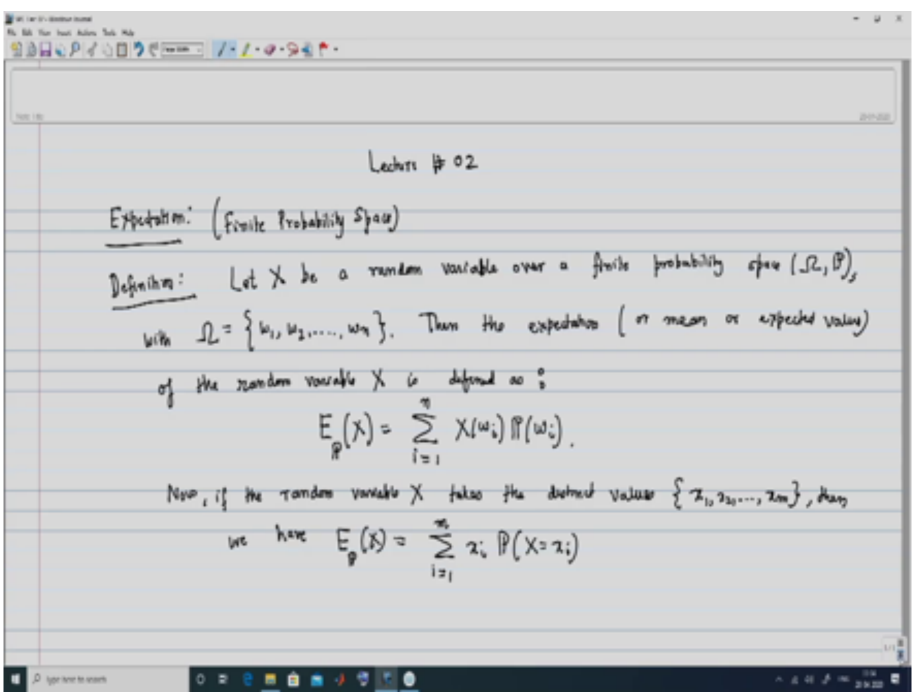
Mathematical Portfolio Theory

Prof. Siddhartha Pratim Chakrabarty
Department of Mathematics
Indian Institute of Technology Guwahati

Module 01: Basics of Probability Theory
Lecture 02: Mean, variance, covariance and their properties

Hello viewers, welcome to this lecture on the MOOCs course on Mathematical Portfolio Theory. In the last lecture, we had talked about probability space both discrete and continuous, we had defined the properties of a probability measure and we had talked about random variables and some of its properties. In todays class, we will look at the two moments, the first two moments in the case of random variables, namely, the mean and variance, and then we will talk about covariances and correlation coefficients. These two moments are of great importance in the case of portfolio theory, because the entire structure of the modern portfolio theory hinges on the mean which will be then related to the expected return and the standard deviation or variance which will be related to the risk in market conditions.

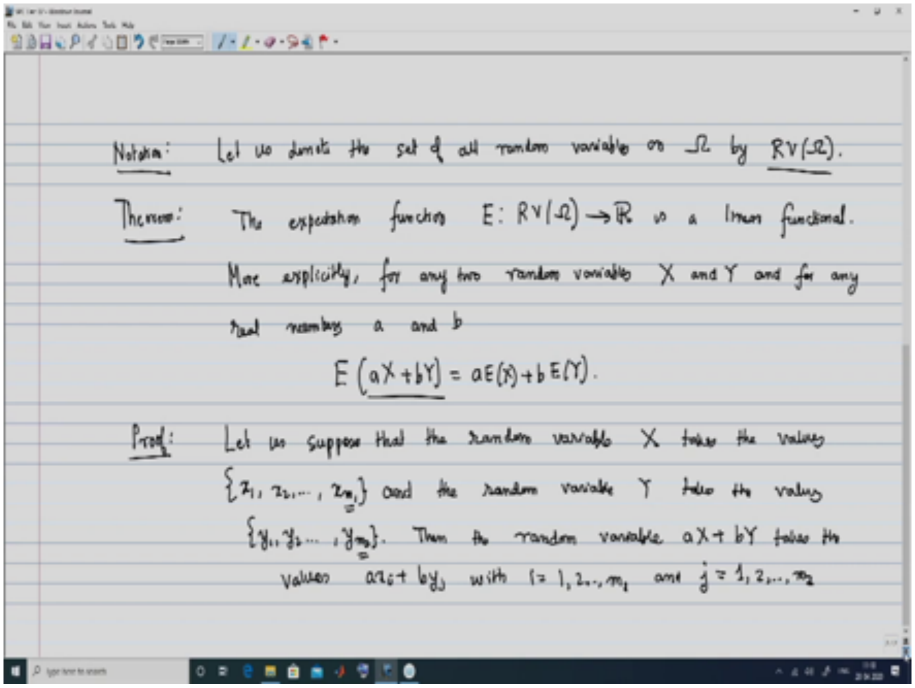
(Refer Slide Time: 01:31)



So, we start this lecture number 2 by first talking about expectation in discrete time. So, first we talk about a expectation and we will talk in the context of a finite probability space and we will first begin naturally with the definition. So, what expectation we are talking about? We are talking about the expectation of a random variable. So, let X be a random variable as already defined over a finite probability space (Ω, P). Remember the Ω was the sample space and P was the probability measure. With Ω now, since this is a finite probability space, so I take my Ω to be comprised of some elementary events $\omega_1, \omega_2, \cdots, \omega_n$.

Then the expectation or sometimes we call it a mean or even expected value of the random variable that is the random variable X is defined as $E_P(X) = \sum_{i=1}^{n} X(\omega_i)P(\omega_i)$. Now, if the random variable X it takes a certain number of distinct values. So, suppose it takes the distinct values and I will enumerate them as $\{x_1, x_2, \cdots, x_m\}$. In that case, we have the expectation $E_P(X) = \sum_{i=1}^{m} x_i P(X = x_i)$.

(Refer Slide Time: 04:53)

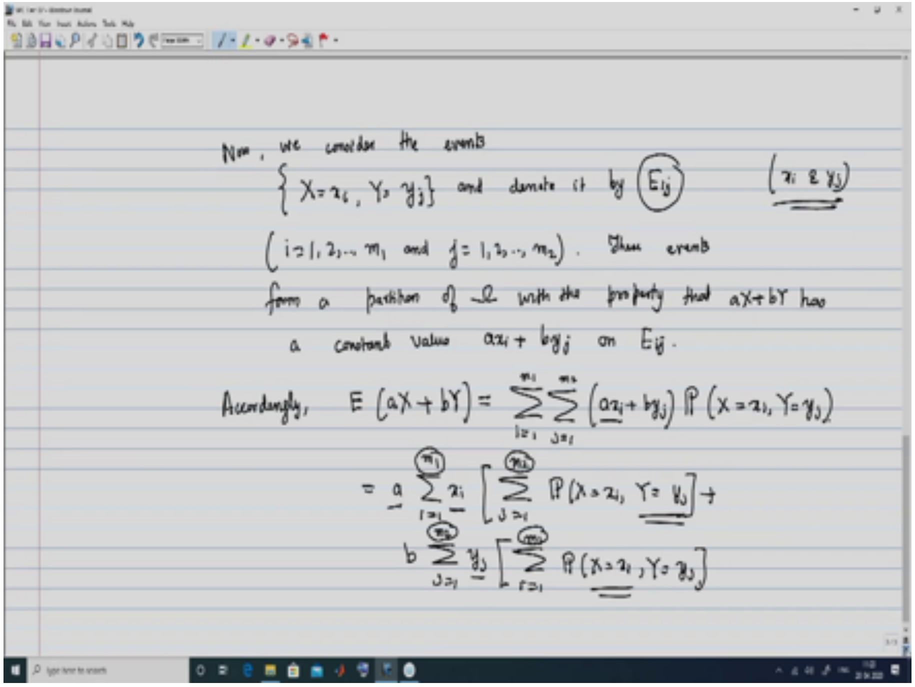


So, let us now introduce a certain notation. So, we say that so let us denote the set of all random variables on the sample space omega with the notation $RV(\Omega)$. Now this is followed by a theorem and we will see you know why we introduce this notation $RV(\Omega)$. So, the expectation function, alright. So, we will basically now talk about the expectation function $E : RV(\Omega) \to R$ is a linear functional. So, I will explain what do we mean by linear functional. So, I will say that more explicitly if we have any two random variables, so for any two random variables which are you know customarily, we choose them to be X and Y and for any real numbers which will customarily choose as a and b. We have the following result that $E(aX + bY) = aE(X) + bE(Y)$, okay. Now let us look at a proof of this a simple proof of this. So, accordingly, we have to start off with the random variable X and the values it takes. So, let us suppose that the random variable X takes the values, say, $x_1, x_2, \cdots, x_{m_1}$ and the random variable Y, similarly it takes the values say $y_1, y_2, \cdots, y_{m_2}$. Then the random variable so we are interested in the random variable $aX + bY$, what does it look like? So, the random variable remember that any linear combination of random variable is also a random variable. So, this random variable $aX + bY$ this will take the values of what form? It is going to take the values of the form $ax_i + by_j$ with $i = 1, 2, \cdots, m_1$ and $j = 1, 2, \cdots, m_2$.

(Refer Slide Time: 08:59)

So, what we are going to do now is we will next consider the event. So now, we will consider the events and the events are what the events are that $\{X = x_i, Y = y_j\}$, we will consider this event and we will introduce a notation for it. So, we denote it by say E_{ij}. So, basically all combinations of this x_i and y_j, all these such events will identify them by E_{ij}, see as before $i = 1, 2, \cdots, m_1$ and $j = 1, 2, \cdots, m_2$. Now this events as you can see, so the corresponding this events form a partition of Ω. So, this is in a very crucial with the property, so consequent to this being a partition of omega it satisfies the property that $aX + bY$ has a constant value $ax_i + by_j$ on E_{ij}. So, accordingly, so what we need to prove is basically we will need to prove this result that

$$E(aX + bY) = aE(X) + bE(Y).$$



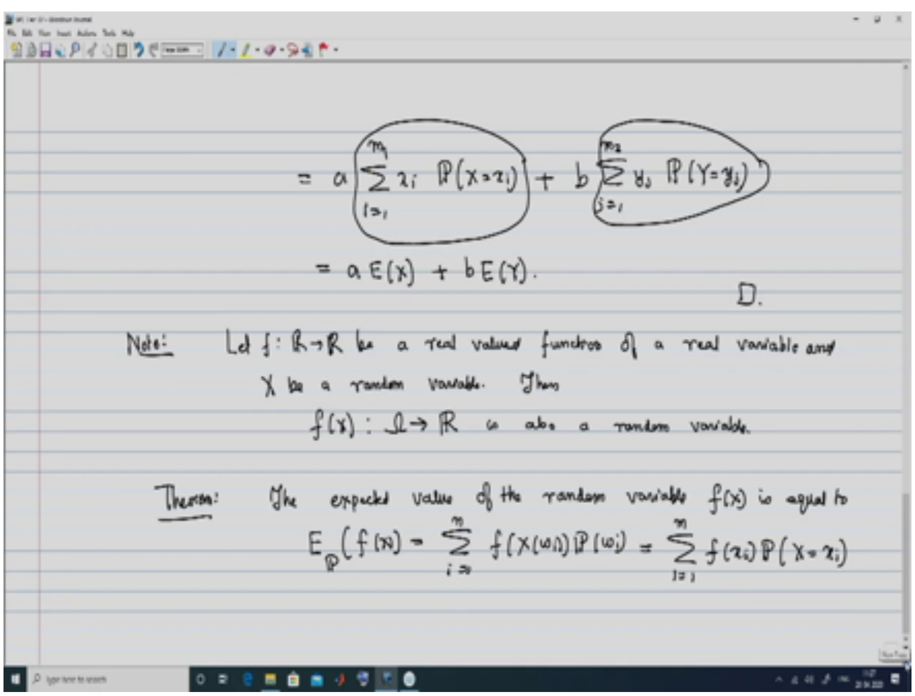
So, accordingly,

$$E(aX + bY) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} (ax_i + by_j) P(X = x_i,\ Y = y_j),$$

this is going to be by definition this is going to be

$$a \sum_{i=1}^{m_1} x_i \left[\sum_{j=1}^{m_2} P(X = x_i,\ Y = y_j) \right] + b \sum_{j=1}^{m_2} y_j \left[\sum_{i=1}^{m_i} P(X = x_i,\ Y = y_j) \right].$$

(Refer Slide Time: 13:34)



So, accordingly, what we have in the next line is that I will have a. So, from here I will have

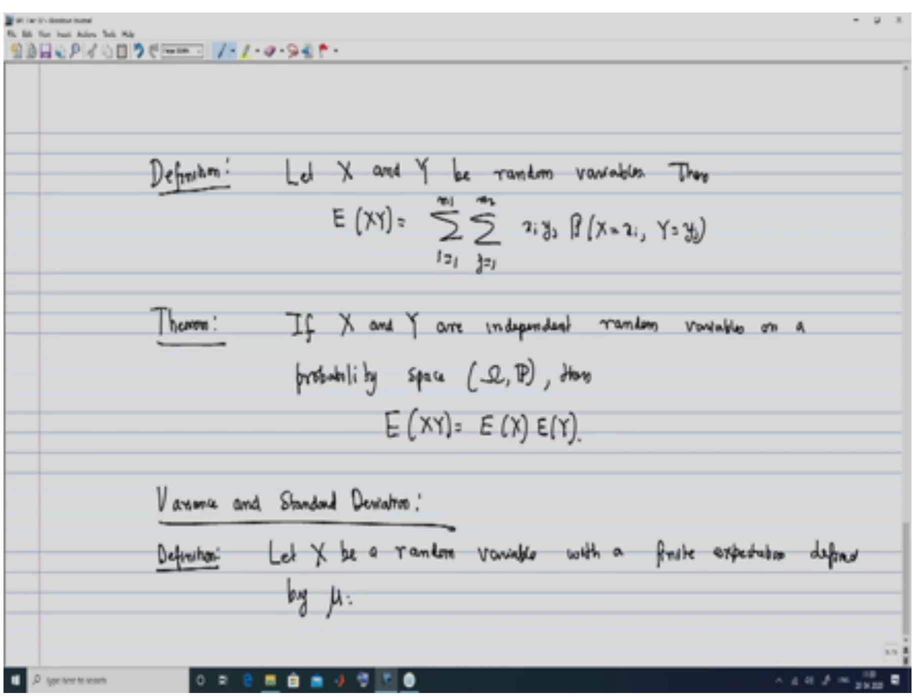
$$a\sum_{i=1}^{m_1} x_i P(X = x_i).$$

Note that here the summation only runs over $j = 1, \cdots, m_2$. So, I am only left with $P(X = x_i)$. In a similar manner, I will have this next term

$$b\sum_{j=1}^{m_2} y_j P(Y = y_i).$$

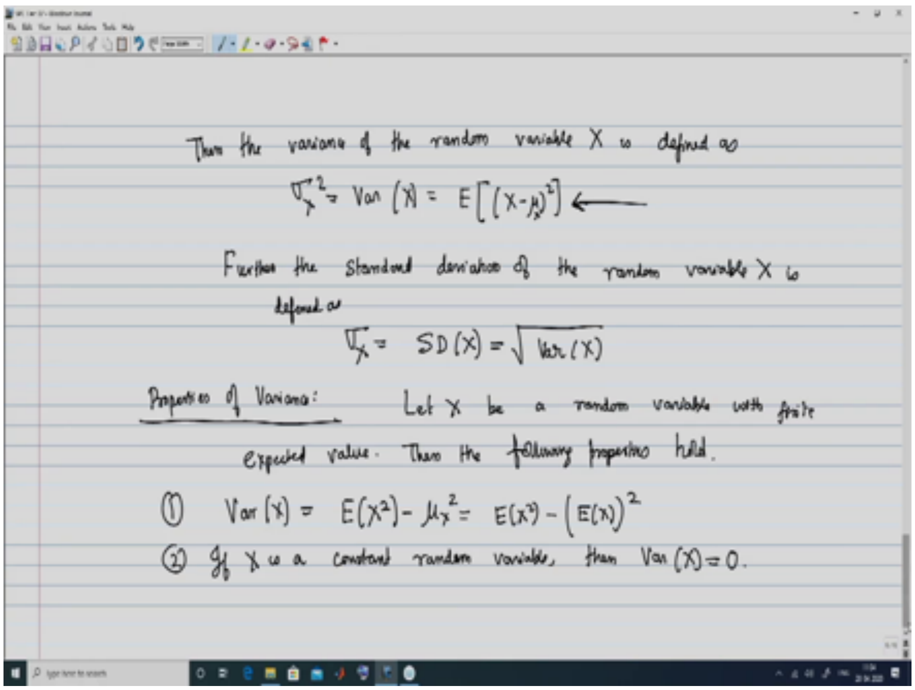
So, these adds up to 1 and so I am only left with probability of $Y = y_j$. So, this is probability of $Y = y_j$. So, now observe carefully this term here, this is nothing but by definition the expectation of $aE(X)$. So, we obtained $aE(X)$ and this term by definition is the expectation of y, so I have $bE(Y)$. So, this completes the proof for the linearity of expectation alright. So, next what we do is we will just make an observation. So, note that so let $f : R \to R$ be a real valued. So, this is R. So, this be a real valued function of a real variable, remember this is from R to R and let X be a random variable. So, then I can now make the observation, then $f(X)$ which will be from $\Omega \to R$, since X is from $\Omega \to R$. So, $f(X)$ will be from $\Omega \to R$ is also a random variable. So, accordingly once I have so earlier I was talking about the random variable X or Y and I talked about the expectation of the random variable X and that of Y and now that I have defined what this that if X is a random variable and $f : R \to R$, then $f(X)$ is also a random variable. So, the natural thing to do now is to define what is going to be the expectation of this newly defined random variable namely $f(X)$. So, this brings us to the next theorem or you can treat it as a definition if you want. So, then the expected value of this newly noted random variable $f(X)$ is equal to. So, this is going to be expectation of $f_P(X)$ as before, this is going to be we take $f(X(\omega_i))P(\omega_i)$ and we run this summation from $i = 1$ to n. And as before if we takes the values $x_1, \cdots, x_m$, then this is simply going to be $f(x_i)P(X = x_i)$ and this time $X = xi$ and this time i will go from 1 to m. Remember that we had taken the random variable $X = x_1, \cdots, x_m$.

(Refer Slide Time: 17:51)



So now, what you want to do is that we have talked about the addition of random variables and the function. So, this brings me to the next definition which is the expectation of the product of random variables. So, let X and Y be random variables, then the expected value of XY, this is defined to be. So, for XY the random variable takes the value naturally x_iy_j with the probability their probability $X = x_i$ and $Y = y_j$, where $i = 1, \cdots, m_1$ as before and $j = 1, \cdots, m_2$. So, the immediate consequence of this definition is the theorem pertaining to independent X and Y. So, if X and Y are independent random variables on a probability space (Ω, P), then $E(XY) = E(X)E(Y)$. So, once we have the definition of expectation in terms of a variable and as well as its linear combination and we have talked about the expectation of the product of two random variables and what happens in the case of both the random variables being independent of each other. So, naturally the next thing that we need to look at is the variance. So, accordingly we start then concept of variance and standard deviation. So, the first thing we look at is we look at the definition. So, again if the variance is for the random variable, so let X be a random variable and since we have already defined expectation. So, we will put the condition that it is this has a finite expectation defined by say μ.

(Refer Slide Time: 20:44)



Then the variance of the random variable X is defined as

$$\sigma_X^2 = Var(X) = E[(X - \mu)^2].$$

So, actually let us go back let us define this by μ_X. So, that there is no ambiguity. So, I will just define this as μ_X, just to identify that this is the finite mean of the random variable X. So, once we have the definition of variance. So, further the standard deviation. So, the immediate fall out of it is the definition of standard deviation. So, the standard deviation of the random variable X is defined as

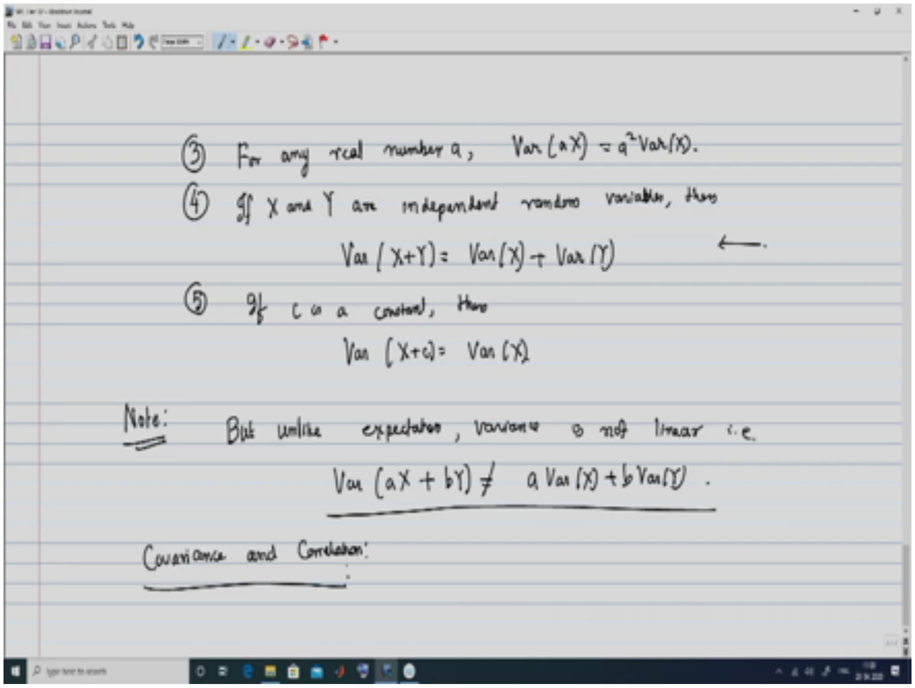
$$\sigma_X = SD(X) = \sqrt{Var(X)}.$$

So, now that we have defined what is now that we have defined what is expectation and variance. So, the next thing that we will do is that and we have looked at a couple of properties of the expectation. So, the next thing that we look at is the certain properties of a variance. So, accordingly we start off with this notion of properties of a variance. So, formally let us I would say that let X be a random variable with finite expected value, then the following properties hold. So, let me enumerate the properties one by one. So, a $Var(X)$ and we look at this definition of $Var(X)$, this can be shown that this reduces. So, the expression on the right hand side here this reduces to

$$E(X^2) - \mu_X^2 = E(X^2) - (E(X))^2.$$

The second property is that if X is a constant random variable, then the variance of this constant random variable $Var(X)$, this is going to be equal to 0.

(Refer Slide Time: 24:18)



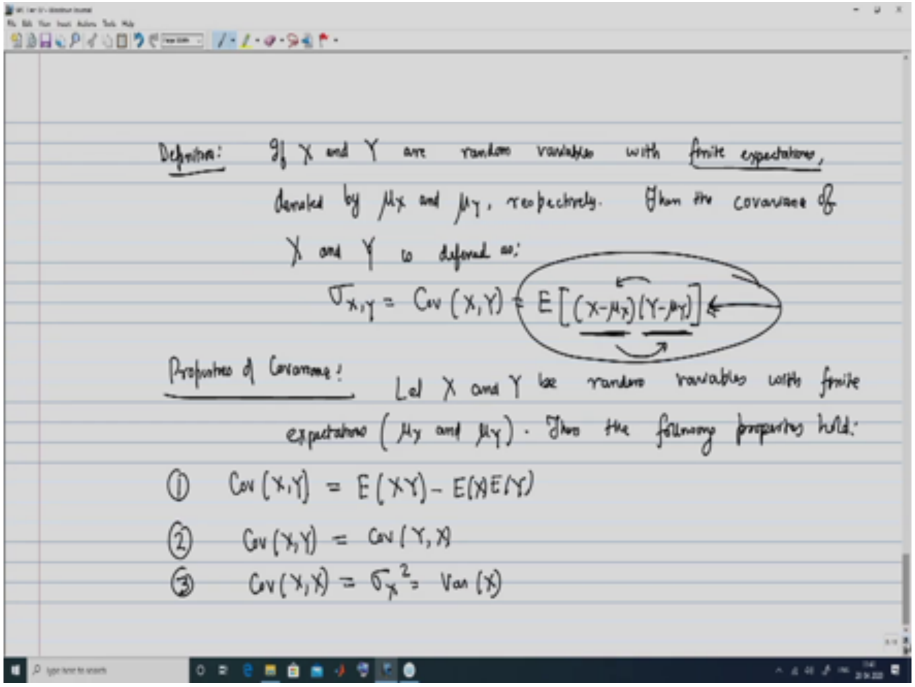
The third property pertains to the scaling. So, for any real number a, $Var(aX) = a^2 Var(X)$. The fourth property this is about the additivity. So, if X and Y are independent random variables and this is very crucial, these are independent random variables. Then $Var(X + Y) = Var(X) + Var(Y)$. And finally if c is a constant, then the variance of its scaling, that is, $Var(X + c)$ is simply going to be $Var(X)$. So, I leave the proof as an exercise. Just make use of the definition in order to prove them. Now I just want to revisit something in the context of this fourth property and make an observation of how variance is distinguished from expectation in terms of one property. So, unlike expectation a variance is not linear, alright. So, that means that, if it was linear, it would have satisfied the property that

$$Var(aX + bY) = aVar(X) + bVar(Y).$$

But this is not necessarily the case as we will see later on for the expression of a linear combination of random variables. So, once you will we present the expression for the variance of a linear combination of a random variables, which are not necessarily independent of each other, then you can easily see that this property actually does not hold. So, I just now mentioned that we look at the variance of a linear combination of the random variables, say $x_1, \cdots, x_n$. However, before we proceed on to do that we have to talk about one more concept that will be required in order to have an expression for the variance of the linear combination of all these random variables and that is basically the covariance. So, covariance is also as we will see later on in the discussion of the modern portfolio theory it is of a great importance. Where it will be related to essentially the joint behavior of the returns of the different assets of, for example, stocks that will constitute a portfolio. But we will discuss the details of that as and when we start talking about modern portfolio theory. So, coming back to our current discussion, let us now move on to what is the definition of covariance. So, we will talk about covariance and one close related concept to covariance, namely, correlation.

(Refer Slide Time: 27:55)

So, first let us start off with the definition. So, here if X and Y are random variables with finite expectations denoted by, so, we will denote them by μX and the expectation of Y to be μY and both of them are



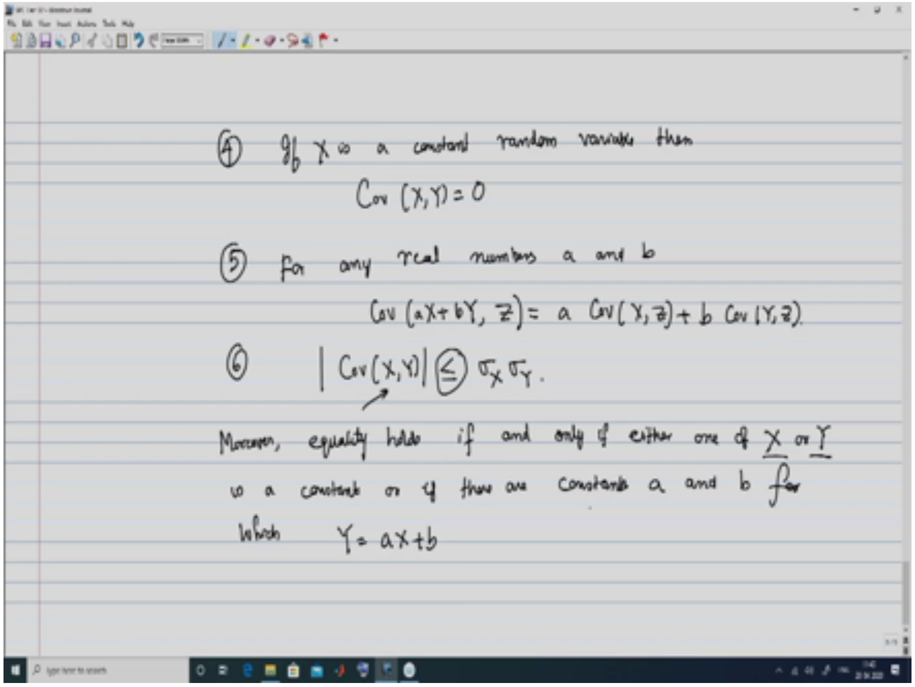
finite. So, we denote them by μX and μY respectively. Then the covariance of X and Y is defined as, so, the notation for this is

$$\sigma_{X,Y} = Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)].$$

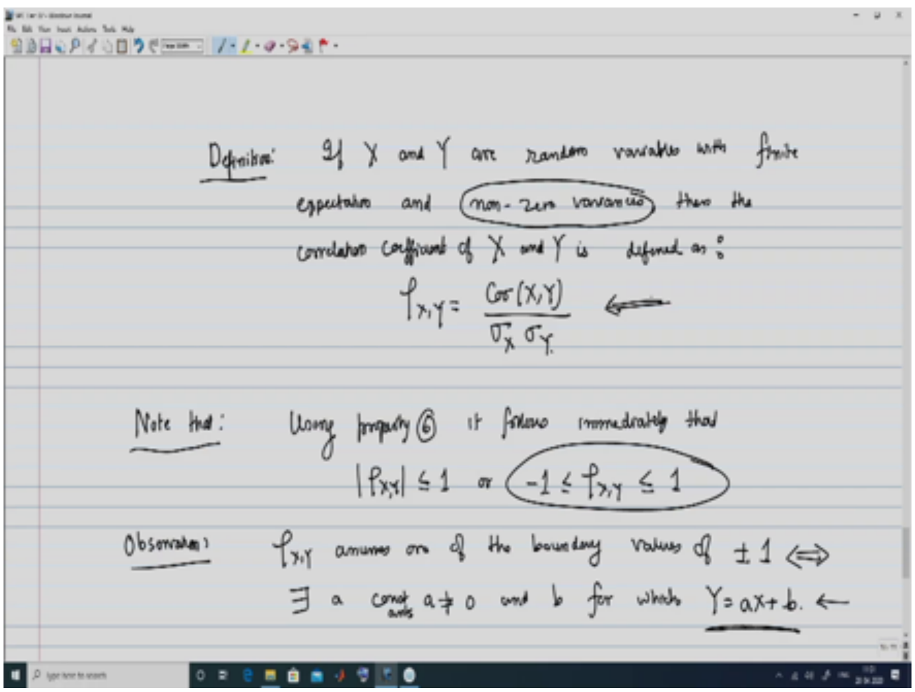
So, that is the expectation of the random variable $X - \mu X$ and $Y - \mu Y$, where μX and μY are both finite expectations. So, let us look quickly have a look at the properties of covariance. So, let X and Y be random variables with finite expectations as before and naturally this is a as before we will denote them by μX and μY, then the following properties hold. So, first property will be the definition of covariance of X and Y and in terms of this value of the expectation and we will give an alternative way of representing this. So, this will turn out to be equal to, so, this expression that we have here this will turn out to be $E(XY) - E(X)E(Y)$. The second property is that the $Cov(X,Y)$ is the same as c$Cov(Y,X)$. So, there is a evident from the fact that $Cov(X,Y)$ is the covariance of this product of these two random variables and covariance of Y is again the product of these two random variables with the positioning of $Y - \mu_Y$ and $X - \mu_X$ being simply exchanged. The third property, if you observe carefully if I take my $X = Y$, then this will simply become $E(X) - \mu_X^2$, which is the $Var(X)$. So, simply $Cov(X,X) = \sigma_X^2 = Var(X)$.

(Refer Slide Time: 31:29)

So, the fourth property is that if X is a constant random variable, then $Cov(X,Y) = 0$. So, this is again evident from the basic definition of a covariance, in which case in this case one of the expectations. So, suppose here the X is a constant random variable. So, expected value of X, if X is the constant value c, then the expected value of X is also going to be equal to c. So, one of the factors in the definition of the covariance, that is, $X - \mu(X) = 0$, which will render the covariance to be equal to 0. So, the next come to the fifth property. So, for any real numbers a and b, $Cov(aX + bY)$, this random variable with the random variable, say, Z this is going to be $aCov(X,Z) + bCov(Y,Z)$. So, this will be used when you later on look at the portfolio theory and talk about the single index model. Finally, one property is that the absolute value of $Cov(X,Y)$ this is going to be less than or equal to $SD(X)SD(Y)$. Moreover, just one more observation, so here this is the inequality. So, I need to sort of look at what happens or under what circumstance the equality will take place. So moreover the equality in this above relation. So the equality holds if and only if either one of X or Y. So, one of the random variables is a constant or if there are constants a and b, for which $Y = aX + b$. So, in the first case if either of them are constant then both the sides are going to be equal to 0 and the equality will hold or the other circumstance in which the equality can hold is one of the random variables is a linear combination of the other in terms of this constants a and b, okay.



(Refer Slide Time: 34:32)

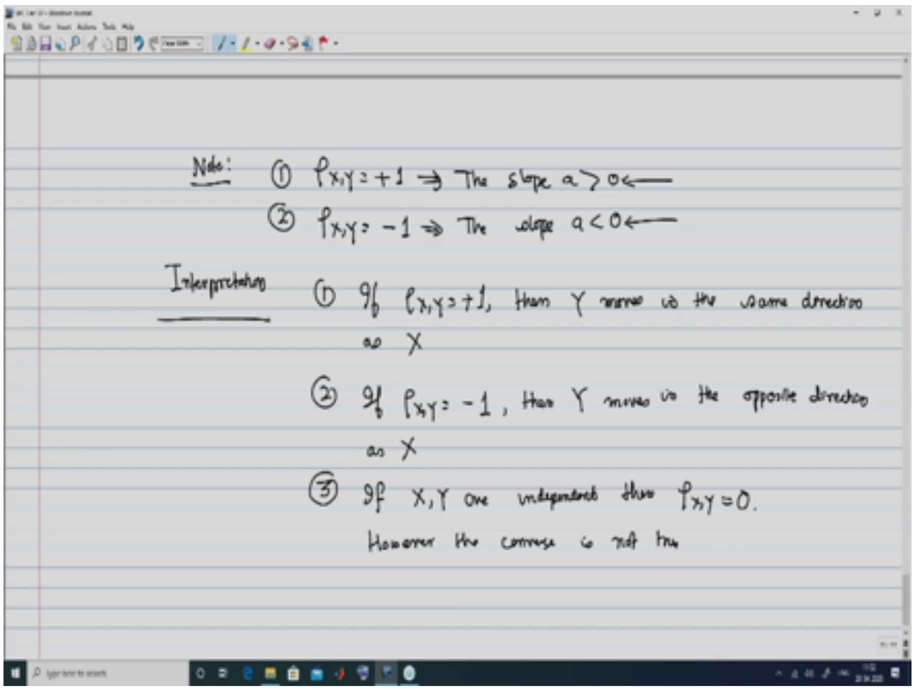


Now, once we have talked about the properties of covariance, the next thing that we look at is we will introduce the definition of correlation. So, again if X and Y are random variables with finite expectation and nonzero variances and we will see once the definition is placed as to why we need nonzero variances. Then the correlation coefficient of X and Y is defined as

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}.$$

So, this is the because we have $\sigma_X \sigma_Y$ in the denominator that is the reason why we needed to have a nonzero variances. So, we note that using property 6. So, remember the property 6 was that the absolute value of $Cov(X, Y) \leq \sigma_X \sigma_Y$. So, if I bring this on the left hand side from there it follows immediately in the context of $\rho(X,Y)$, it follows immediately that absolute value of $\rho(X,Y) \leq 1$ or this means that $-1 \leq \rho(X,Y) \leq 1$. So, this is an important property. So, next we make one observation in the context of the fact that $-1 \leq rho(X_1) \leq 1$. So, $\rho(X,Y)$, this assumes one of the boundary values of ± 1, if and only if there exist a constant a not equal to 0 and b so there are actually two constants, for which $Y = aX + b$. So, from this observation and the definition of ρ that we have here, it couple of things immediately follows.

(Refer Slide Time: 37:52)



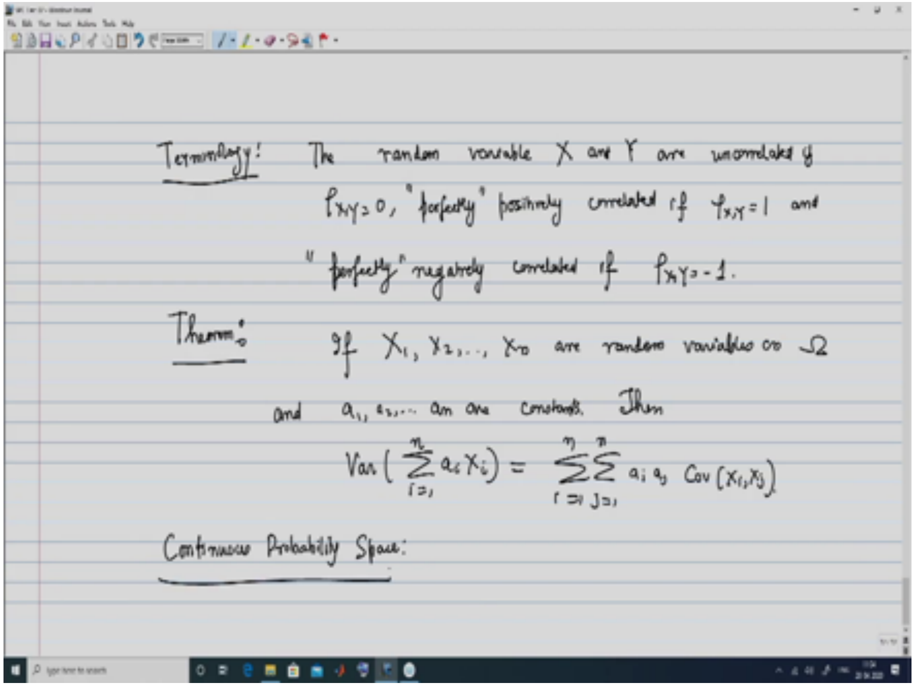
Immediately it follows that $\rho(X,Y) = 1$, this implies that the slope of this line that is $a > 0$ and for the other extreme value of $\rho(X,Y) = -1$ it follows that the slope $a < 0$. So, let us now look at the interpretation of this. So, if $\rho(X,Y) = 1$. So, this means that $Y = aX + b$ with a positive. So, then it means so I am making this inference from the fact that $Y = aX + b$ with a being positive in this case of $\rho(X,Y) = 1$. So, then I can conclude that then Y moves in the same direction as X and likewise if $\rho(X,Y) = -1$. So, again I look at the linear relation and take into account the fact that the slope $a < 0$. So, then Y moves in the opposite direction as X. So, one last interpretation remains and this is again in terms of independence. So, if X and Y are independent then $\rho(X,Y) = 0$. So, it is obvious from the fact that if they are independent of each other then $Cov(X,Y) = E(X)$. So, remember that we had this property of $Cov(X,Y) = E(XY) - E(X)E(Y)$. So, if X and Y are independent, then $E(XY) = E(X)E(Y)$. So, covariance will become 0 and consequently naturally your $\rho = 0$. However, we need to be cautious that the converse is not necessarily true, okay.

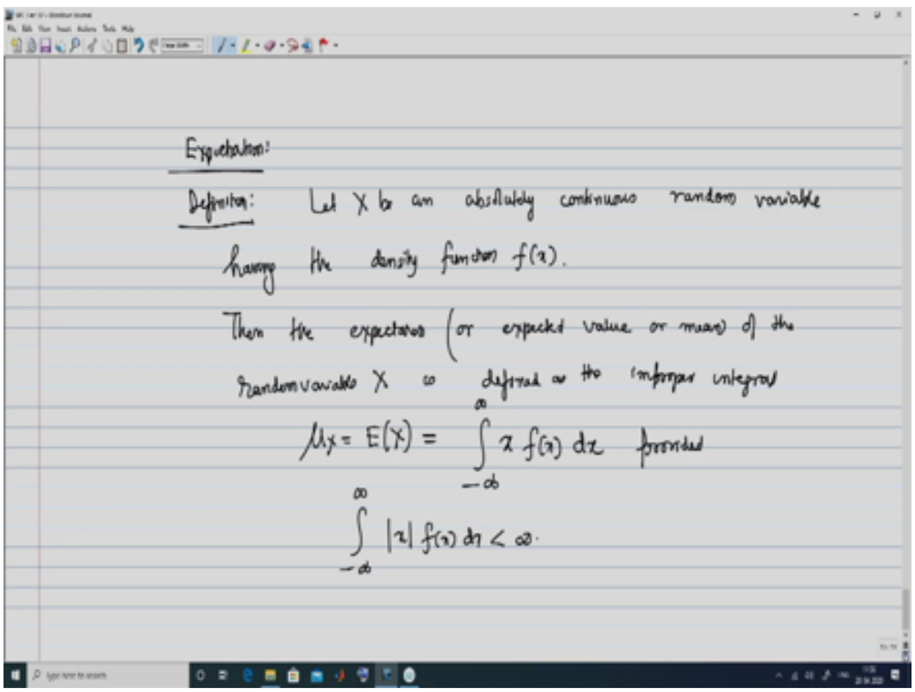
(Refer Slide Time: 40:35)

Now, let me note down a couple of terminology that will frequently use. So, the random variables so again this is related to the values of ± 1 and 0 of $\rho(X,Y)$. So, the random variable X and Y are uncorrelated if $\rho(X,Y) = 0$, perfectly positively correlated if $\rho(X,Y) = 1$ and perfectly negatively correlated if $\rho(X,Y) = -1$, okay. So, we conclude this discussion on covariance and correlation with one theorem related to variance. So, linear combination of random variables. So, if $X_1, X_2, \cdots, X_n$ are random variables on omega and $a_1, a_2, \cdots, a_n$ are naturally the corresponding constants. Then the linear combination that is $\sum_{i=1}^{n} a_i X_i$. The formula for this is given by double $\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)$. Now, next so we have discussed elaborately on the discrete space. So, let us now move on to continuous probability space and talk about the expectation and the variance in the continuous probability space.

(Refer Slide Time: 43:08)

As before we start off with expectation and we have the definition first. So, let X be an absolutely





continuous random variable having the density function. So, we have to specify the density function $f(x)$. So, then the expectation or expected value or mean of the random variable X is defined as the improper integral. So, as before I will use

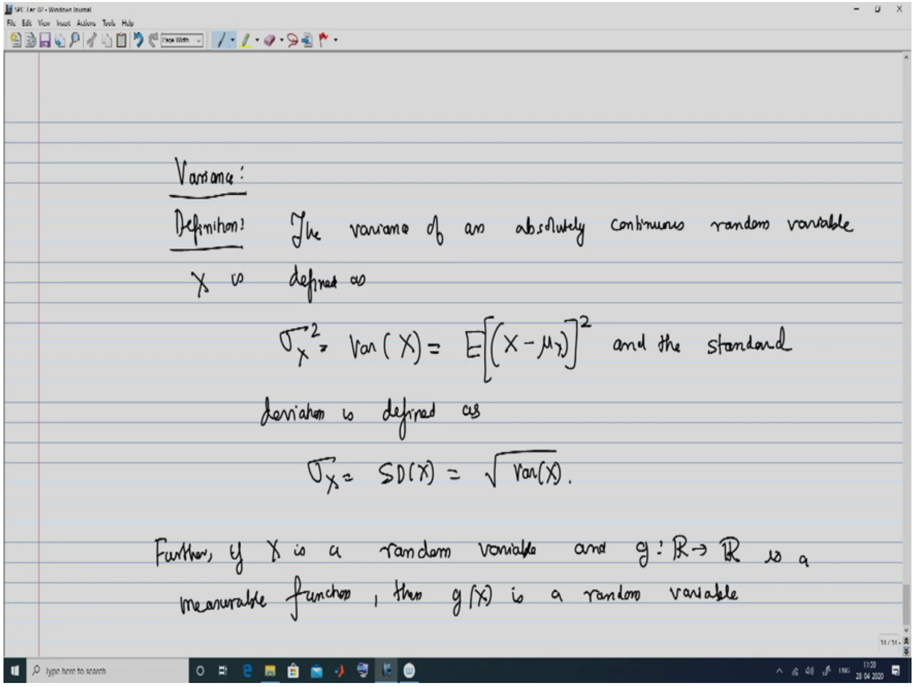
$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

. So, this is similar to $\sum X_i P(X = x_i)$. However, we need the condition provided that

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

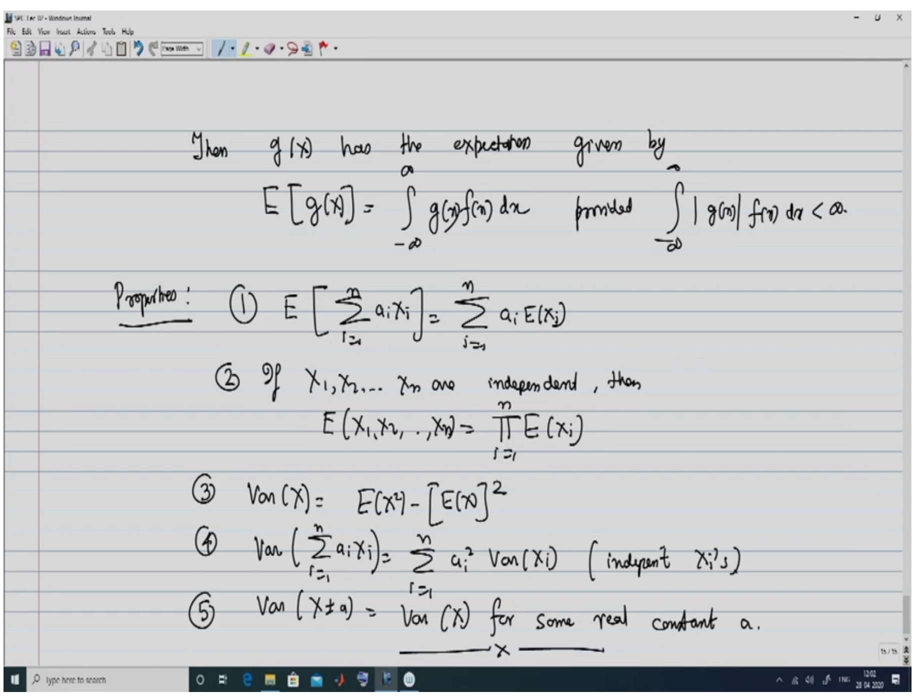
and finally we come to variance.

(Refer Slide Time: 44:59)



So, for this so let us look at the definition the variance of an absolutely a continuous random variable X is defined as was the case with the finite probability space. So, this is $\sigma_X^2 = Var(X) = E(X - \mu_X)^2$ and consequently the standard deviation is defined as $\sigma_X = SD(X) = \sqrt{X}$. Further, if X is a random variable and say $g : R \to R$ is a measurable function, then $g(X)$ is a random variable.

(Refer Slide Time: 46:52)



So, recall that we had something similar in the finite space, then $g(X)$ has the expectation given by and the notation for this would be

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx,$$

provided

$$\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty.$$

So, just to wind up this, we will just briefly note the properties both for expectation as well as variance. So,

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E(X_i).$$

So, expectation satisfies the linearity property. Secondly, if $X_1, X_2, \cdots, X_n$ are independent, then

$$E(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} E(X_i).$$

The third property is

$$Var(X) = E(X^2) - [E(X)]^2.$$

The fourth property is that

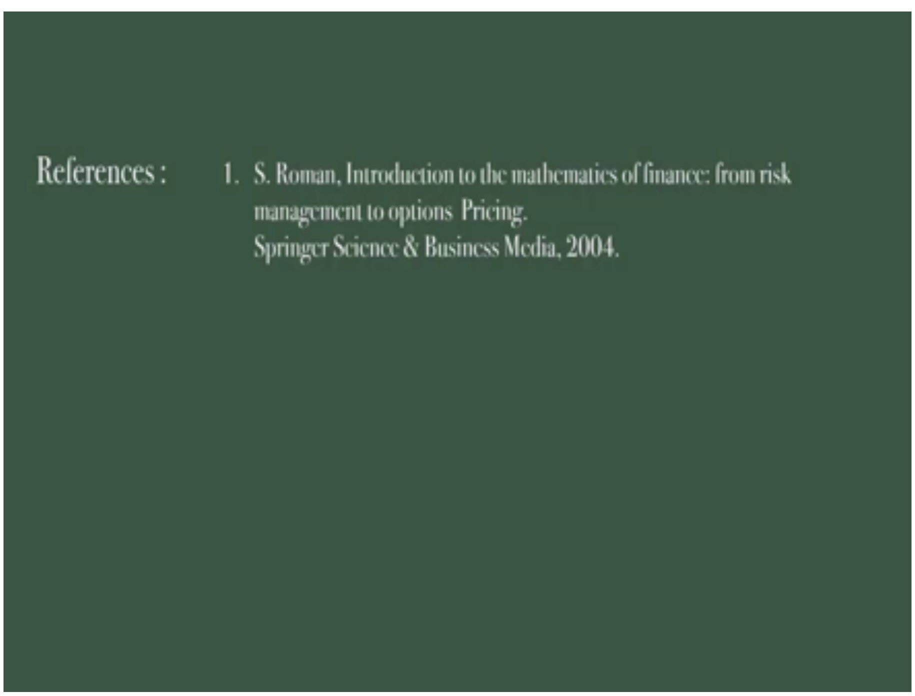
$$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i),$$

for independent X_i's. And the last property is

$$Var(X \pm a) = Var(X),$$

for some real constant a. So, this brings us to the end of this lecture, this lecture just to recall was focused on essentially three main concepts namely the expectation, the variance and the covariance. For the expectation we defined it both in the finite discrete space as well as the continuous space and we have we did the same in the case of variance and we took the concept of variance. And then we extended in the case of covariance and we define what is the correlation coefficient and all this concepts that is the expectation variance and the covariance as well as the correlation concept. They all basically form the fulcrum of the discussion of the modern portfolio theory. So, in the next class, we will talk a little bit about estimation in the context of covariances and we will talk about some important distributions that are relevant in the context of this particular course.

(Refer Slide Time: 50:35)



Thank you for watching.