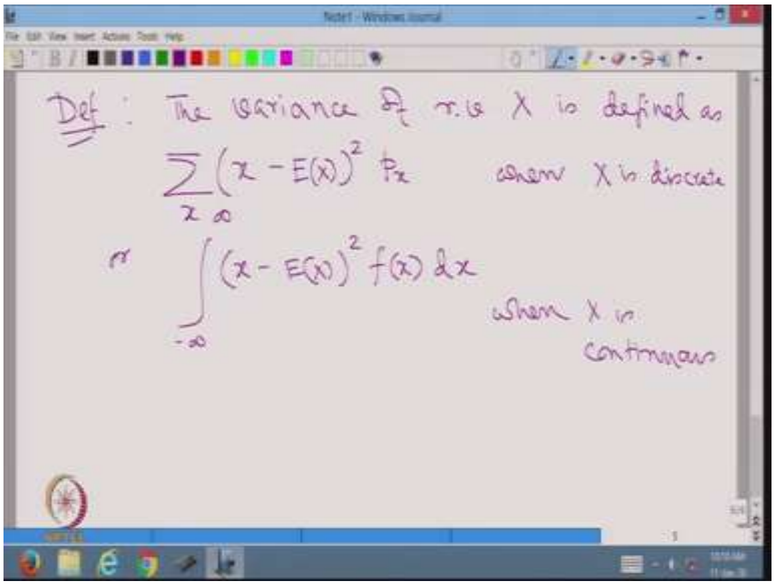


Note that, these two distributions are very different. Say, binomial 100 comma 1 by 4 takes values in 0, 1, 2, up to 100, but binomial 150, 1 by 6 takes values in 0, 1, 2, up to

150. Hence, even though they have the same expectation, we need additional properties to differentiate between them.

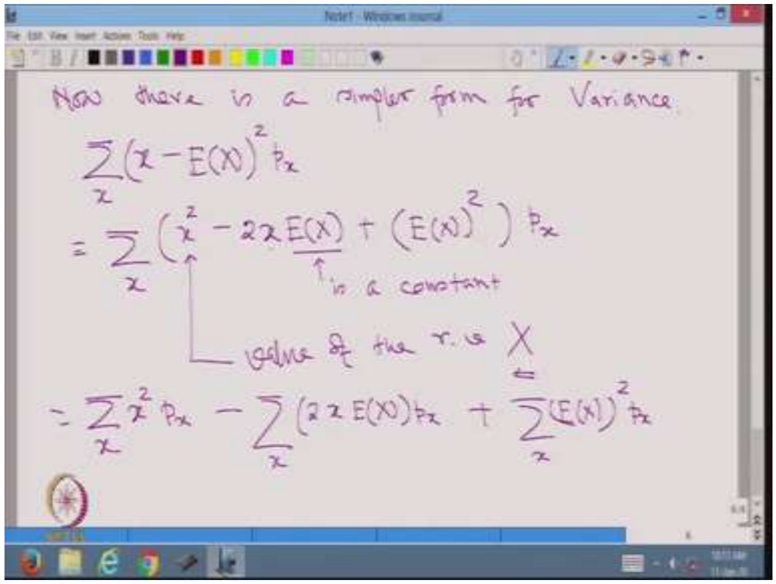
One such property is called variance which measures, which gives an idea of how the values taken by the random variables is spread around their expectations. Variance of a random variable is defined as follows.

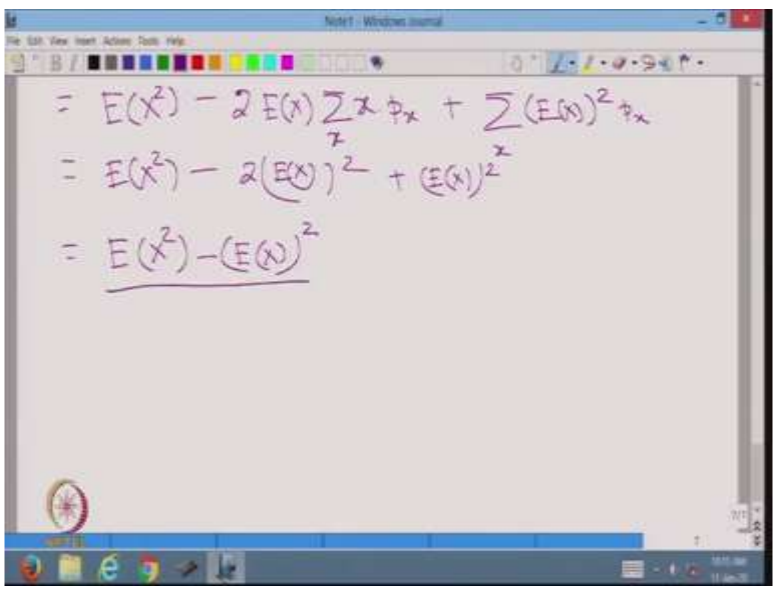
(Refer Slide Time: 17:12)



Definition, the variance of a random variable X is defined as sigma over X, X minus expected value of X whole square into px when x is discrete and it is integration minus infinity to infinity, x minus expected value of x whole square into fx dx when x is continuous.

(Refer Slide Time: 18:12)





Now there is a simpler form for Variance. I am showing it for discrete but the same will apply for continuous also. Sigma over x, x minus expected value of x whole square px, we can write it as sigma over x, x square minus 2x into expected value of x plus expected value of x whole square into px.

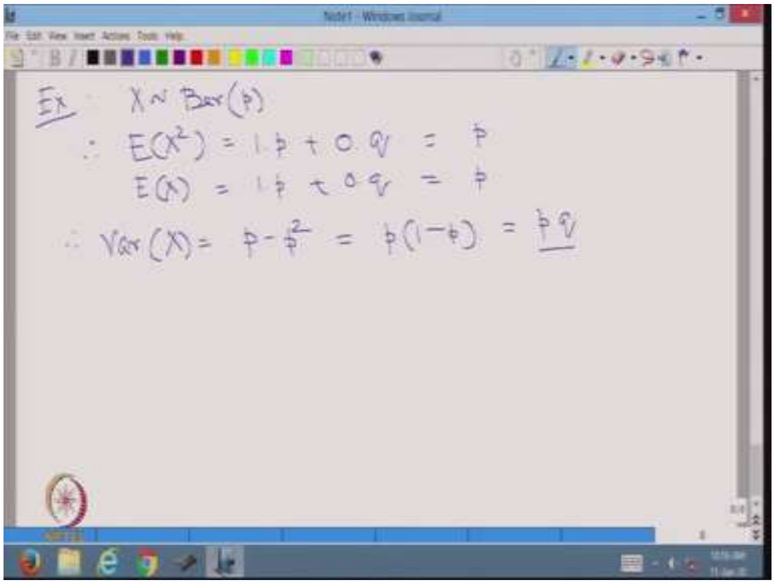
I want to emphasize on two things. One is that expected value of x is a constant. So we can take out of the summation as I will show later, but another thing is about notation, you can notice that here I am using small x which means value of the random variable x.

When we use capital x then that gives the name of the random variable. I have used it before also, but (())(19:51) did not explain. So in case some of you have the confusion, I am explaining it that when you are using small value that means it is a particular value that the random variable can take. But when you are using capital, then that is the name of the random variable. And as you can understand that here this is a possible value for the random variable, but x itself is the random variable.

That is why when I am writing expectation of x, I am putting it in capital X. So this is equal to sigma over x, x square Px minus sigma over x to expected value of x Px plus sigma over x expected value of x whole square into Px which is equal to expected value of x square minus 2 into expected value of x as it is a constant we take out of the summation sigma over x, x Px plus sigma over x expected value of x whole square times Px.

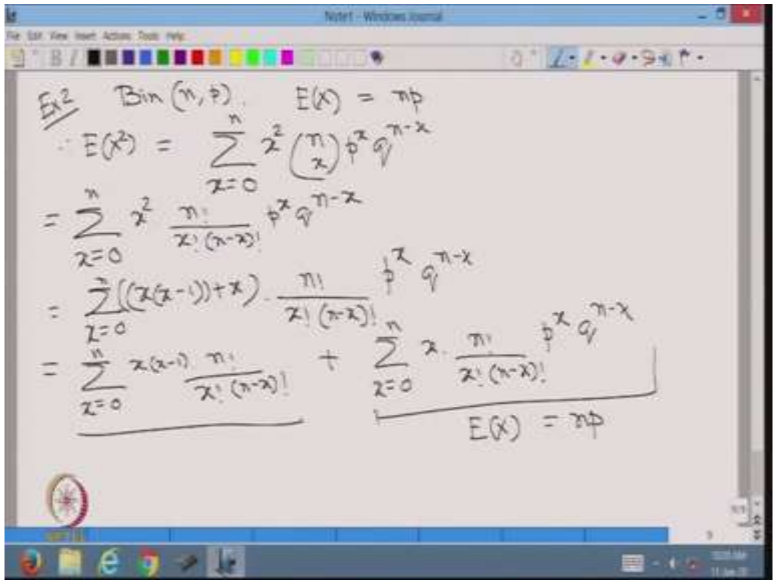
And this we can write it as expected value of x square minus 2 into expected value of x whole square plus expected value of x whole square. Because once we take out expected value of x square out of the summation, the remaining things adds to 1, is equal to expected value of x square minus expected value of x whole square. This is a very convenient formula for computing variance.

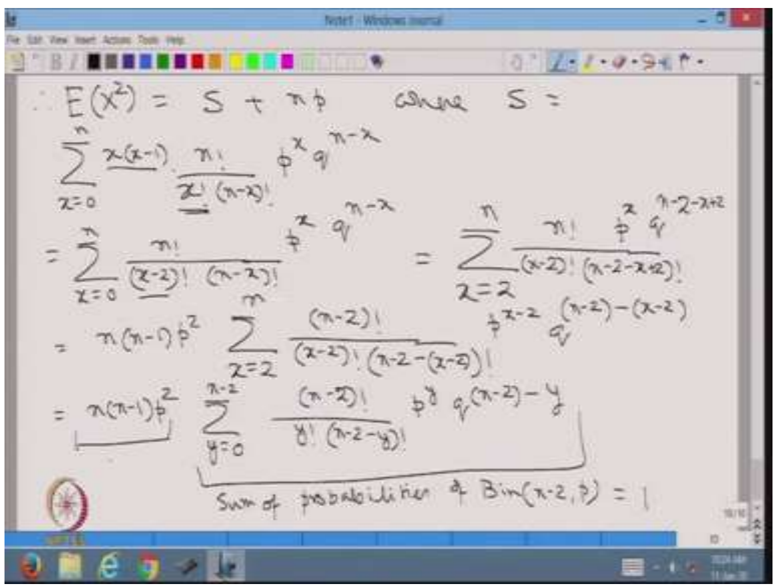
(Refer Slide Time: 22:15)



So let us give some example. X is Bernoulli with parameter plus, therefore, expected value of x square is equal to 1 times p plus 0 times q is equal to p. And expected value of x is also 1 times p plus 0 time q is equal to p. This is obvious because x takes values only 1 and 0, 1 with probability p and 0 with probability q. Therefore, variance of x is equal to p minus p square is equal to p into 1 minus p is equal to pq.

(Refer Slide Time: 23:15)





Example 2, binomial n, p, we know expected value of x is equal to n times p. Therefore, expected value of x square we can compute as follows, this is sigma over x is equal to 0 to n x square ncx p to the power x, q to the power n minus x, is equal to sigma over x is equal to 0 to n x square into factorial n upon factorial x factorial n minus x p to the power x, q to the power n minus x.

Now we write it as sigma x is equal to 0 to n, x into x minus 1 plus x. So I am breaking x square into two parts, it is x into x minus 1 plus x into n factorial into x factorial into n minus x factorial, p to the power x, q to the power n minus x, is equal to sigma x is equal to 0 to n, x into x minus 1 into n factorial into x factorial n minus x factorial plus sigma x is equal to 0 to n, x into n factorial, x factorial into n minus x factorial, p to the power x, q to the power n minus x.

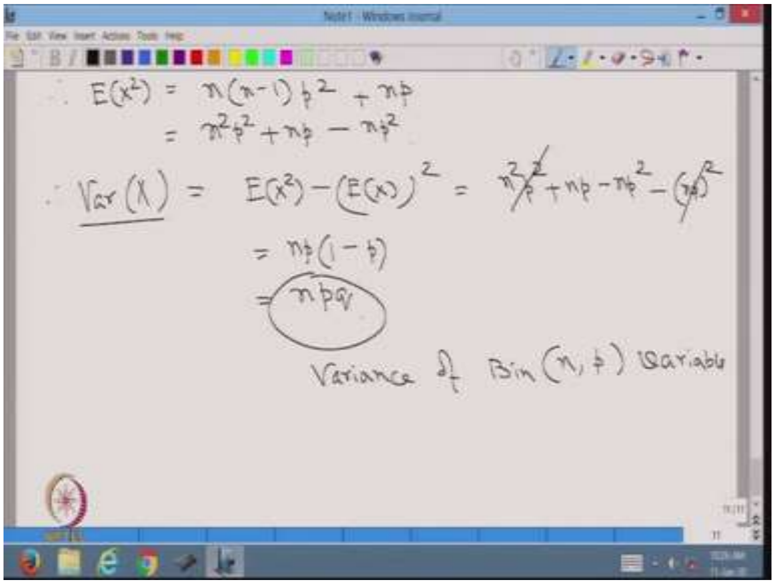
This part is, we know, is expected value of x which is is equal to np. So we need to calculate this part let us call it, say S, therefore, expected value of x square is equal to S plus np where S is equal to sigma, x is equal to o to n, x into x minus 1 into n factorial x factorial n minus x factorial p to the power x q to the power n minus x.

Now we cancel x into x minus 1 with factorial x, so two terms will be cancelled. So what we are left with is x is equal to 0 to n, n factorial x minus 2 factorial n minus n p to the power x q to the power n minus x. Since, negative, factorial of a negative number has no meaning, therefore we can change the limit of the sum from x is equal to 2 to n.

So this becomes x is equal to 2 to n, n factorial x minus 2 factorial, n minus 2 minus x plus 2 factorial, p to the power x, q to the power n minus 2 minus x plus 2, which is equal to, if we take n and n minus 1 common and therefore out of the summation and p square out of the summation, then we get, it is x is equal to 2 to n, n minus 2 factorial upon x minus 2 factorial with n minus 2 minus x minus 2 factorial, p to the power x minus 2 q to the power n minus 2 minus x minus 2.

A small change of variable will make it sigma y is equal to 0 to n minus 2, n minus 2 factorial upon y factorial n minus 2 minus y factorial p to the power y q to the power n minus 2 minus y, what is this sum? This is sum of probabilities of binomial n minus 2 comma p. Hence this is equal to 1. Therefore S comes out to be n into n minus 1 p square.

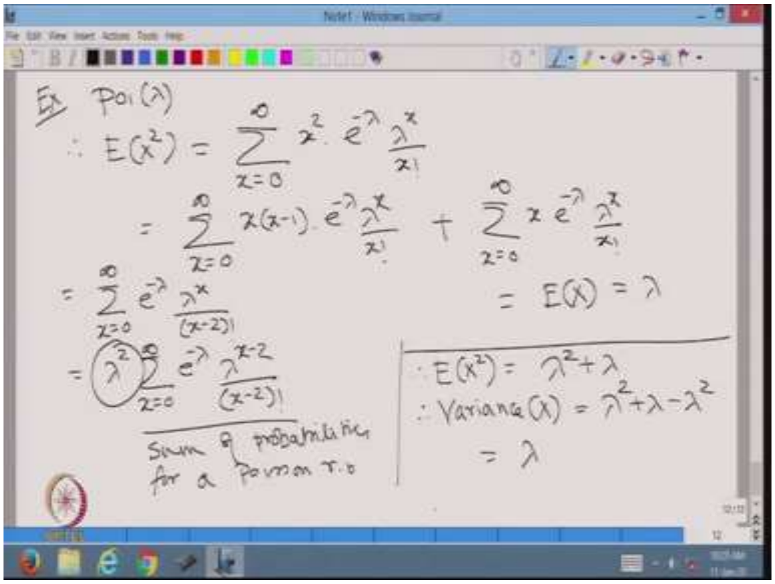
(Refer Slide Time: 28:59)



Therefore, expected value of x square is equal to n into n minus 1 p square plus np is equal to n square p square plus n p minus n p square. Therefore, variance of x, which we write like that, is equal to expectation of x square minus expectation of x whole square is equal to n square p square plus n p minus n p square minus n p whole square.

So this cancels with this, therefore we are left with n p into 1 minus p which is equal to np q. So that is the variance of binomial n comma p variable.

(Refer Slide Time: 30:19)

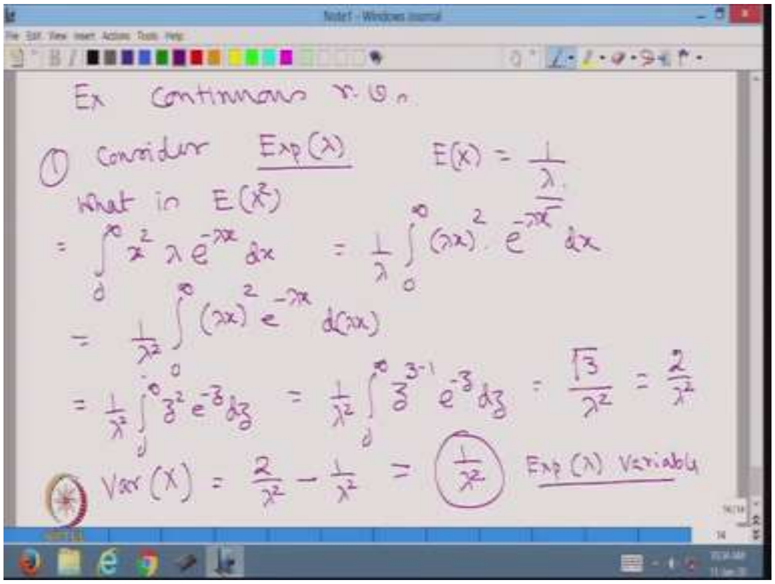


Let me do one more example from discreet case. Poisson with parameter lambda, therefore expected value of x square is equal to sigma x is equal to 0 to infinity x square into e to the power minus lambda, lambda power x upon factorial x is equal to in a very similar way we can write it as x is equal to 0 to infinity.

X into x minus 1 to e to the power minus lambda, lambda power x upon factorial x plus sigma x is equal to 0 to infinity, x e to the power minus lambda, lambda power x upon factorial x. This is, as before, is going to be expected value of x is equal to lambda. This on the other hand, we can write it as sigma x is equal to 0 to infinity, e to the power minus lambda, lambda power x upon x minus 2 factorial is equal to.

If we take lambda square out of the summation, sigma x is equal to 0 to infinity e to the power minus lambda, lambda to the power x minus 2 upon factorial x minus 2, sum of probabilities for a Poisson random variable. Therefore, this comes out to be lambda square, therefore expected value of x square is equal to lambda square plus lambda. Therefore, variance of x is equal to lambda square plus lambda minus lambda square is equal to lambda.

(Refer Slide Time: 32:43)

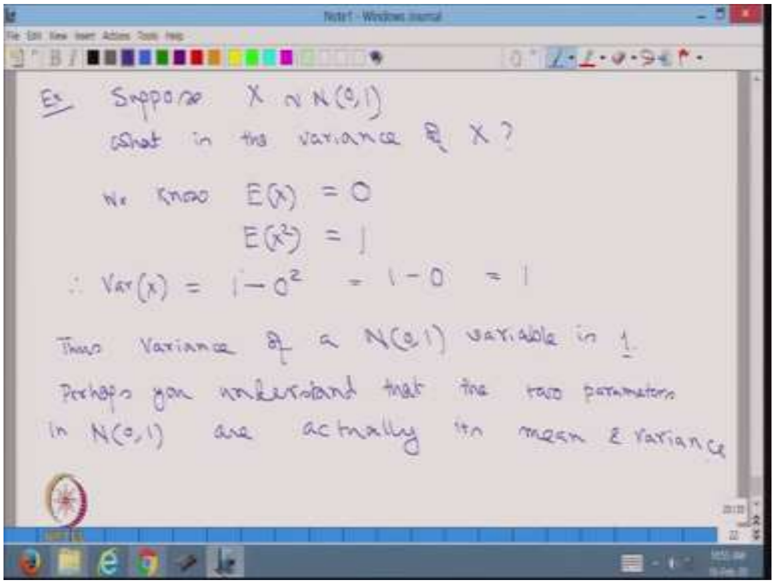


Example of continuous random variables, consider exponential lambda, we know that expected value of x is equal to 1 upon lambda. Question is what is expected value of x square? This is equal to integration 0 to infinity x square into lambda e to the power minus lambda x dx is equal to 1 upon lambda into 0 to infinity lambda x whole square into e to the power minus lambda x dx is equal to 1 upon lambda square 0 to infinity lambda x square e to the power minus lambda x into d lambda x.

Now this is is equal to 1 upon lambda square 0 to infinity z square e to the power minus z dz is equal to 1 upon lambda square 0 to infinity z to the power 3 minus 1 e to the power minus z dz. This is coming out to be gamma 3 upon lambda square. Now gamma 3 is equal to factorial 2, therefore this is 2 upon lambda square.

Therefore, variance of x when x is exponential lambda variable is equal to 2 upon lambda square minus expected value of x whole square, which is 1 upon lambda square is equal to 1 upon lambda square. So this is the variance of exponential lambda variable.

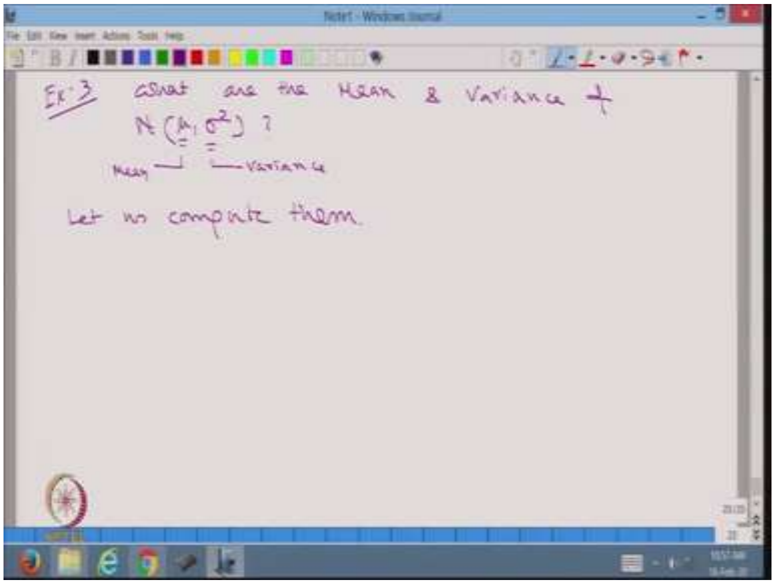
(Refer Slide Time: 35:34)



Example, suppose is distributed as normal 0, 1, what is the variance of x? We know expectation of x is equal to 0 and expectation of x square is equal to 1. Therefore, variance of x we can understand is 1 minus 0 square is equal to 1 minus 0 is equal to 1.

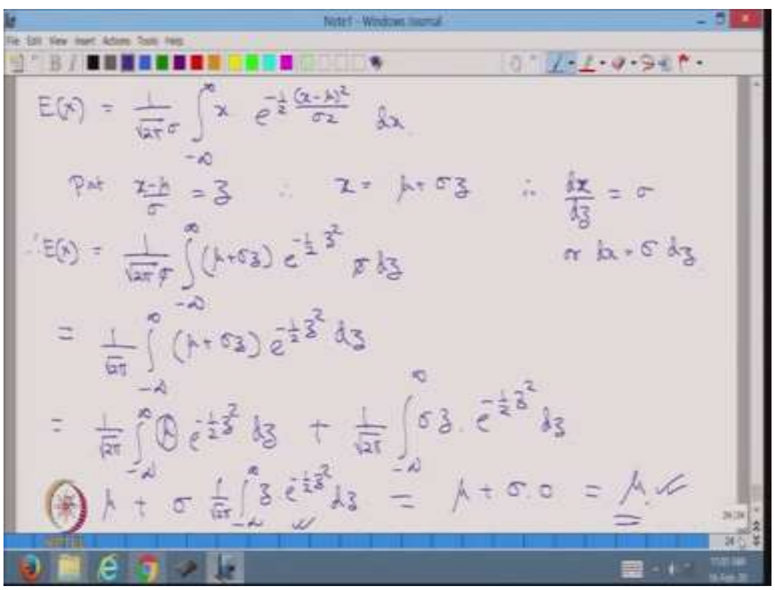
Thus, variance of a normal 0, 1 variable is 1. Now perhaps you understand that the two parameters in normal 0, 1 are actually its mean and variance.

(Refer Slide Time: 37:19)



So example 3, what are the mean and variance of normal Mu comma sigma square? So we understand that the mean is going to be mu and variance is going to be sigma square. Let us compute them.

(Refer Slide Time: 38:06)



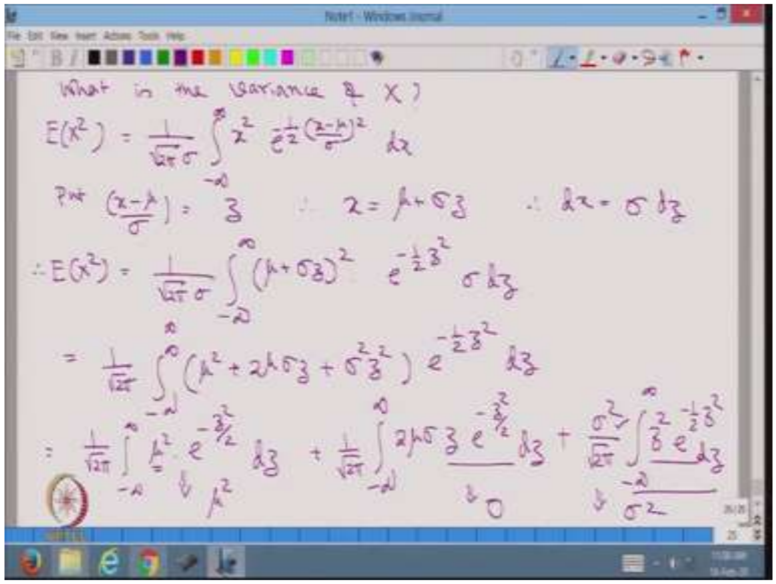
Expectation of x is equal to 1 over root of a 2 Pi sigma integration minus infinity to infinity x into e to the power minus half x minus Mu whole square upon sigma square dx.

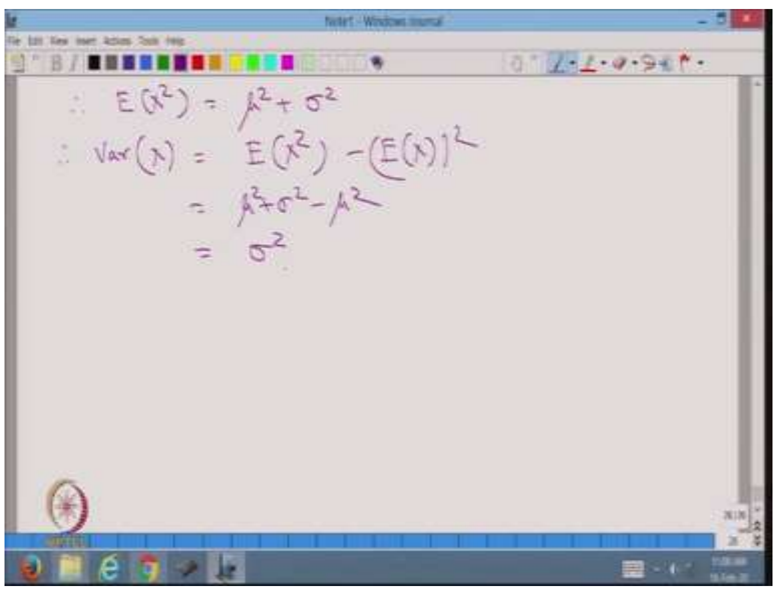
Put x minus Mu upon sigma is equal to z. Therefore, x is equal to mu plus sigma z. Therefore, dx dz is equal to sigma or dx is equal to sigma dz.

Therefore, expectation of x is equal to 1 over root over 2 Pi sigma minus infinity to infinity mu plus sigma z into the power minus half z square sigma dz is equal to now sigma gets cancelled 1 over root over 2 Pi integration minus infinity to infinity mu plus sigma z e to the power minus half z square dz is equal to 1 over root over 2 Pi minus infinity to infinity mu e to the power minus half z square dz plus 1 over root over 2 Pi minus infinity to infinity sigma z into e to the power minus half z square dz.

Now if I take out mu, we get the integration to be the integration of the normal, standard normal PDF. Therefore, this gives us mu plus sigma times 1 over root over 2 Pi minus infinity to infinity z into e to the power minus of z square dz and this is the expectation of normal 0, 1 variable, therefore 0. Therefore, the end result is mu plus sigma dot 0 is equal to mu. Thus, we get that expected value of x is equal to mu.

(Refer Slide Time: 41:04)





What is the variance of x? Expectation of x square is equal to 1 over root over 2 Pi sigma minus infinity to infinity x square e to the power minus half x minus mu by sigma whole square dx. Put x minus mu by sigma is equal to z. Therefore, x is equal to Mu plus sigma z and as before. Therefore, dx is equal to sigma dz.

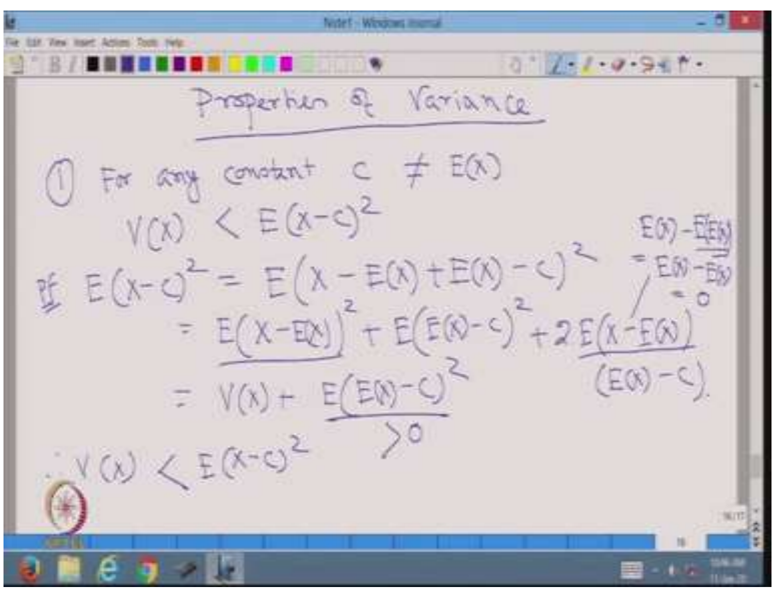
Therefore, expected value of x square is equal to 1 over root over 2 Pi sigma minus infinity to infinity mu plus sigma z whole square into e to the power minus half z square sigma dz is equal to 1 over root over 2 pi minus infinity to infinity mu square plus 2 mu sigma z plus sigma square z square into e to the power minus half z square dz.

Is equal to 1 over root over 2 Pi minus infinity to infinity mu square into e to the power minus z square by 2 dz plus 1 over root over 2 Pi minus infinity to infinity 2 mu sigma z e to the power z square by 2 dz plus sigma square upon root over 2 Pi integration minus infinity to infinity z square into e to the power minus half z square dz.

Now this is, I am integrating a constant mu square with the normal, standard normal PDF. Therefore, this will give me mu square. This is effectively computing the expected value of standard normal which we know is equal to 0 and this is expectation of x square when x is standard normal. Therefore, this integration will give us 1 and therefore the outcome is from this sigma square we get sigma square.

Therefore, expected value of x square is equal to mu square plus sigma square. Therefore, variance of x is expectation of x square minus expectation of x whole square is equal to mu square plus sigma square minus mu square is equal to sigma square.

(Refer Slide Time: 44:52)

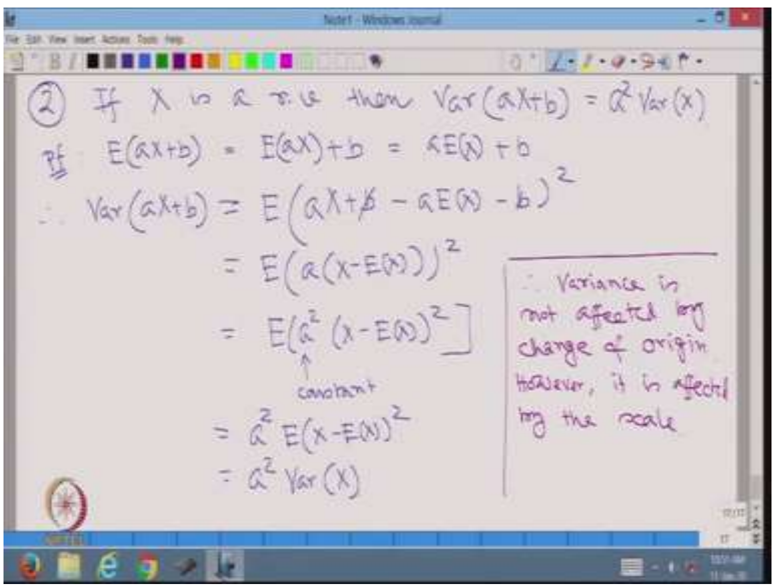


Let us now investigate certain properties of variance. For any constant c not equal to expected value of x, variance of x is less than expected value of x minus c whole square. This is very simple since expected value of x minus c whole square is equal to expected value of x minus expected value of x plus expected value of x minus c whole square is equal to expected value of x minus expected value of x whole square plus expected value of x minus c whole square plus 2 times expected value of x minus expected value of x into expected value of x minus c.

Since this is a constant and expected value of x minus expected value of x is going to be 0, because this is expected value of x minus expected value of x is equal to, because it is a constant. Therefore, it is expected value of x minus expected value of x is equal to 0.

Therefore what we are getting is equal to this term is variance of x plus this term is a positive quantity. It is greater than 0 therefore, variance of x is less than expected value of x minus c whole square.

(Refer Slide Time: 47:20)

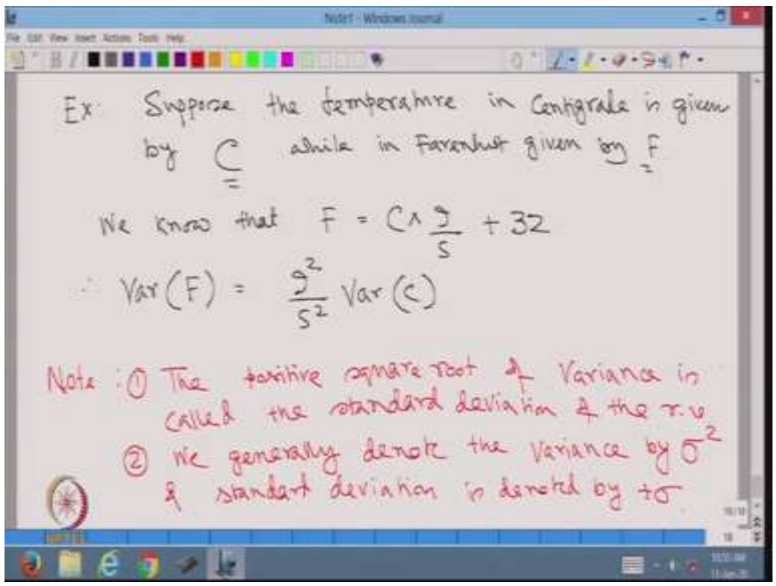


Property 2, if x is a random variable, then variance of ax plus b is equal to a square into variance of x. Proof, expected value of ax plus b is equal to expected value of ax plus b is equal to a times expected value of x plus b.

Therefore, variance of ax plus b is equal to expected value of ax plus b minus a times expected value of x minus b whole square, which is equal to expected value of a into x minus expected value of x plus 0 because they cancel each other square, is equal to expected value of a square into x minus expected value of x whole square.

Now since this is a constant, and we know that constant times expectation, expectation of a constant times a random variable is constant times expectation of the random variable. Therefore, we write it as a square into expected value of x minus expectation of x whole square is equal to a square into variance of x. Therefore, variance is not affected by change of origin. However, it is affected by the scale.

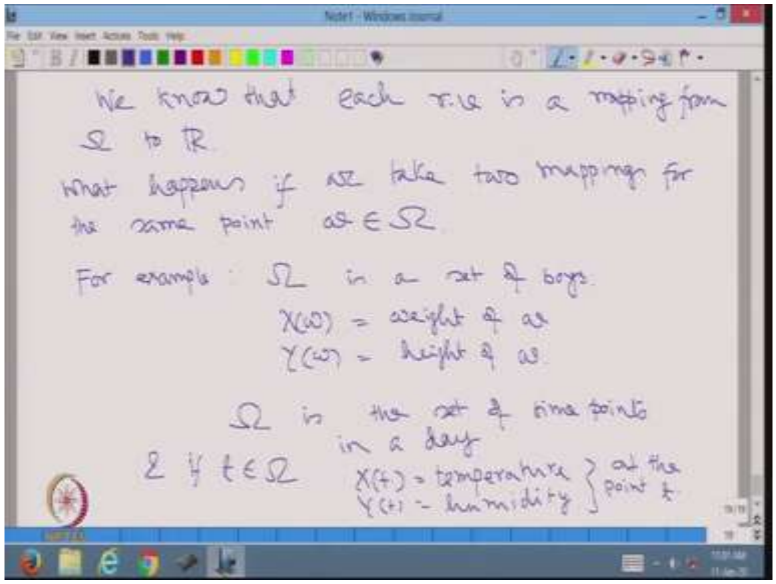
(Refer Slide Time: 50:00)

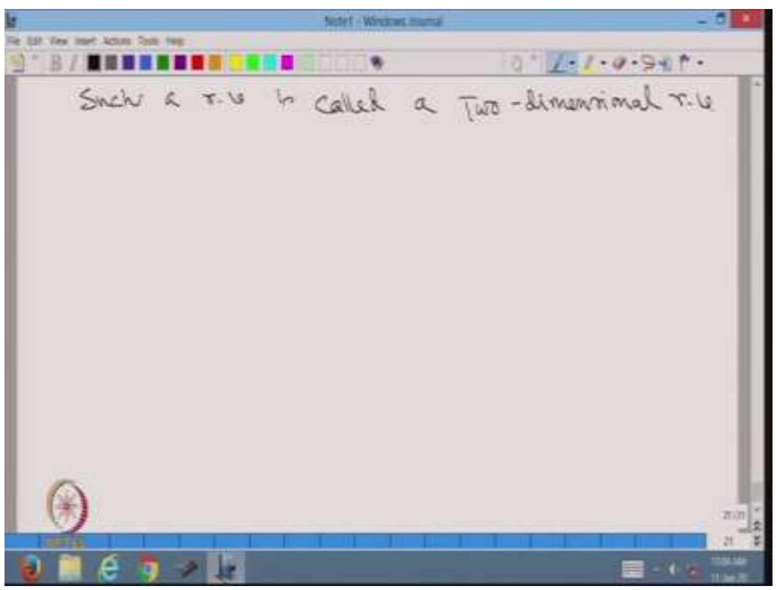


Example, suppose the temperature in centigrade is given by C while in Fahrenheit given by F. So if we consider these two to be the random variables, then we know that F is equal to C into 9 by 5 plus 32. Therefore, variance of F is equal to 9 square upon 5 square into variance when it is measured in terms of centigrade.

The positive square root of variance is called the standard deviation of the random variable. We generally denote the variance by sigma square. Sigma square, because variance has to be always positive, so the positiveness of the measurement is taken care of by the square and standard deviation is denoted by plus sigma.

(Refer Slide Time: 52:57)





So, so far we have discussed mean and variance of a random variable and we know that each random variable is a mapping from omega to R. What happens if we take two mappings for the same point omega belonging to omega?

For example, omega is a set of boys and X omega is equal to weight of omega and Y omega is equal to say height of omega. In a similar way, if omega is the set of time points in a day and if t is a time point belonging to omega, Xt is equal to temperature and Yt is equal to humidity at the point t.

In practice, we often consider such two different random variables and that is important because we may like to see the relationship between these two random variables for illustration how much the weight of a person is affected by his or her height. Such a random variable is called a two-dimensional random variable and I shall start with that in the next class. Okay, students. Thank you so much.