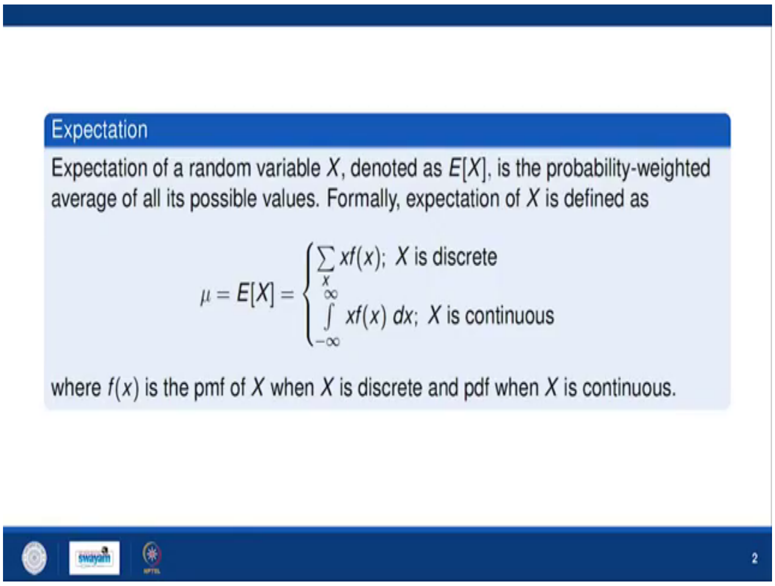
Essential Mathematics for Machine Learning
Prof. S. K. Gupta
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture – 33
Expectation and Variance

Hello friends, welcome to lecture series on Essential Mathematics for Machine Learning. So, in the last lecture we have seen what probability is and how we can find out conditional probability, basic concepts of Bayes theorem, etcetera. Now we will see how we can find out Expectation and Variance of a distribution function, of a random variable sorry. We want to find out expectation and variance of a random variable; how we can find?

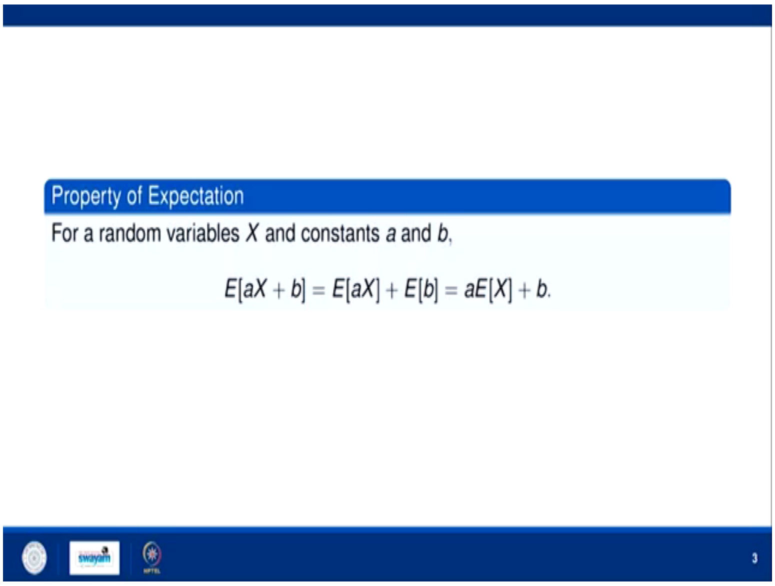
(Refer Slide Time: 00:53)



So, what do you mean by expectation first? So, expectation of a random variable X, denoted by E of X is the probability weighted average of all its possible values. So, basically sometimes we call, we also call expectation of X as a mean of a random variable X. So, formally expectation of X is defined as, we denote it by mu also which is expectation of X and for discrete case it is simply summation of x f x, where f x is the pmf of X.

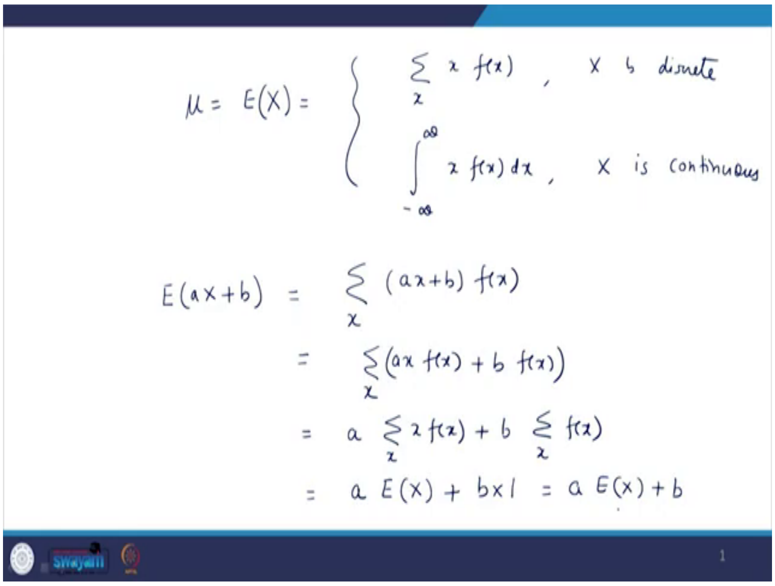
And if it is continuous, then it is integral from minus infinity to plus infinity x f x, where f x is in this case when X is continuous; in this case f is simply pdf ok, probability density function.

(Refer Slide Time: 01:51)



Now, now we have a important property of expectation that is for a random variable X and a constants a and b; expectation of a X plus b is nothing, but a times expectation of X plus b. So, so let us discuss this thing now.

(Refer Slide Time: 02:10)



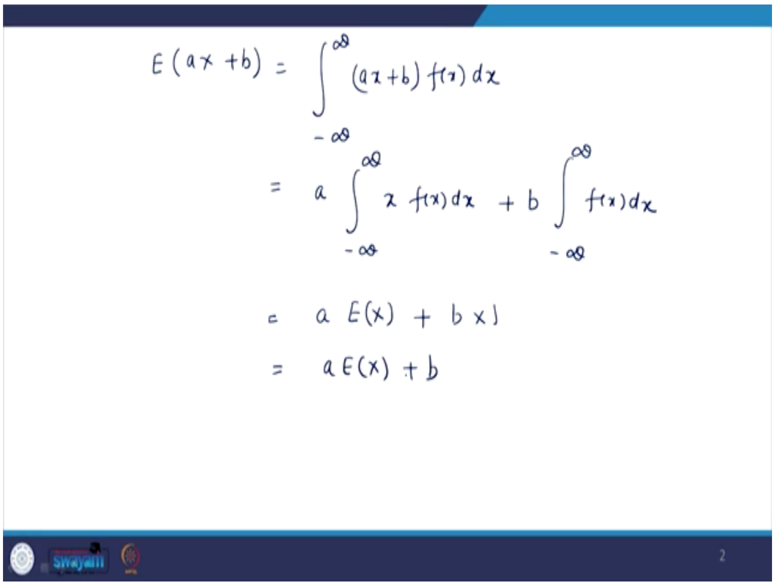
So, we have seen that if we have a random variable X, then expectation of X is given by is given by a summation over x, x f x, if X is discrete as we discussed and it is integral from minus infinity to plus infinity x f x d x, if X is continuous.

Now, suppose you want to find out expectation of a X plus b; if X is discrete, then how we can define expectation of X, simply given by summation x f x. So, it will be simply summation of a X plus b into f x summation over x, if x is discrete. So, this is given by summation a x f x plus b times f x summation over x. And there is this further can be written as a times summation over x x f x plus b times summation over x f x, and this can be further written as a times.

Now, this is what this is nothing, but expectation of X as per the definition of expectation of X when X is discrete. So, this is nothing, but expectation of X plus b times. Now summation of f, because f here is pmf; so this is nothing but 1, summation of f x is 1. So, this is into 1. So, we can say that it is a times expectation of X plus b.

So, basically we can say that, expectation follows linearity that is a X plus b is same as a times expectation of X plus b, this is for the discrete case.

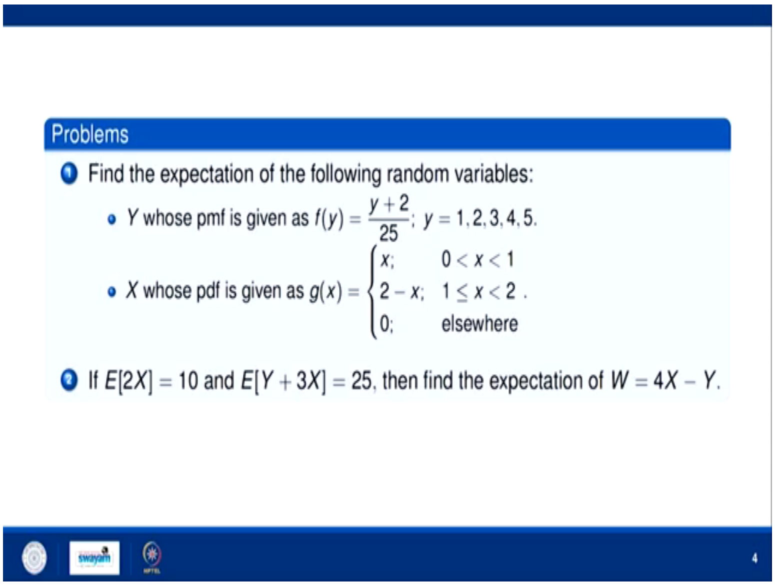
(Refer Slide Time: 04:27)



Now, similarly if X is continuous in that case, expectation of a X plus b will be given by integral minus infinity to plus infinity a x plus b into f x d x. So, which is nothing, but a times integral minus infinity to plus infinity x f x d x plus b times integral minus infinity to plus infinity f x d x. And as per definition of expectation of X, this term is nothing, but expectation of X.

So, it is a times expectation of X plus b and again this is 1; because here f is nothing, but pdf of X. So, this is into 1 again. So, this is equal to a times expectation of X plus b. So, we can say that expectation of a X plus b is always a times expectation of X plus b. So, this is the important property of expectation of X.

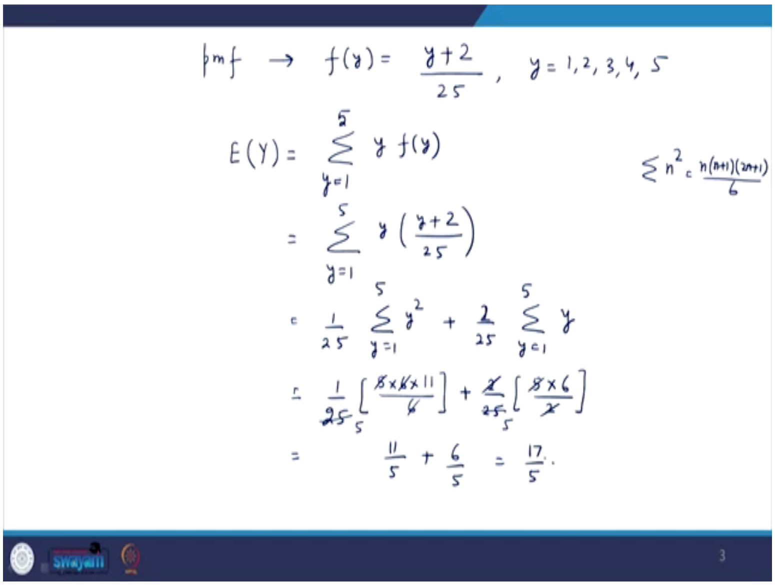
(Refer Slide Time: 05:32)



Now, let us discuss few problems based on expectation.

Suppose you have a discrete distribution, here you have a discrete distribution given by y plus 2 upon 25, this is basically pmf. This is probability mass function of this random variable y which is given by y plus 2 upon 25 y is varying from 1 to 5. And you want to calculate, you want to find out the expectation of this variable, this random variable Y. So, how can you find?

(Refer Slide Time: 06:06)



So, here what is, what is pmf? Pmf of y is it is f y, it is given by y plus, this is y plus 2 upon 25, where y maybe 1, 2, 3, 4, 5. Now you want to calculate expectation of Y, random variable Y. So, this is as per definition this is nothing, but y into f y ok, as per the definition of expectation of Y and here y is varying from 1 to 5; because the values of y maybe 1, 2, 3, 4 or 5.

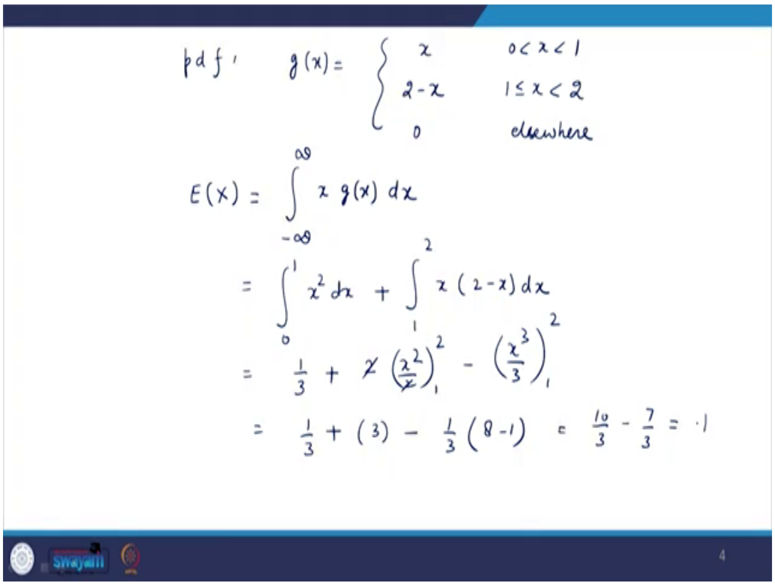
So, this is further equal to y varying from 1 to 5 y times y plus 2 upon 25. So, this is 1 by 25 times summation y varying from 1 to 5, this is y square plus 1 by 25 times; in fact 2 upon 25 times summation y from 1 to 5, it is y. Now this is nothing but this is further equal to 1 by 25 times.

So, this is 1 square plus 2 square plus 3 square up to 5 square. So, we know that either we can find out directly the sum up to, sum of squares up to 5 or we can use the formula; we know that sum of n square is nothing, but n, n plus 1 2 n plus 1 by 6, sum of squares of natural numbers. So, we can use this expression here. So, it is basically, n is 5 here; so it is 5, 5 into 6 into, when you substitute 5 here, it is 11, 11 by 6 and plus 2 upon 25 times. It is sum of natural numbers up to 5, by the formula we know that it is nothing, but n n plus 1 by 2.

So, it is 5 into 5 plus 1 is 6 upon 2. So, now, when you, when you simplify this. So, 6, 6 cancels out and it is 5, 5 times cancel out. So, it is 11 by 5 plus it is 5 times cancels out and it is 2, 2 cancels out. So, it is 6 upon 5. So, this comes out to be 17 upon 5. So, this is how we can compute expectation of this random variable Y for this problem.

Now, if we have a pdf for a continuous random variable given by this expression. So, in this case, how you can find expectation? So, expectation can be find in this way for this example.

(Refer Slide Time: 08:58)



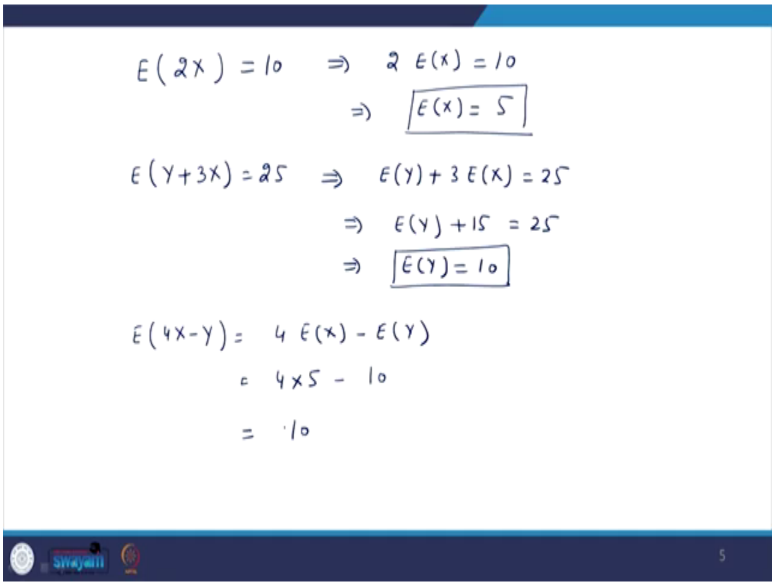
Now, what is pdf here? Pdf is given by g x which is x when x is between 0 and 1 and it is 2 minus x when x is between 1 and 2 and 0 elsewhere.

Now, you want to come, you want to find out expectation of X. So, this is as per definition, it is integral minus infinity to plus infinity x g x d x, ok. Here pdf is given by g x, so g x; g x will come here. It is further given, further written as, now this integral is 0; otherwise only accept 0 to 2, 0 to 1 is defined like x and 1 to 2 it is 2 minus x. So, we can say it is 0 to 1 given by x, so it is x square d x plus and 1 to 2 it is x into 2 minus x d x.

So, this will be x cube by 3 which is 1 by 3 plus this is 2 into x square upon 2 from 1 to 2 minus; it is x square, that is x cube upon 3 from 1 to 2. So, this is 1 by 3 plus 2 cancels out. So, it is 4 minus 1 is 3 minus, it is 1 by 3 times it is 8 minus 1. So, the final answer will be it is 9 plus 1, 10, this is 10 by 3 minus it is 7 by 3.

So, that will be nothing, but 1. So, in this way we can find out expectation of the random variable X for this pdf. Now this is a property based problem of a simple problem, it is known to us that expectation of 2 X is 10, expectation of Y plus 3 X is 25; then how can you find out expectation of 4 X minus Y?

(Refer Slide Time: 11:07)

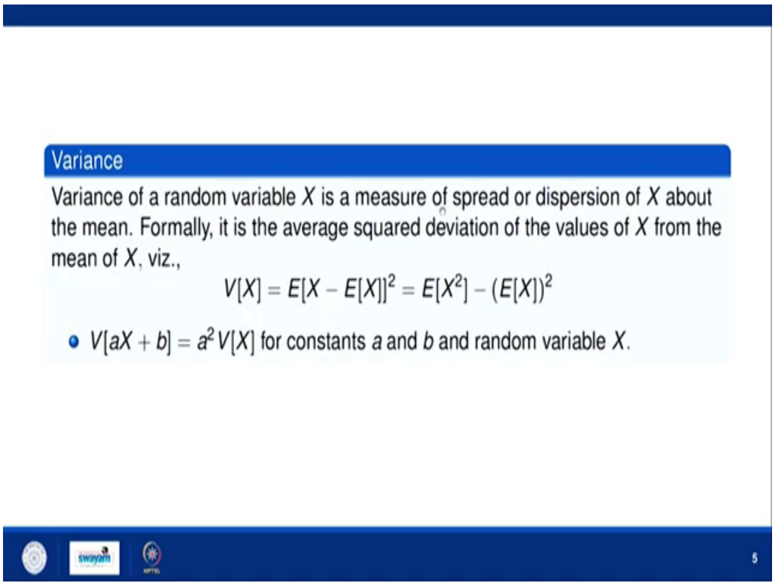


So, now in this problem expectation of 2 X is known to you; expectation of 2 X is what, is 10. So, this implies 2 times expectation of X is 10 as by the property of expectation of X; we know that expectation of a X plus b is nothing, but a times expectation of X plus d. So, this is further written as expectation of X is 5.

So, this is expectation of X. Now expectation of Y plus 3 X is 25. So, expectation of Y plus 3 X is 25. So, this implies expectation of Y plus 3 times expectation of X is 25. So, you can substitute expectation of X is 5 here. So, this implies expectation of Y plus 15 equal to 25. So, this implies expectation of Y equal to 10.

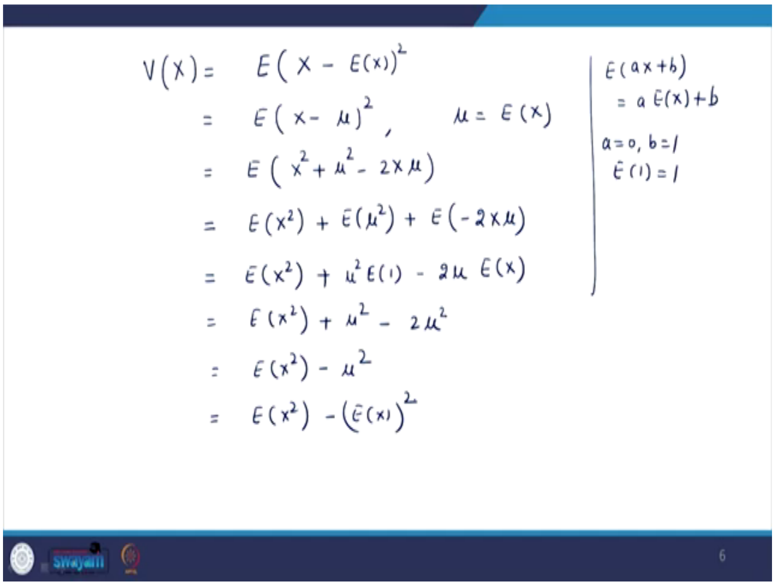
Now what you want to, what you want to find? We want to find, we want to find out expectation of 4 X minus Y. So, expectation of 4 X minus Y will be 4 times expectation of X minus expectation of Y. So, this will be 4 times; expectation of X is what, is 5 we have calculated here minus, expectation of Y is what, 10. So, it is 20 minus 10, that is 10. So, 10 is the final answer.

(Refer Slide Time: 12:44)



Now, let us come to variance. How can you find out variance of a random variable X? So, variance of a random variable X is a measure of a spread or dispersion of X about the mean. Mean is nothing, but expectation of X. So, if we want; if, we want to find out the how much random variable it disperse is spread about the mean that is, that measures by variance of X. And mathematically it is given by, it is given by expectation of X minus mean whole square; that means it is simply average squared deviation of the values of X from the mean, from the mean of X.

(Refer Slide Time: 13:35)



So, how can we define variance? Now you can see here, a variance of X is nothing, but expectation of X minus expectation of X whole square.

So that means, simply expectation of X minus mu whole square; if we denote mu by expectation of X. So, let us simplify this. So, it is a minus b whole square. So, it is a square plus b square minus 2 a b, ok. Now you can further simplify it, it is expectation by the linearity property; it will be expectation of X square plus expectation of mu square plus expectation of minus 2 X mu.

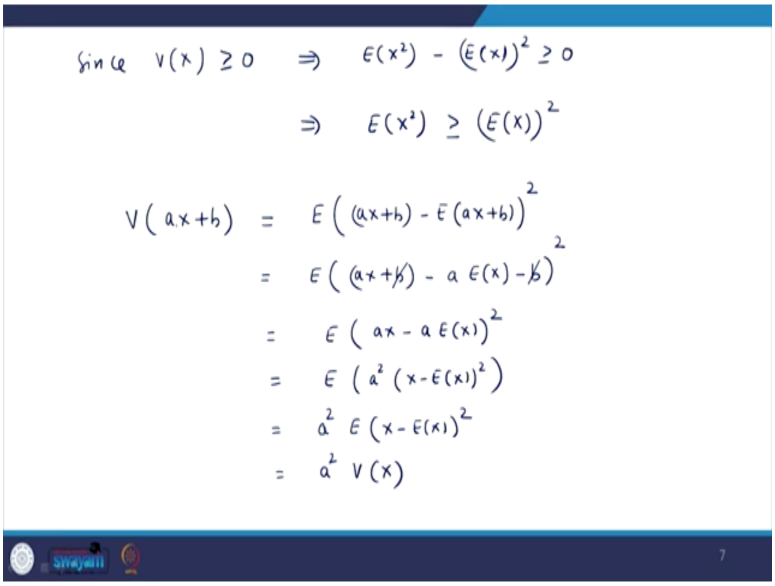
So, it is expectation of X square plus; now mu square, mu square is a fixed value, mean of a mean of a random variable is a fixed quantity. So, this will come out. So, this is simply mu squared times expectation of 1. And here minus 2 mu will come out and it will be nothing, but expectation of X; because here also minus 2 times mu is also a fixed quantity.

Now, what is the expectation of 1? Expectation of 1 is nothing, but 1; because if because, if you find expectation of 1, because you see expectation of a X plus b as we already derived, it is a times expectation of X plus b. If you take a as 0 and b as 1; so we get expectation of 1 equal to 1.

So, in this way it is expectation of X square plus mu square minus 2 mu. And what is the expectation of X here? It is mu; so it is into mu that is mu square. So, this will be expectation of X square minus mu square and that will be expectation of X square minus expectation of X whole square.

So, in this way we will we can get; if we want to find out variance of a X, that how much it is disperse about mean, then that can be measured by this expression or by this expression, the simplified one, that is variance of X is expectation of X square minus expectation of X whole square.

(Refer Slide Time: 16:09)



So, we can say that, since variance of X is always greater than equal to 0. So, that implies expectation of X square minus expectation of X whole square is always greater than equal to 0. So, from this we can conclude that expectation of X square is always greater than equal to expectation of X whole square.

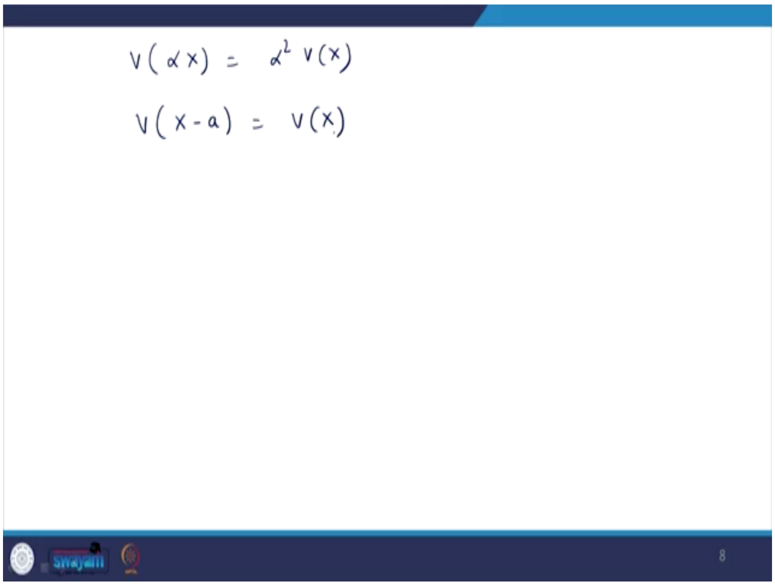
So, this is one result that, expectation of X square is always greater than or equal to expectation of X whole square. Now if we want to calculate variance of a X plus b, so how, for constants a and b. So, how we can find out variance of a X plus b? So, variance of a X plus b as if we apply the basic definition, so that is nothing, but expectation of, expectation of a X plus b minus expectation of a X plus b whole square, this is as per definition of variance of X.

Instead of X, now we are having a X plus b. So, you simply replace X by a X plus b. Now this is further equal to expectation of a X plus b. Now here you apply the linearity property of expectation of X, so that is a times expectation of X minus b. So, this b will cancel with this b that is; so that means that, further implies it is expectation of a X minus a times expectation of X whole square and that is further equal to expectation of a square into X minus expectation of X whole square.

Now, again this is a constant quantity, so it can, it can taken out by the linearity property of E, I mean expectation. So, that is a square times expectation of X minus expectation of X whole square. So, and this is nothing, but variance of X. So, we can say that, a square times variance of X. So, we can say that, if we, if we add or subtract a fixed number from each, from each random variable X; then variance, variance is unaffected.

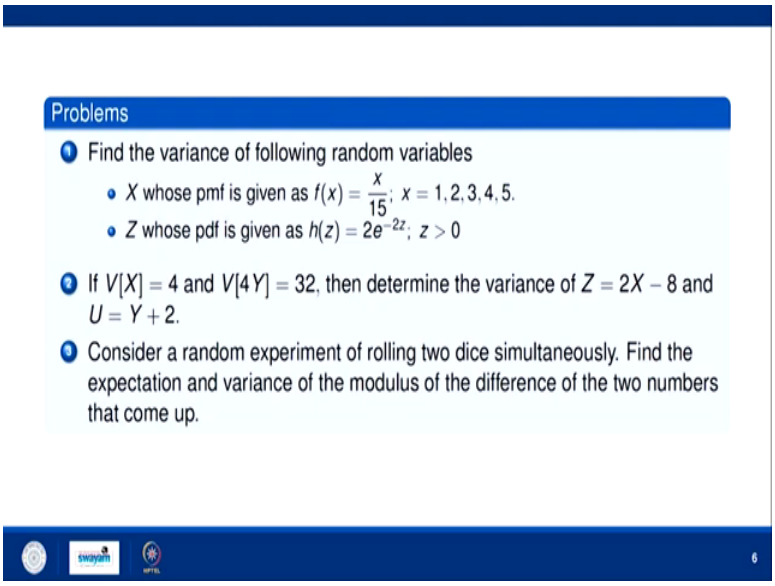
However if you multiply X by a certain number say, certain constant say a; then it will be nothing, but a square times variance of X.

(Refer Slide Time: 18:46)



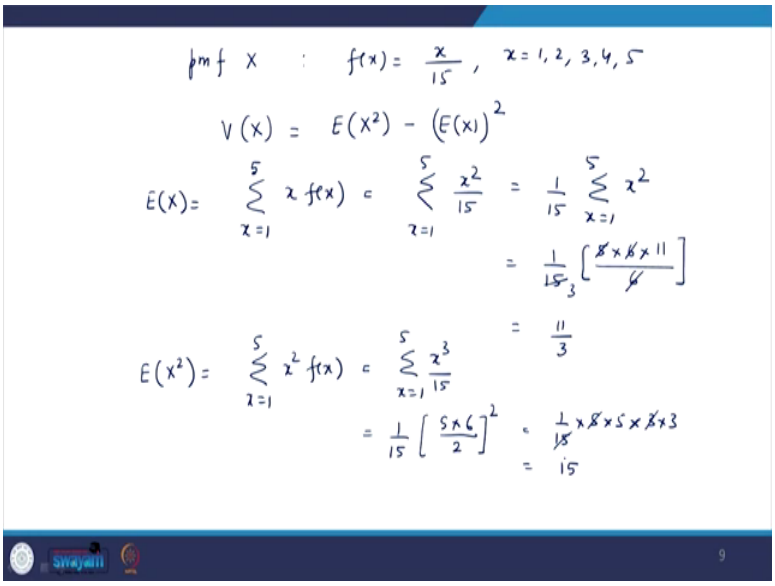
I want to say that, if you want to find out variance of alpha X, so that simply variance is alpha squared times variance of X; we want to compute variance of X minus a, so that nothing, but variance of X itself. So, that will not; I mean the, if you add or subtract something with a random variable, that will not affect the result on variance of X.

(Refer Slide Time: 19:12)



Now, let us discuss few problems based on variance of X, ok. Now suppose you find, we want to find out variance of the random variable X whose pmf is given by x by 15 when x is from 1 to 5, these are discrete case. So, how we can find variance of this?

(Refer Slide Time: 19:34)



So, here, here what is the pmf of X, that is simply given by x upon 15, where x is between 1, 2, 3, 4, 5.

Now, we want to compute variance of X. So, variance of X as per the formula, it is as per the expression, it is expectation of X square minus expectation of X whole square. So, we want to compute two things; first expectation of X square, and then expectation of X. So, let us compute expectation of X first. So, it is a discrete case. So, summation will come and x f x, and x is varying from again 1 to 5.

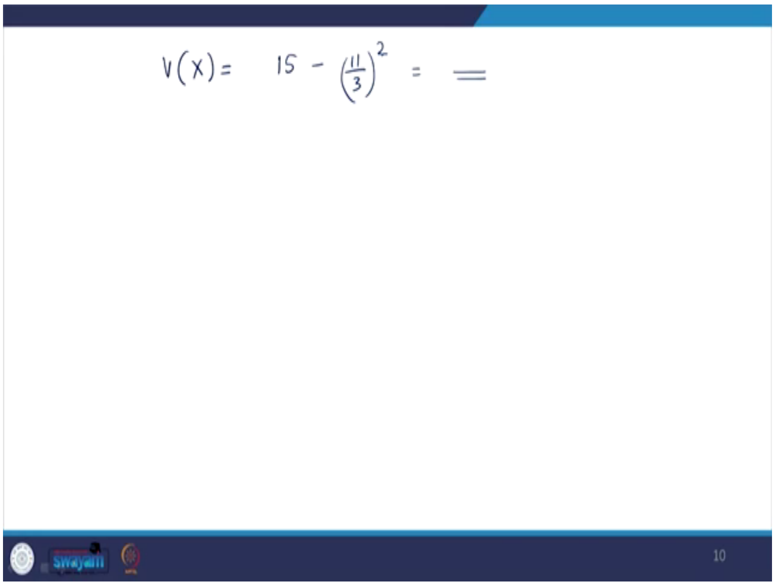
So, this is summation, now this is x square upon 15; you can substitute x as x upon 15, x is varying from 1 to 5. So, 1 upon 15 will come here and it is summation x varying from 1 to 5, x square. Now again it is, it is what? It is nothing, but sum squares of natural numbers up to 5.

So, as by the formula of sum of squares, it is 5 into 5 plus 1 that is 6 into 2 n plus 1 that is 11 upon 6 and n plus 1, 2 n plus 1 by 6. So, this cancels out, 5 into 3 times cancels out; so it is 11 upon 3.

Now let us compute expectation of X square. So, it is summation x square f x, x is varying from 1 to 5. So, it is nothing but summation, x is varying from 1 to 5, it is x cube upon 15. So, how can we compute x cube upon 15? So, it is 1 upon 15 times, sum of cubes of natural number up to 5; so that is n n plus 1 upon 2 whole square ok, you can find directly also. So, that is nothing, but it is 1 upon 15 into 5 into 5 into 3 into 3, ok.

So, that is this cancels out, so it is 15. So, what is the final answer now? So, we want to compute variance of X.

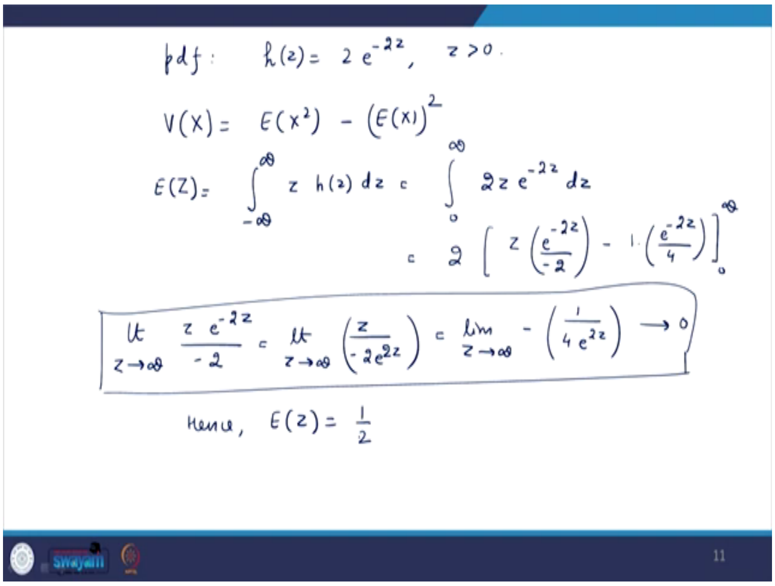
(Refer Slide Time: 22:08)



So, variance of X is now given by; first expectation of X square, from here is 15. So, it is 15 minus 11 upon 3 whole square. So, it is 15 minus 11 upon 3 whole square.

So, you can simplify this, you can simplify this and that will be the variance of this random variable X, ok. Now what is the, what is the variance of random variable Z whose pdf is given by this expression? So, now, let us compute for this continuous case, this is a continuous case.

(Refer Slide Time: 22:51)



So, now pdf is given by 2 e raise to power minus 2 z and z is greater than 0.

Now, for this, you want to compute variance of X. So, again very for variance of X, we have a simple expression that is expectation of X square minus expectation of X whole square. Now first expectation of X, expectation of X is now it is a continuous case, so it would be integral minus infinity to plus infinity. So, it will be z into x z, your h z d z. So, basically here instead of x, we are having z, anyway that is not a problem.

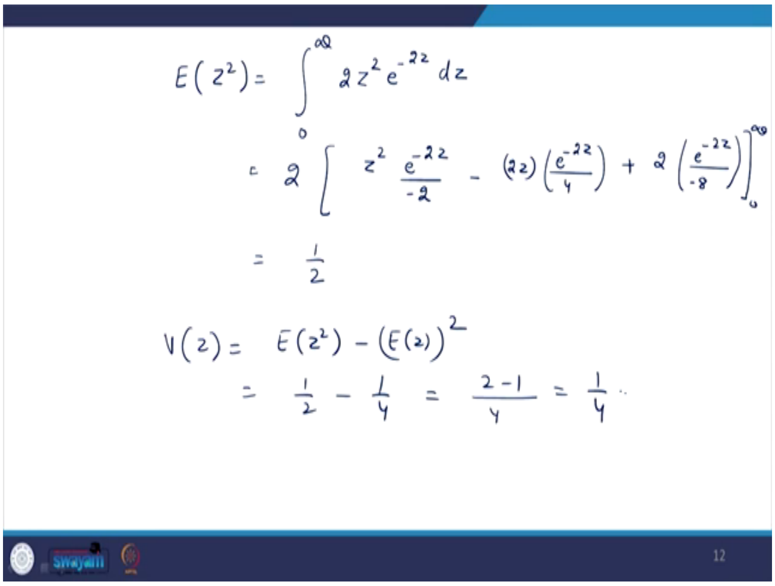
So, it is given by, now it is 0 to infinity; it is z 2 into z e raised to power minus 2 z d z, because we assume that otherwise it is 0. I mean only for z equal to greater than 0, it is defined like this; otherwise it is 0. So, now, it is 2 times; now we have to integrate this, so we have to apply by parts, integral by parts. So, when you integrate by part directly, so it is minus 2 z upon minus 2 and minus derivative of first and integral of second; that means, e raise to minus 2 z upon 4 and that is from 0 to infinity.

So, now, it is further equal to; now first, what will happen if z tending infinity here? So, let us try to compute limit z tending to infinity, z into e raise to power minus 2 z upon minus 2, it is equal to limit z tending to infinity z upon minus 2 e raise to power 2 z. As z tend to infinity, it will, it will tends to infinity upon infinity; you can take minus outside.

So, this will be nothing, but. So, when it is infinity upon infinity, you can apply L hospital rule. So, it is limit z tending to infinity minus the derivative of numerator, and derivative of denominator will come in the denominator; so it is 4 times e raised to power 2 z. And as z tend to infinity, this will tends to infinity and this will tends to 0, ok. Here also when z is tending to infinity, this will tends to 0. So, what is the final value of this?

So, hence expectation of Z will come out to be. So, when z is tending to infinity, both terms are tending to 0. So, from the upper limit it is 0; from the lower limit when z is tending to 0, it is 0, because z is here and it will be 1 upon 4. Negative, negative will be positive, so it is 1 upon 4 into 2 that is half. Now let us compute expectation of Z square, then only we can compute variance of Z.

(Refer Slide Time: 26:13)



So, what will be expectation of Z square? Expectation of Z square will be, now it is integral 0 to infinity, it is 2 z square; I mean here it is what, it is 2 e raise to power minus 2 z.

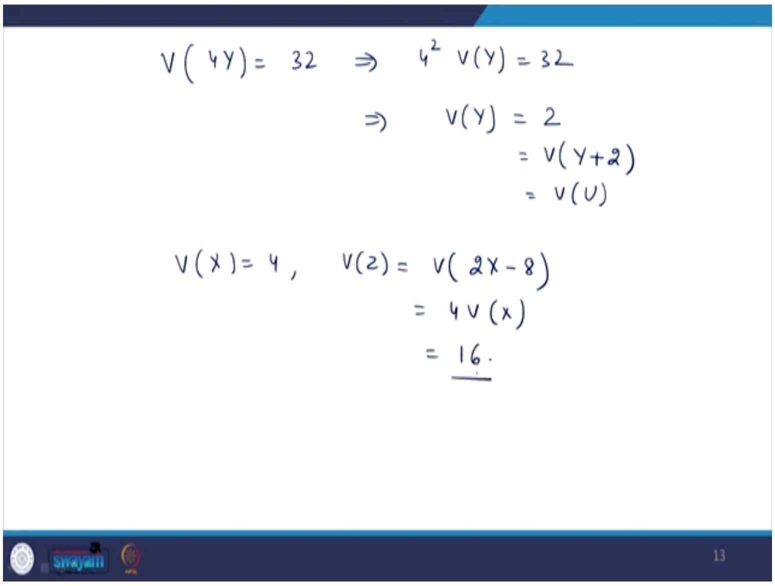
So, it is z square e raised to power minus 2 z d z, so it is 2. Now again you can integrate, it is z square e raised to power minus 2 z upon minus 2 minus derivative of first integral of second plus derivative of first next integral of second. So, basically you have to apply by part two times, you will get the same answer.

Now, again when z is tending to infinity, by the same processor as we did here, in the by the same processor this will also tends to 0. This will tends to 0 and this will also tends to 0; so as z tend to infinity, all the three terms are tending to 0. Now when z is tending to 0, it is 0, it is 0 and this will be 1 by 8; I mean minus 1 by 8.

So, when you are taking 4 times, so it is nothing, but 1 by 2 again. So, what will be variance of Z finally? So, variance of Z which is this expression; so this will be 1 by 2 minus 1 by 4 and that will be nothing, but 1 by 4. So, in this way we can find out variance of, variance of this pdf, variance of random variable Z for this pdf.

Now, the another problem is basically based on the property of variance of X. So, if variance of X is 4 and variance of 4 Y is 32; how can you find out variance of 2 X minus 8 or variance of Y plus 2? So, here a variance of Y plus 2 will be same as variance of Y; we know that, because addition or subtraction will not affect the value of the variance.

(Refer Slide Time: 28:38)

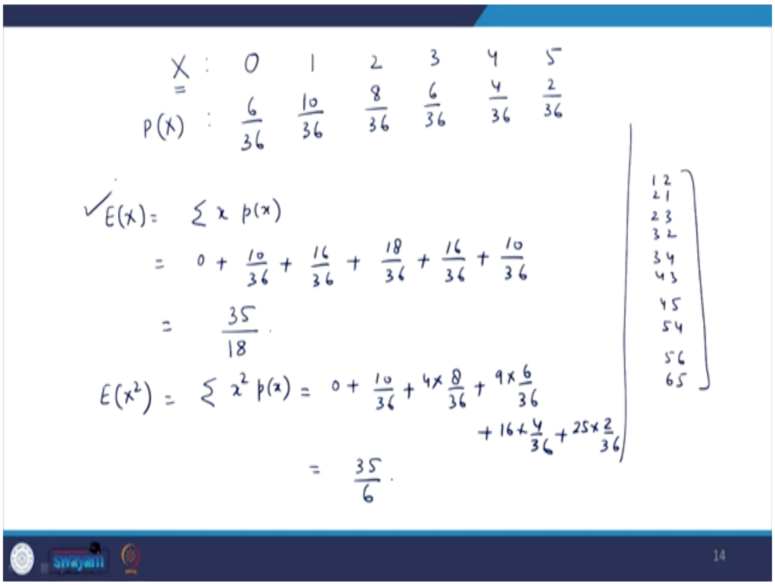


So, and what is variance of Y from here? See if variance of 4 Y is 32, if variance of 4 Y is 32; so this implies 4 square times variance of Y is 32. So, this implies variance of Y is it is 2, which is same as, which is same as variance of Y plus 2 that is variance of U; U or V, variance of U, yes, ok.

Now from the first one, variance of X is 4. Now, you want to compute variance of Z. So, if variance of X is 4 and till variance of Z, which is variance of 2 X minus 8 will be nothing, but 4 times variance of X which is 16. So, using the property of, simple property of variance we can find out these two values. Let us solve one more problem based on this; consider random experiment of rolling two dice simultaneously. So, you are rolling two dice simultaneously, and you want to find out expectation and variance of modulus of difference of two numbers that come up.

So, here random variable is basically, here the random variable is basically the modulus of difference of two numbers.

(Refer Slide Time: 30:05)



So, what will be X? So, here X denotes the modulus of difference of two numbers. Now the difference of when two dice are thrown, the difference, the minimum difference between the two dice will be 0, modulus; because if both numbers are same on the dice, then 1 minus 1 will be 0, 2 minus 2 will be 0.

So, the minimum difference is 0; it may be 1 also, suppose you are having 1 or 2 or 2 or 1 modulus is 1, it may be 2 also, 3, 4. What will be the, what will be the maximum difference? So, maximum difference will be, when one die has a minimum value and the other die has a maximum value. So, one die has a minimum number that is 1 and the other die has a maximum number that is 6. So, difference will be 5.

So, the maximum difference will be 5 only. So, this, this will be the different values with this random variable can take. Now what is the, what is the probability of this? Now when it is 0, the 0 it will be when it is 1 1, 2 2, 3 3, 4 4, 5 5, 6 6; that means 6 cases, 6 cases out of 36,.

Now, when it will be 1? It will be 1, when you are having 1 2 or 2 1, it will be having say 2 3, 3 2, it will be having 3 4 or 4 3, it will be 4 5 or 5 4; 1, 2, 3, 4, 5, 6, 7, 8. So, it will be having 5 6 or 6 5. So, these are the 6 cases, these are the 10 cases sorry, these are 10 cases when a difference is 1; you can see the modulus of difference is 1.

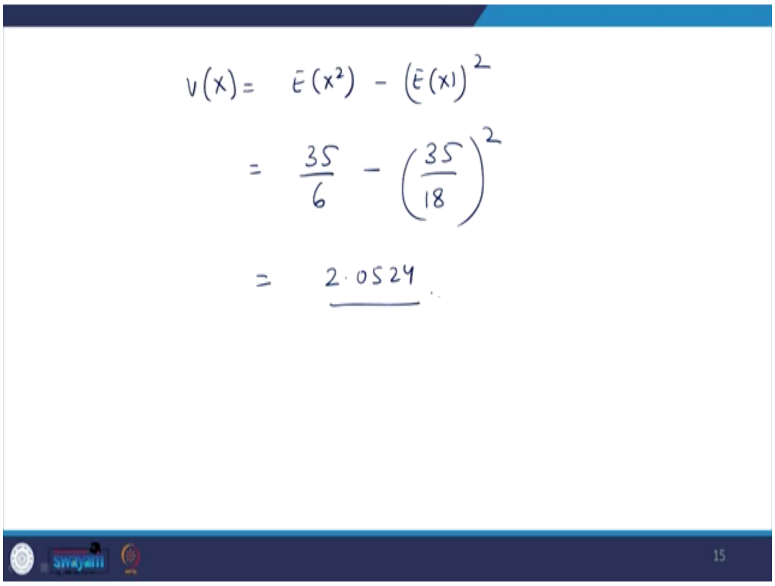
So, these are 10 cases. Similarly when difference is 2, so there will be 8 cases; you can check. Now it is 6 cases, it is 4 cases; these are 4 cases, and for 5 these are the two cases only 1 6 or 6 5. Now if you want to compute variance or X variance or expectation of action, mean of this random variable.

So, mean will be given by summation x p x. So, this will be 0 into this is 0 plus 10 by 36 plus 16 by 36 plus 18 by 36 plus 16 by 36 plus 10 by 36 and that will be nothing, but 35 by 18. If you simplify, so that will be nothing, but 35 by 18; you can simplify and you can find out,.

Now, what is expectation of X square? Expectation of x square is summation x square p x. So, that will be this into this is 0 plus 10 by 36 plus 4 into; you have to make square of this and then multiply with the respective probability, so that will be 8 by 36. Similarly 9 into 6 upon 36 plus 16 into 4 upon 36 plus 25 into 2 upon 36 and that simply will be given by 35 upon 6 that you can verify.

Now, what will be variance of X now?

(Refer Slide Time: 33:40)



So, variance of X will be. So, this is expectation and variance of X will be as per the formula it is expectation of X square minus expectation of X whole square. Expectation of X square is what? It is 35 by 6 minus expectation of X is what, 35 by 18.

So, this value will come out to be nearly it is 2.0524; so you can, you can simplify and see. So, in this way if you are having these type of problems, where you have to where some random variable is given to you and you have to find out expectation and variance, that also you can find.

So, basically in this lecture we have seen that if a random variable is given to you, maybe discrete or random; then how can we find out expectation that is mean or variance of that random variable.

In the next lecture we will see that, what are the various distribution functions and how we can, we can find out their mean or variance?

Thank you.