Dealing with Materials Data: Collection, Analysis and Interpretation

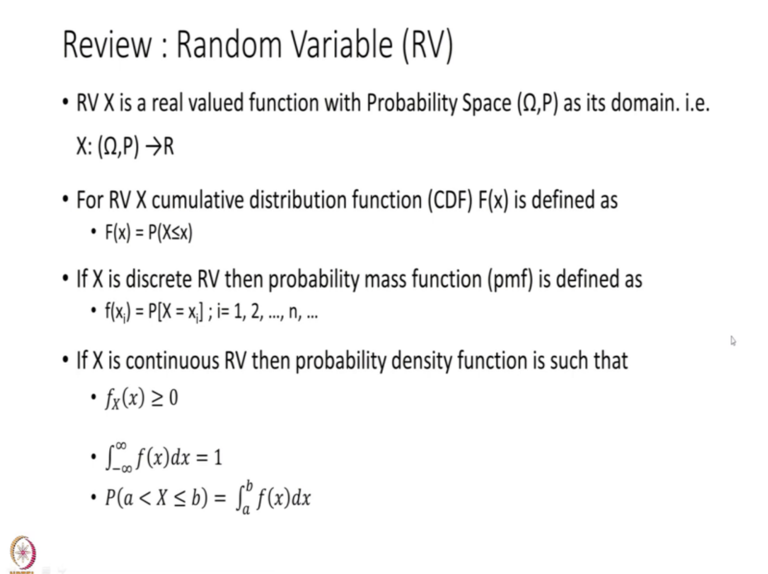
Professor. Hina. A Gokhale
Department of Metallurgical Engineering and Materials Science,
Indian Institute of Technology, Bombay.

Lecture 05

Random Variable and Expectation 2

Hello and welcome to the course on Dealing with Materials Data. In the previous session we introduced to what is known as random variables and from this session onward we would like to talk about its expectations, its moments and many other properties of random variables. Let us review what we did in the past.

(Refer Slide Time: 00:40)



Random variable we defined as a real valued function with probability space omega and P as its domain and then we said that there are, there is always a cumulative distribution function capital F of x is attached to any random variable x and that is defined as

$$\mathbf{F(x) = P\ (X \leq x).}$$

If X is a discrete random variable, then we defined a probability mass function which is nothing but

$$\mathbf{f(x_i) = P[X = x_i]\ ; \quad i = 1, 2, \ldots\ n, \ldots}$$

Please is remember, it is a discrete random variable and therefore it takes countable values.

In case, x is a continuous random variable then there is a probability density function attached to it and it is defined as

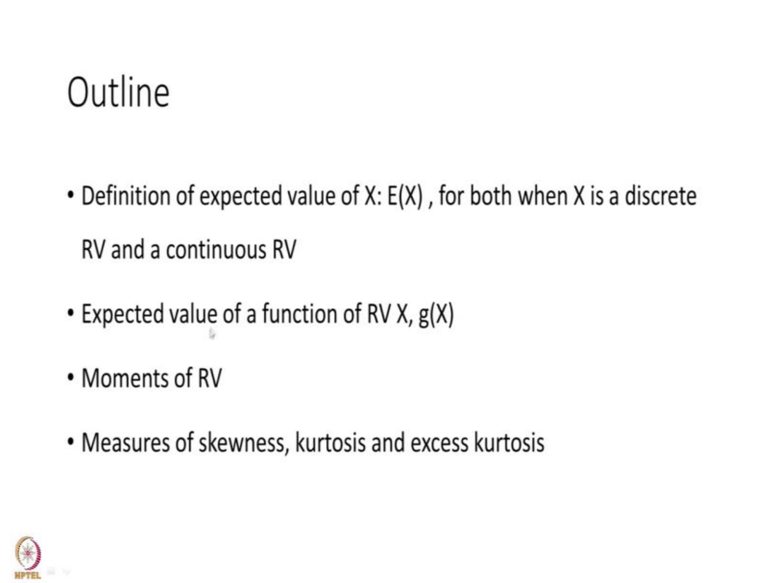
$$f_X(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$P(a < X \leq b) = \int_a^b f(x)dx$$

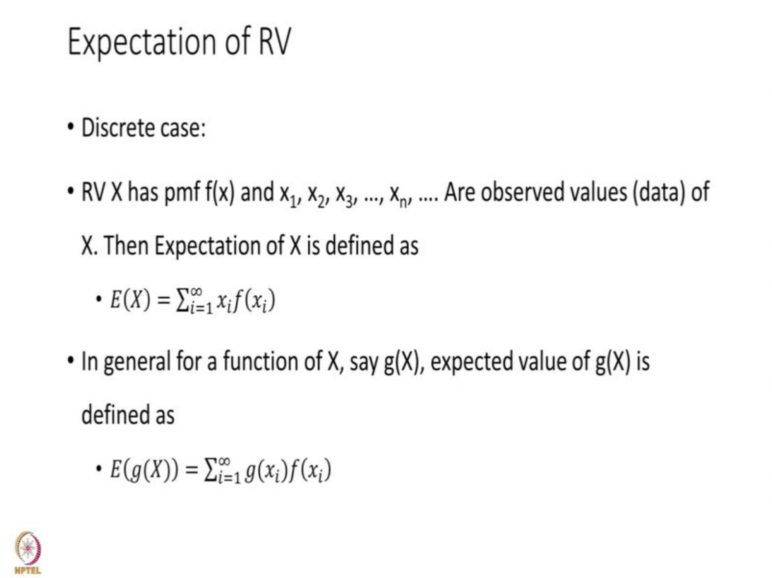
We gave some examples in the previous session and now we move on with the further properties of random variable.

(Refer Slide Time: 02:34)



So, in this session what we plan to do is give a definition of expected value of X, which is given as E(X) for both discrete random variable and continuous random variable. The expected value of a function of a random variable g(X) , we will define that also. We will define what is known as moments of a random variable and then we will give some more details on the measures of skewness, kurtosis and excess kurtosis of any random variable X.

(Refer Slide Time: 03:20)



So, to begin with let us talk about what is expectation of a random variable. See when you have a data or you have a random experimentation or a random event which has a certain output which we called a random variable. We would generally be curious to know that what do we estimate, what do we expect from this data. In other words, what is the expected general value of this data. Like you remember, in the discrete, In the case of a descriptive statistics or exploratory data analysis we said that, x bar or the average value of x gives the general location of where the data lies. Similarly, expected value of x is also a general location where the random variable x lies when it takes different values.

So, expected value of x is of a random variable is like taking a mean value of the random variable. It is actually defined as a mean value of a random variable. However, in this session we want to introduce expectation as a function and therefore, we start that random variable which is a discrete random variable has a probability mass function small f(x) attached to it and we find that x_1,x_2,x_3, x_n are the values, observed values of this random variable X.

Then the expectation of x is defined as expected value of X or E(X) is equal to

$$E(X) = \sum_{i=1}^{\infty} x_i f(x_i)$$

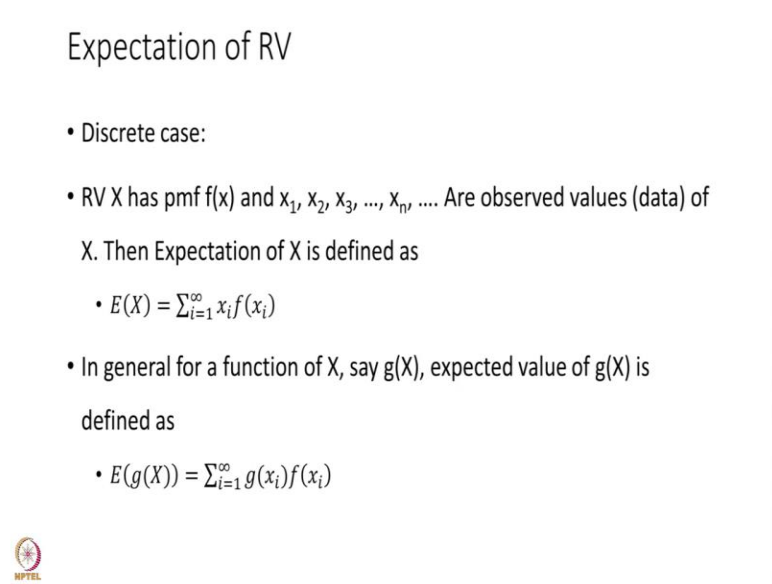
You remember $f(x_i)$ is a probability mass function. So, you can see that this is something like an weighted average of xi, because summation of xi *f(x_i) is 1 so this is like a weighted average of an xi and that is called expected value of x.

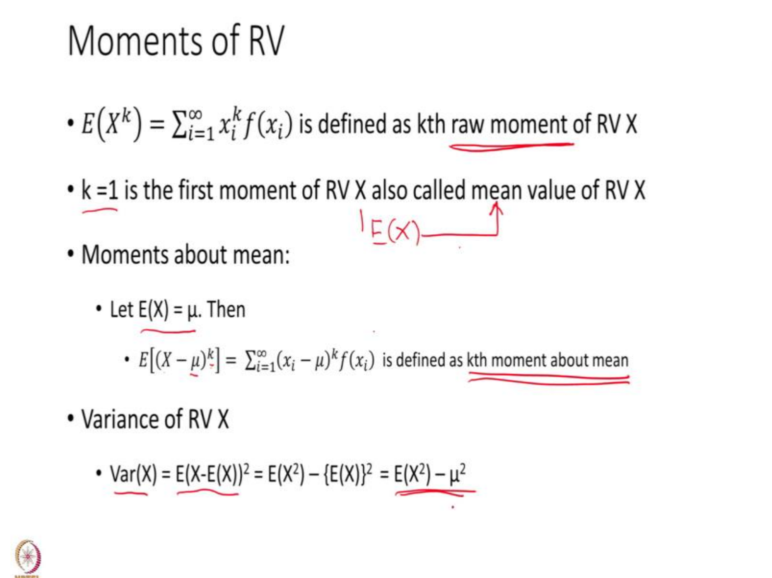
In general for a function of X, say g(X) an expected value of g(X) is defined as

$$E(g(X)) = \sum_{i=1}^{\infty} g(x_i)f(x_i)$$

This is again the probability mass function of the random variable.

(Refer Slide Time: 06:04)





Let us talk of its moments. Now, expected value of a random variable to the power k or $E(X^k)$. Remember that this is also a function of random variable. So, $E(X^k)$ is the kth raw moment of random variable X.

$$E(X^k) = \sum_{i=1}^{\infty} x_i^k f(x_i)$$

This is define as kth raw moment of RV X.

Please remember now we are going to have two kinds of moments. We go back and we okay. I want you to remember that this is called a raw moment of random variable x. If k is equal to 1, then the first moment of random variable X is nothing but expected value of x and it is also called, a mean value of random variable of x.

Let $E(X) = \mu$.Then

$$E[(X-\mu)^k] = \sum_{i=1}^{\infty} (x_i - \mu)^k f(x_i)$$

is defined as kth moment about mean.

If expected value of x is meu, mean then, expected value of x minus mean to the power k is called the kth moment about mean. This is the difference I want you to remember. Something is called a moment about mean and the other one is called the raw moment.

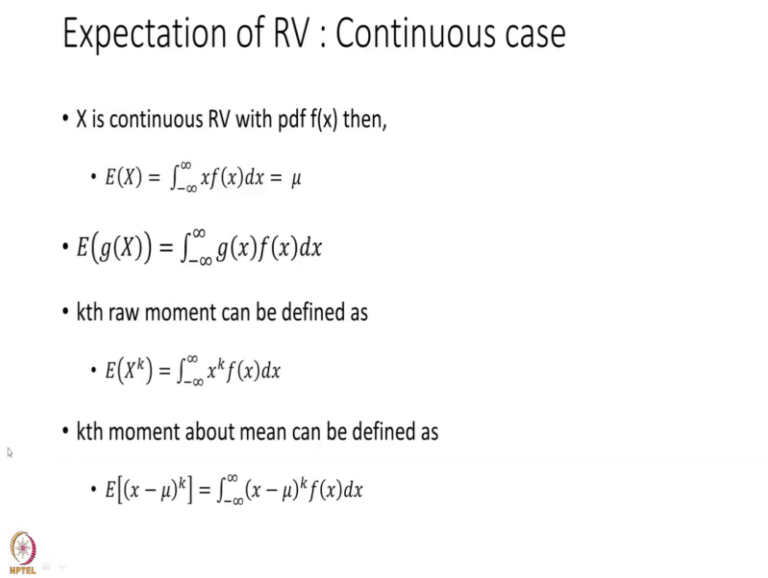
Variance of X is nothing but the second moment about mean of a random variable X.

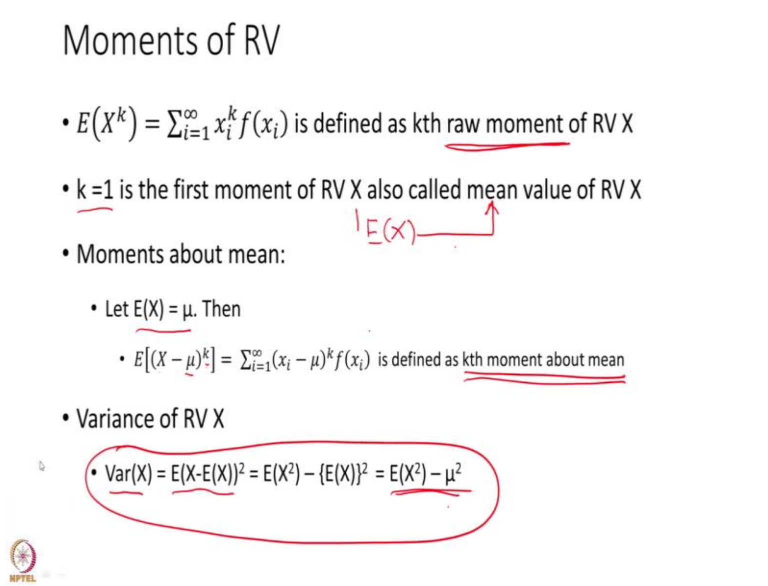
$$\mathbf{Var(X) = E(X-E(X))^2 = E(X^2) - \{E(X)\}^2 = E(X^2) - \mu^2}$$

Please recall, this exercise in a different format we have done it for the descriptive statistics when we defined a when we defined the variance of a data X_1, X_2, X_3, X_n This is for a generalized definition, for any random variable X.

Let us move on.

(Refer Slide Time: 08:30)





Expectation of a random variable in continuous case, So please remember. In general the discrete case we have a probability mass function, in continuous case we have a probability density function and whatever the definition we have for the probability mass function generally in the case of continuous random variable it gets replaced by the integration instead of summation.

So, expected value of X. is first raw moment. Expected value of a function of a random variable X is defined as

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \mu$$

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

The k^{th} raw moment can be defined as,

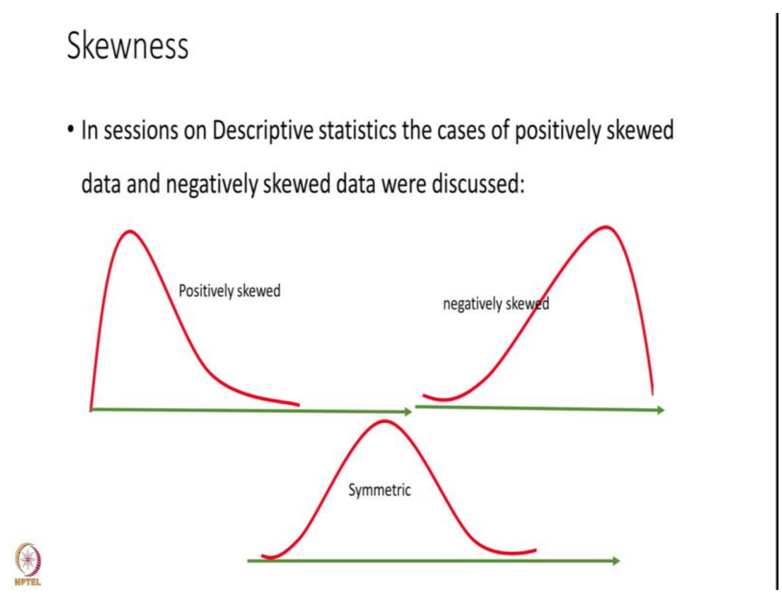
$$E\big(X^k\big) = \int_{-\infty}^{\infty} x^k f(x)dx$$

The kth moment about mean can be defined as

$$E[(x-\mu)^k] = \int_{-\infty}^{\infty} (x-\mu)^k f(x)dx$$

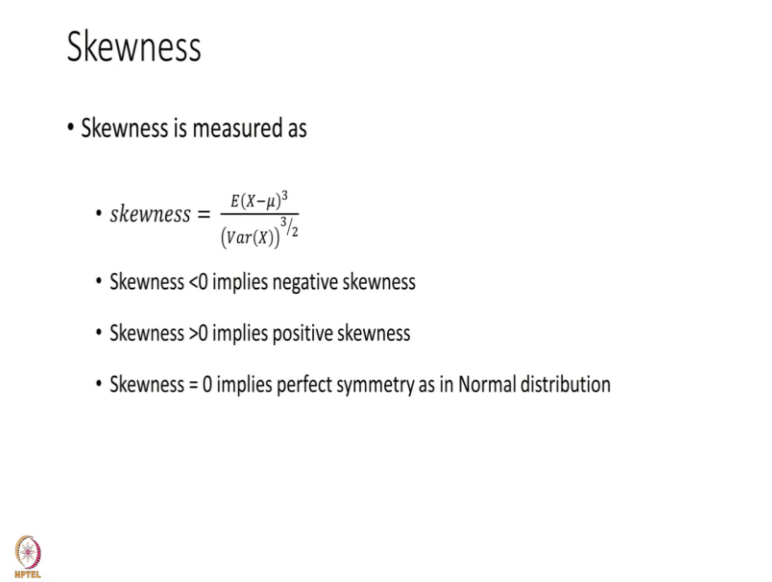
If you recall that as such this formula which is shown here applies to both continuous and random case. So, please note that.

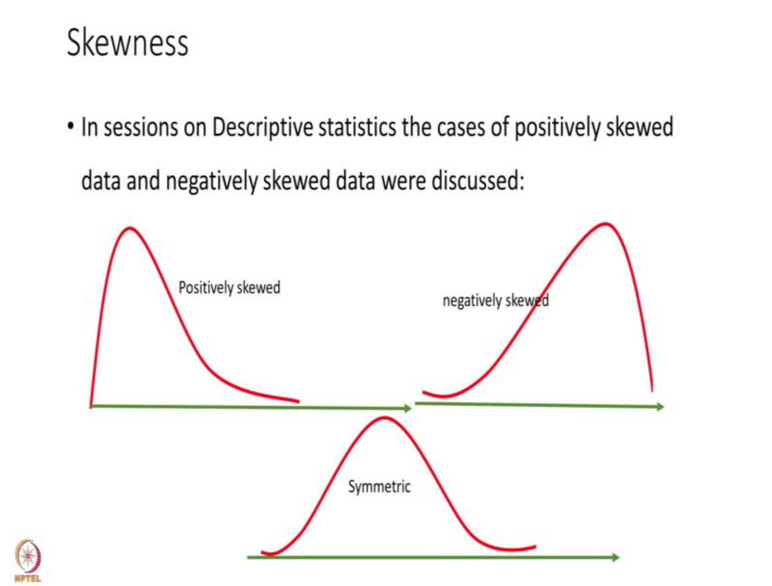
(Refer Slide Time: 10:17)



So, then we come to another important measure which we have briefly introduced in the descriptive statistics. Now we put it in here again. If a data is positively skewed it has a long tail on the right side and if the data is negatively skewed it has a long tail on the left side. This is a symmetric data in which we do not have long tail anywhere and it is very nicely symmetric well shaped curve.

(Refer Slide Time: 10:54)





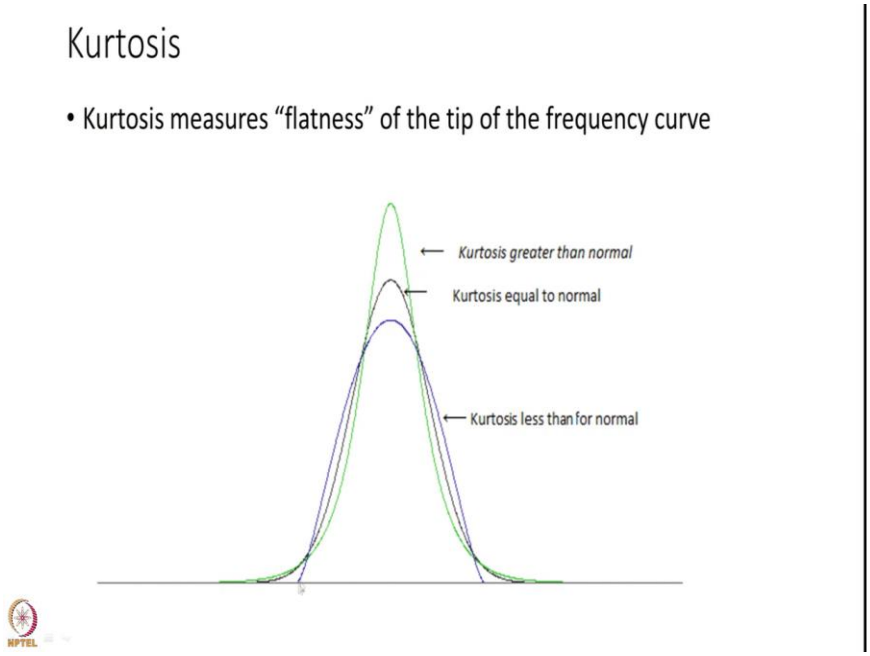
The skewness in terms of expectations is measured as expected value of in other words it's the

$$skewness = \frac{E(X-\mu)^3}{(Var(X))^{3/2}}$$

So that this becomes a unit less quantity .

If skewness is less than 0 it implies the negative skewness. If skewness is greater than 0 it implies the positive skewness and if skewness is 0 it implies the perfect symmetry as in the normal distribution. So, going back to the previous slide here the skewness value as defined as a ratio of the third moment about mean divided by the variance to the power 3 by 2. This will have a positive value because it is a positively skewed, this will have a negative value because it is negatively skewed with a perfect symmetry skewness will be 0.

(Refer Slide Time: 12:13)



Let us move to another quantity which is called Kurtosis. Kurtosis measures the flatness of the tip of the frequency curve. So, here again this black curve is the perfectly symmetric normal distribution curve. This is a sharper tip which is greater than the normal tip and this is a flatter tip which is less than the normal tip.

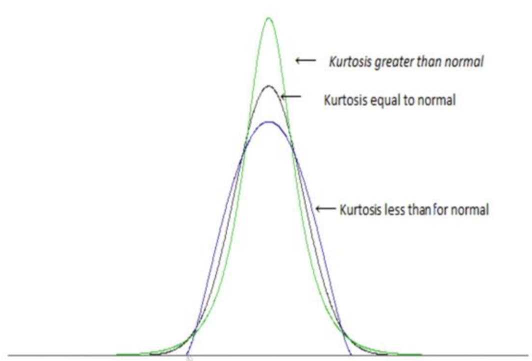
(Refer Slide Time: 12:47)

Kurtosis

- Kurtosis is defined as
  - $kurtosis = \frac{E(X-\mu)^4}{(Var(X))^2}$
- Kurtosis for normal distribution is 3
- Excess kurtosis is defined as
  - $\frac{E(X-\mu)^4}{(Var(X))^2} - 3$
- If Excess kurtosis > 0 , tip of the curve is sharper than Normal distribution
- If Excess Kurtosis < 0 then, tip of the curve is flatter than Normal distribution

Kurtosis

- Kurtosis measures "flatness" of the tip of the frequency curve



This is defined by the fourth moment about mean. Define as

$$kurtosis = \frac{E(X-\mu)^4}{(Var(X))^2}$$

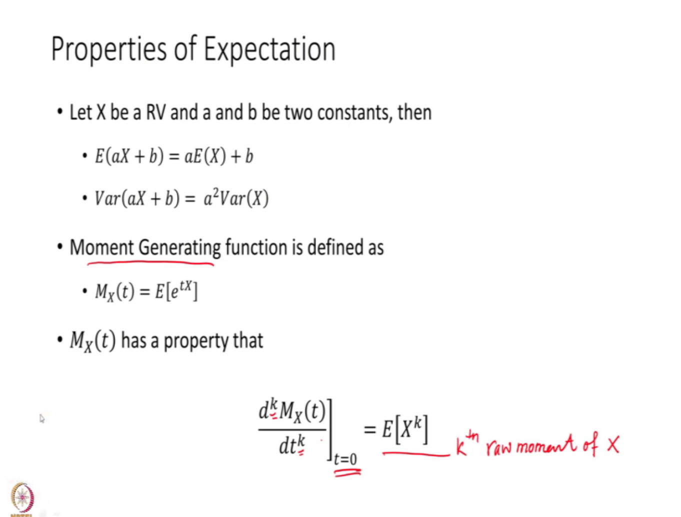
Please remember this variance is already a square term it is second moment about the mean. So you have to match the units and therefore you have the different powers of the variance. So that you have a unit less quantity to, to compare across the data in order to check the kurtosis of the data.So, kurtosis for a normal distribution is always 3.That is if you go back this is a perfectly normal curve the black line if you can see here this is a black line. And this is a kurtosis equal to normal and that kurtosis is generally 3. It is 3 not generally 3 it is 3.

Therefore there is also another definition called excess kurtosis. In which you take this measure of kurtosis since subtract 3 out of it.

$$\frac{E(X-\mu)^4}{(Var(X))^2} - 3$$

If the excess kurtosis is greater than 0 then the tip of the curve is sharper than the normal distribution. If the excess kurtosis is smaller than 0 then the tip of the curve is flatter than the normal distribution. So, here you will have excess kurtosis greater than 0, here you will have excess kurtosis less than 0, this is excess kurtosis will be 0.

(Refer Slide Time: 14:28)



Let us look at certain properties of this expectation. These are very simple properties because you can imagine that expectation is nothing but an integral value and therefore you can show that for a random, any random variable X a and b are two constants. Please remember I am not now going by having a discrete and not discrete case. This is I am saying any random variable X and a and b are 2 constants.

Then

$$E(aX + b) = aE(X) + b$$

$$Var(aX + b) = a^2 Var(X)$$

The moment generating function this is another quantity is defined as $M_X(t)$ which is

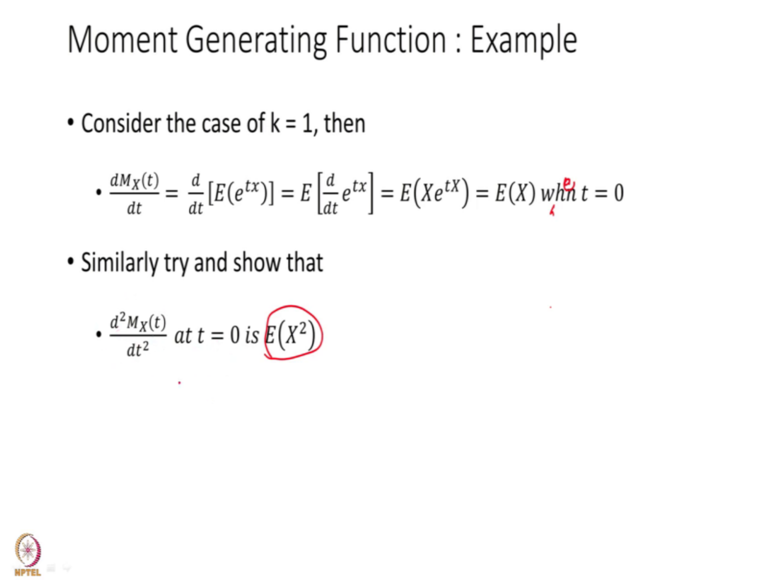
$$M_X(t) = E[e^{tX}]$$

Remember this is a function of X. So we know how to deal with the expected value of a function of X.

The beauty why is it called a moment generating function? Because if you take a kth derivative of this moment generating function and you put t is equal to 0 you get the kth row moment, kth raw moment of the random variable x and therefore it is called a moment generating function. This name is called a moment generating function because it generates the kth raw moment of random variable x by taking the kth derivative with respect to t of the moment generating function and putting t is equal to 0.

$$\left.\frac{d^k M_X(t)}{dt^k}\right]_{t=0} = E[X^k]$$

(Refer Slide Time: 16:40)

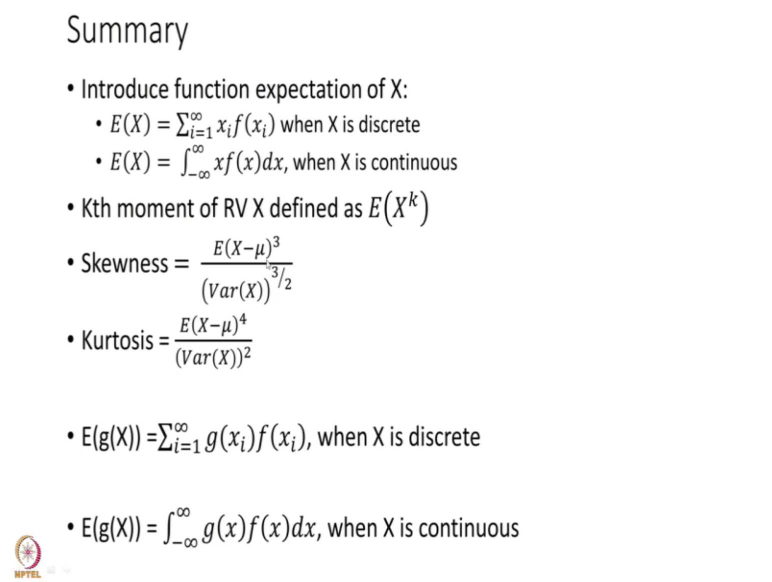


This is I have shown a simple example by working it out. If you take the first derivative of moment generating function t it is derivative of expected value of exponential t to the x. Which is expected value this is interchangeable.

$$\frac{dM_X(t)}{dt} = \frac{d}{dt}[E(e^{tx})] = E\left[\frac{d}{dt}e^{tx}\right] = E(Xe^{tX}) = E(X) \; when \; t = 0$$

Similarly, I leave it to you, you can try and show that the second derivative gives you the second raw moment of the random variable x and you can generalized it to show that the kth moment, kth raw moment will be derived from the kth derivative at t equal to 0 of moment generating function.

(Refer Slide Time: 18: 01)



So, let us summarize what we learned right now, we introduced a function called expectation of x and we defined it for a discrete case as an expectation of x is a summation i is equal to 1 to infinity xi times f of xi which is a probability mass function of x, x and then we defined in the case of x is continuous we defined it as an integral from minus infinity to infinity x f of x dx. We defined the kth raw moment of x define as expected value of x to the power k.

We gave a measure of skewness and kurtosis. These are all also called coefficients of skewness and kurtosis. And we showed with respect to normal distribution what values they take. Skewness less than 0 indicate negative skewness. Skewness greater than 0 indicates positively skewed data. Kurtosis less than 0 indicate that tip is flatter, flatter. And when kurtosis is greater than 0 it says that the tip of the curve is sharper. We also learn that an expected value of a function of an X can also be defined as an in case of discrete as a summation, i is to 1 to an infinity g(xi), f(xi). Otherwise you integrate in case x is a continuous function.

Thank you!