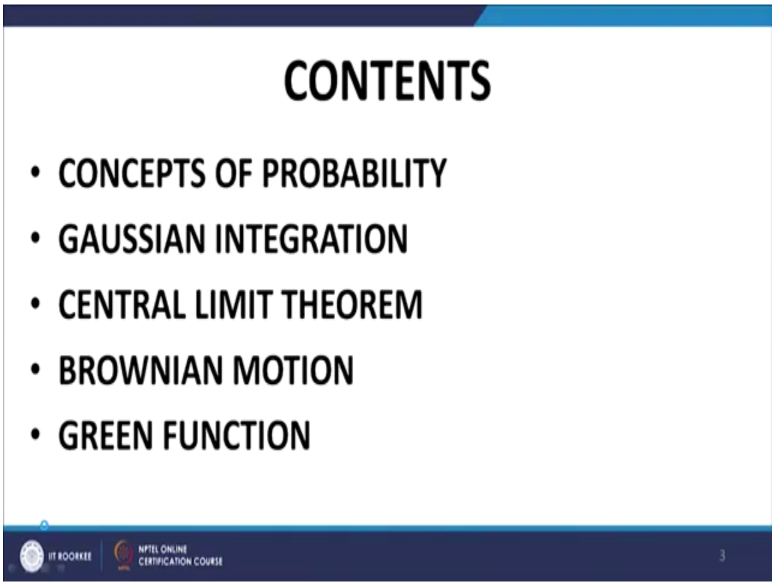
**Path Integral Methods in Physics & Finance**
**Prof. J. P. Singh**
**Department of Management Studies**
**Indian Institute of Technology, Roorkee**

**Lecture – 03**
**Probability, Generating Functions**

Welcome back, in the last lecture I gave you a feel about what path integrals are; in a sense it they are the integration of functionals or functions of functionals of variables. And then I discussed the relevance of path integrals in the context of quantum mechanics.

To continue with that; to continue a more formal description of path integrals, I need certain prerequisites; to make this course self contained to some extent, I shall be covering them at reasonable speed and not withstanding that; it may take me about two lectures to cover up most of the things.

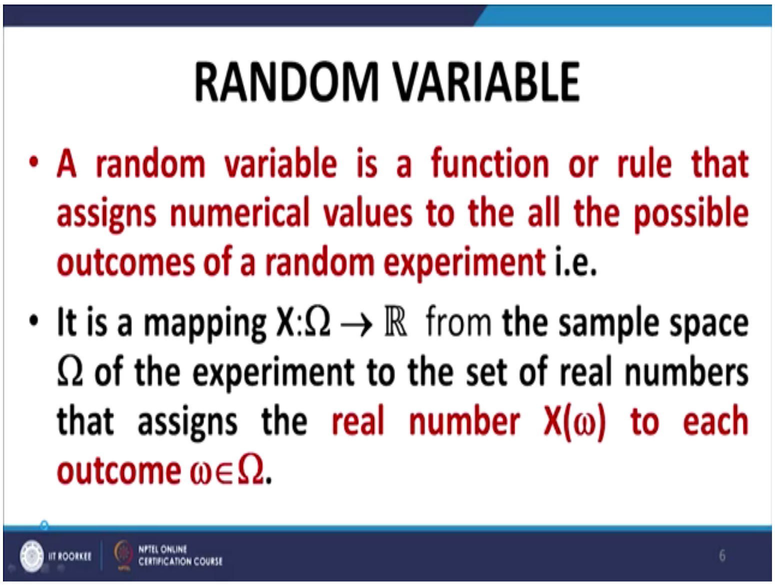
(Refer Slide Time: 01:07)




The topics that, I propose to cover as prerequisites; basic concepts of probability, Gaussian integration, the Central Limit Theorem, Brownian motion and brief introduction to the theory of Green functions.

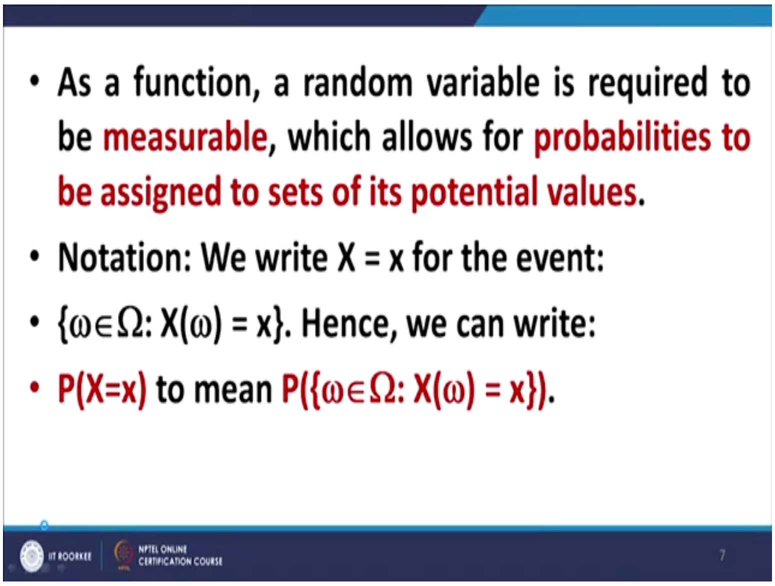
So, let us start; when we have an random experiments, an experiment whose outcome we are unable to predict with any kind of certainty; we are not precisely; able to precisely predict, the outcome of an experiment; we call it a random experiment. The set of all possible outcomes constitutes the sample space of that experiment. Whatever outcomes are possible on performance of that experiment, they constitute the sample space.

(Refer Slide Time: 01:57)




Now, a random variable is the variable; which is a mapping; which maps the elements of the sample space on to real numbers; in other words, every outcome is assigned by assigned a real numbers by some kind of a rule and that rule constitutes the random variable. So, random variable is a; then is a mapping from the sample space on to the set off real numbers.

(Refer Slide Time: 02:36)

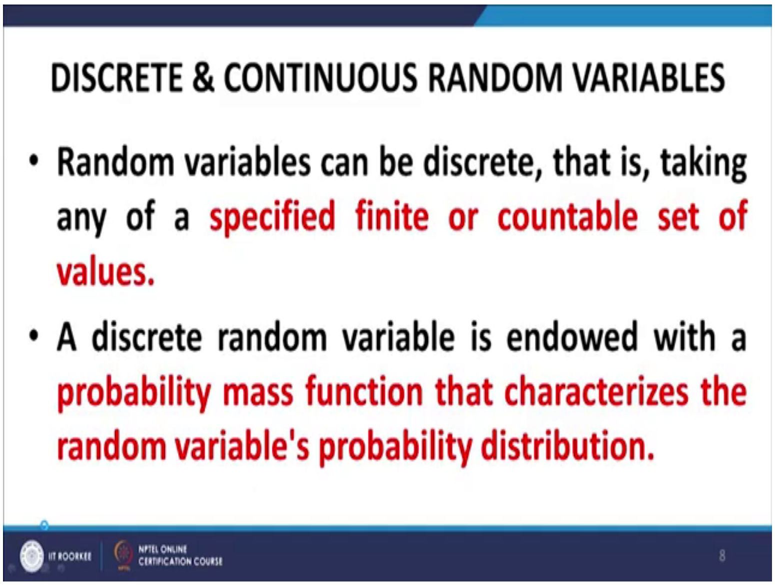



Now, this random variables may be discrete or the random variables may be continuous; depending upon weather the co domain of the random variable is discrete or continuous. In other words, if the numbers that are assigned by the mapping, the random variable to various outcomes of the experiment; if they form a discrete set or a countable set, then it is said to be a discrete random variable.

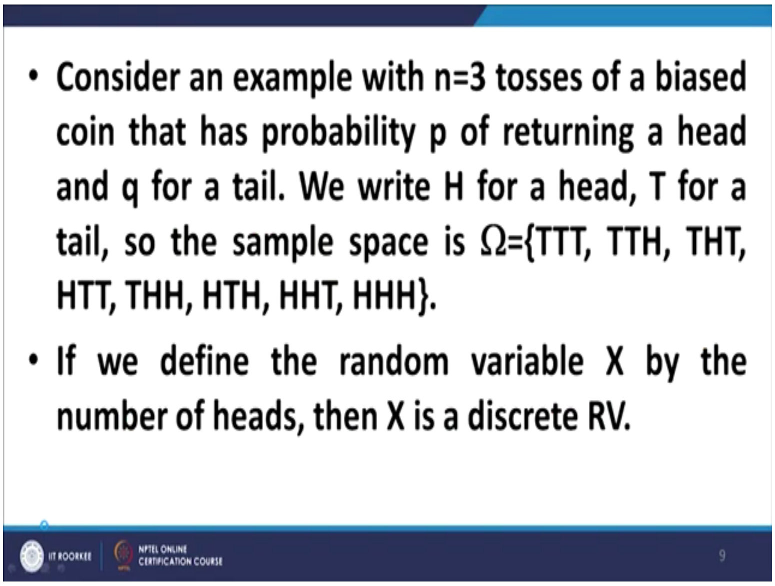
On the other hand, if the assignment by the random variable can take any value in an interval or a collection of intervals, then it constitutes a continuous random variable. And there is an abbreviation here which is very common very popularly used; we use the abbreviation which is here P; the capital alphabet is generally used to represent a random variable.

And the small alphabet is used to represent a particular value; a particular realization of that random variable. I will come back to it, but basically we use the abbreviation P; capital X equal to small x to represent the fact that the mapping represented by capital X of an element, small omega within the sample set is mapped on to number; real number small x.

(Refer Slide Time: 04:01)



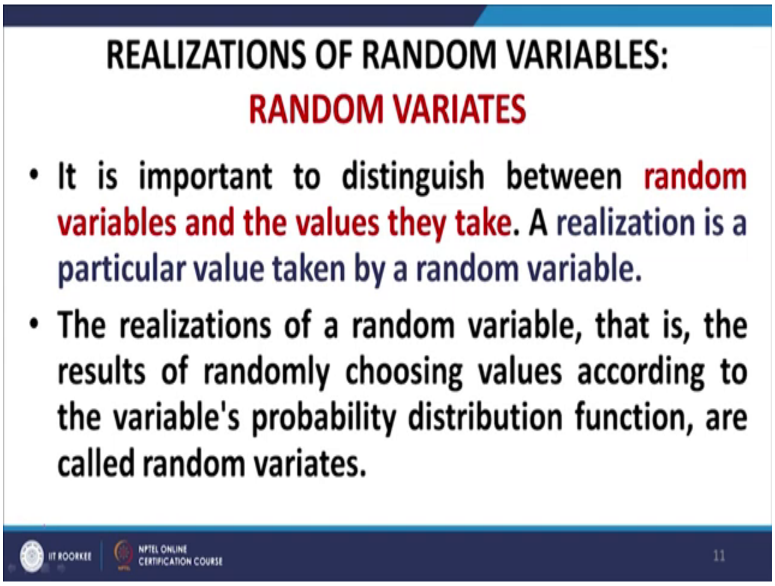
(Refer Slide Time: 04:03)




So, discrete in continuous random variables; discrete random variables when we have the co domain or the numbers that are assigned by the mapping, they are discrete numbers, they are countable; then it is a discrete random variable. And if they can take any value within a particular interval; it constitutes a continues random variable.

For example, if you are tossing 3 coins and you define random variables by the number of heads; then, obviously, the random variables can take the value 0 heads, 1 head, 2 head, 3 heads. So, these this is the example of a discrete random variable. The age of students in a class; constitutes an example of a continuous random variable.

(Refer Slide Time: 04:52)



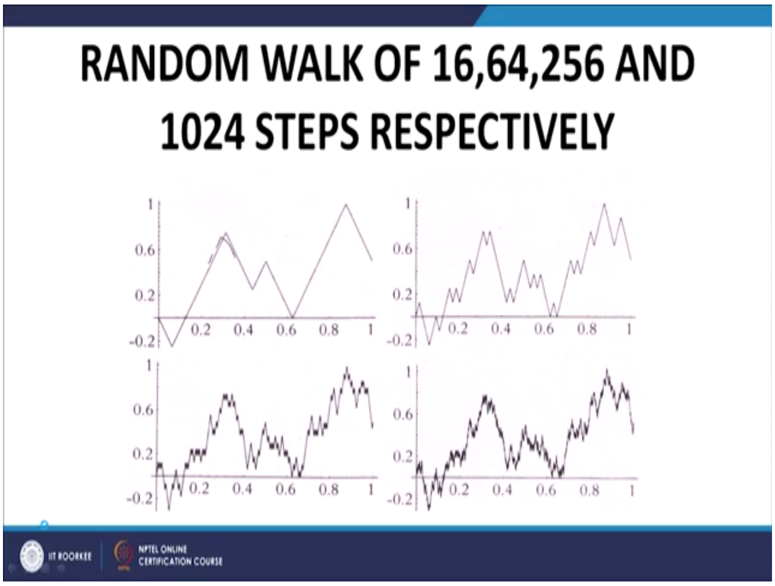

Now, there is an important point; we need to differentiate between the random variable itself which is the mapping and the values that the random variable can take. For example, if you toss 3 coins simultaneously; then you can define if you like, you can define a random variable as the number of heads that arise on the toss of 3 coins simultaneously; obviously, then the possible values that can; that the random variable can take 0, 1, 2 and 3.

Now, any particular value say you make a toss of these coin and you return 2 heads in the; as the outcome, then in that case the 2 heads constitutes an outcome and the 2 heads constitutes a realization, a particular realization of the random variable.

This has to be contrasted with the mapping itself; the this is, these 2 heads is a number where as the random variable itself is a mapping; so, we need to be careful of this restriction.

Normally, the mapping is represented by the capital alphabet and the value the on a particular realization that the random variable takes is assigned a small number.

(Refer Slide Time: 06:18)

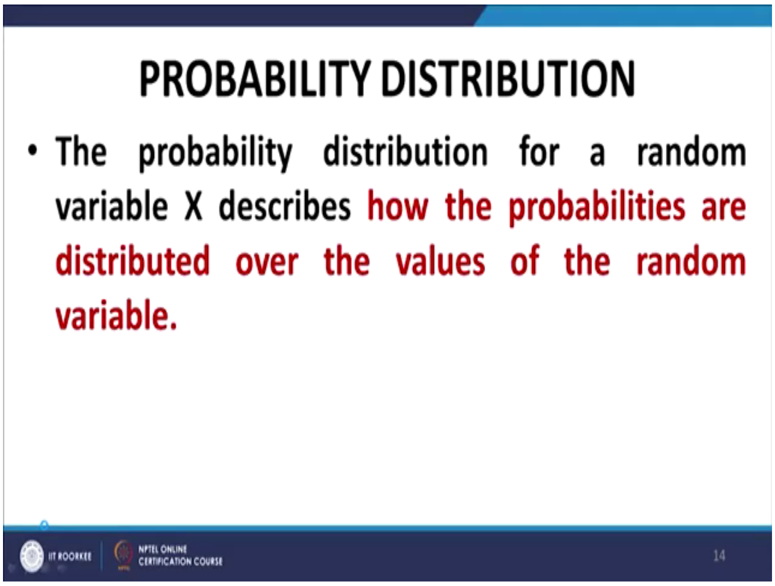



For example, if you look at the slide; now, this is the slide of a random walk; random walk is a sequence of random variables. In this particular case, we have an unbiased random walk; so the probability of up and down is the same; that does not matter, it is sequence the collection of random variables which is index by time.

So, at let us say at t equal to 1, you make a first t equal to 0; you start from the origin, t equal to 1, you make a coin toss and if the; if you return a head, you move up one step and if you return a tail, you move down one step, t equal to 2; you repeat the same process; this is an example of a possible outcome that may have arisen in a particular experiment.

Now, it is certainly not necessary that if the experiment is repeated, we get a replica, a precise replica of this particular outcome, so these are called realizations of the experiment. Realizations may be; must be distinguished from what the random variable itself is; realization is the outcome a particular outcome or a particular number assigned to a particular outcome of the random variable.

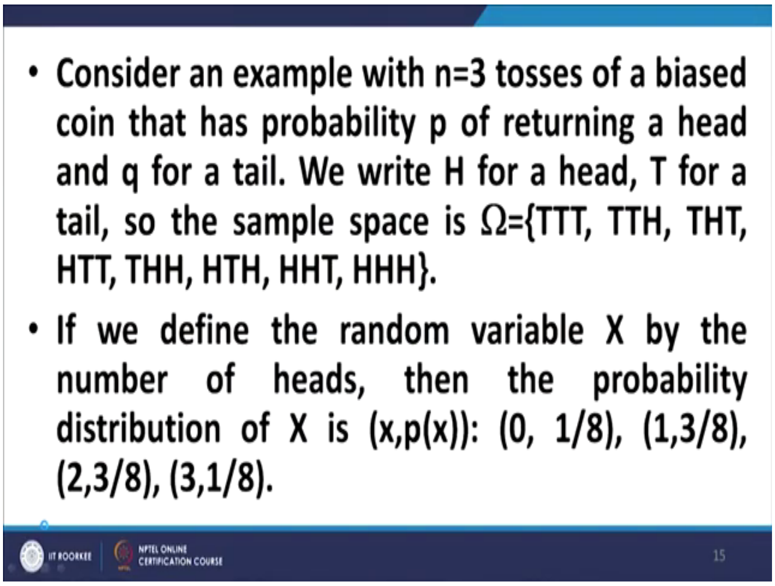
(Refer Slide Time: 07:36)




Now, we come to the concept of probability distribution; as I mentioned random variables assign a real numbers to the various possible outcomes which are enlisted in some form or the other; which are enlisted or which are present in the sample space. You have the entire set of outcomes, you are constituting the sample space and for each particular outcome, you assign a real number.

Now, on the basis of some either a subjective reasoning or objective experimentation; you assign certain probabilities to the various outcomes. This combination of the various values that the outcomes represent and the probabilities that, those values are likely to take, constitute the probability distribution. Let us continue with that example of a simultaneous toss of 3 coins and let us again define the random variable by the number of heads, then the various possible outcomes are 0 head, 1 head, 2 and 3 heads.

(Refer Slide Time: 08:55)

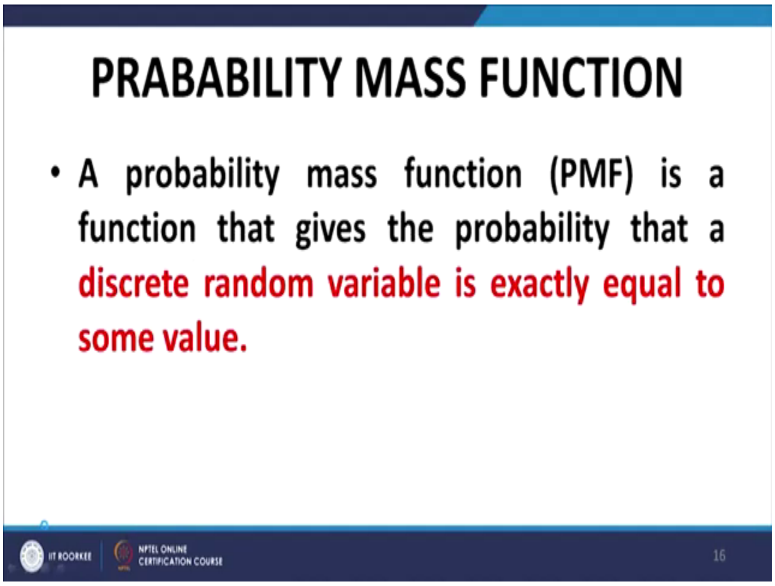



And the probabilities of their happening are; as you can see, the probability of their happening is 0 heads is 1 by 8, 1 head is 3 by 8, 2 heads is 3 by 8 and 3 heads is again 1 by 8.

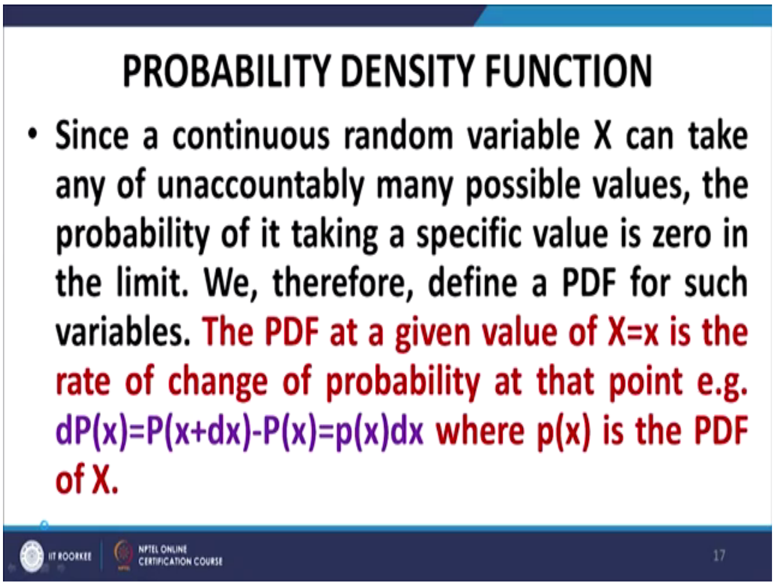
So, this combination represents a probability distribution where you have the various values of the random variable enlisted, together with the respective probabilities of occurrence of those values. Now, a particular; in the case of a discrete random variable, a particular probability is assigned to the occurrence of a particular value of the random variable, a particular realization of that random variable and that is what is called the probability mass function.

For example, in the previous slide, if I select; if I want to know the probability of 0 heads, then the probability mass function of 0 heads is 1 by 8.

(Refer Slide Time: 09:47)



(Refer Slide Time: 09:48)

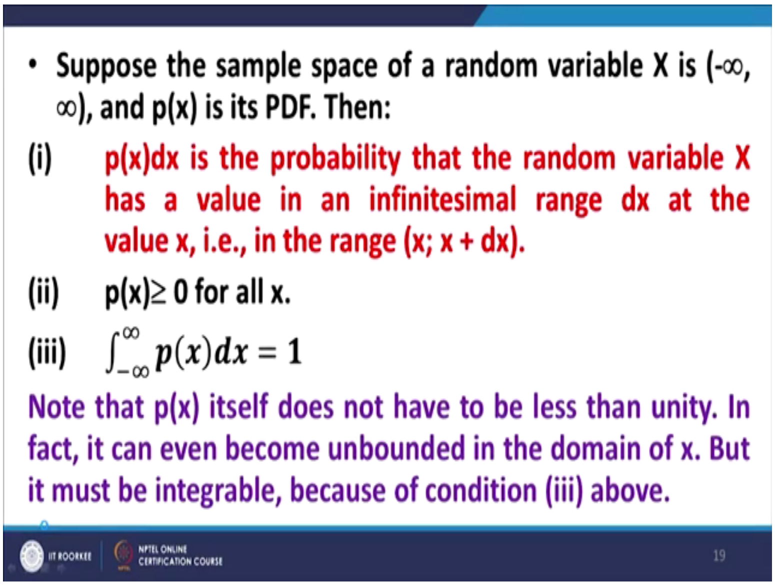



And that is how the definition goes; however, in the case of a continuous random variables where the random variable can take value within a real interval that means, it can take one of an infinite number of values in the interval. The probability of the random variable taking precisely a particular value; taking a particular point in the real interval approach is 0.

Because there are infinite numbers of possible outcomes and when you have an infinite number of possible outcomes, then the probability of one particular outcome approaches a 0 naturally. Therefore, in such a situation, it becomes meaningless to define a probability mass function, it is more appropriate to define a probability density function.

We define the probability density function as the function p x such that p x; d x, p x; d x constitutes the probability of finding of the random variable taking a value between x and x plus d x, where d x is very small. In other words, the probability of the random variable lying in the interval x to x plus dx is given by p x; d x and p x that particular quantity p x is called the probability density function.

(Refer Slide Time: 11:29)



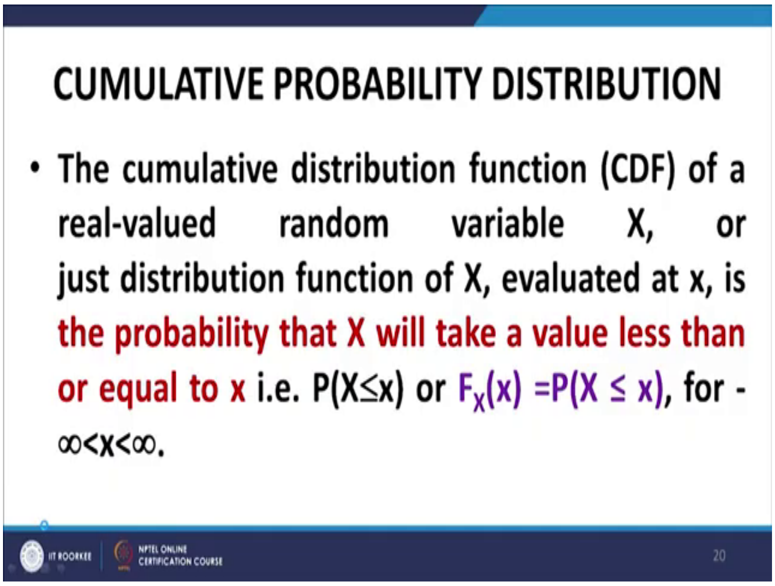

Now, naturally the probability density function needs to satisfy certain probabilities and because by enlarge, we shall dealing with continuous distributions; let us talk of little bit more about it; p x; dx is the probability that the random variable X has a value in an infinitesimal range x to x plus dx because, if I workout the probability of the random variable taking a value precisely X that will approach 0 because the random variable can take an infinite number of values. So, any particular value it approaches 0.

So, then; obviously, the probability invariable has to be positive or 0, it cannot be negative. And thirdly the integral of the probability over the entire range of values; of the realized possible values of the random variable has to be equal to 1; that is the definition of probability.

If you integrate or if you sum the probabilities over all possible outcomes, it has to be equal to 1.

But, the important thing here is to note that p; x itself need not necessarily be less than 1, at every point in a given interval; on which it is defined. It is not necessary that at every point p; x needs to be less than 1, p; x can in fact, be unbounded as well; provide, it is integrable and this condition 3, that is integral p; x, dx is equal to 1 is satisfied; p; x need not be restricted to 1; of course, it has to be integrable.

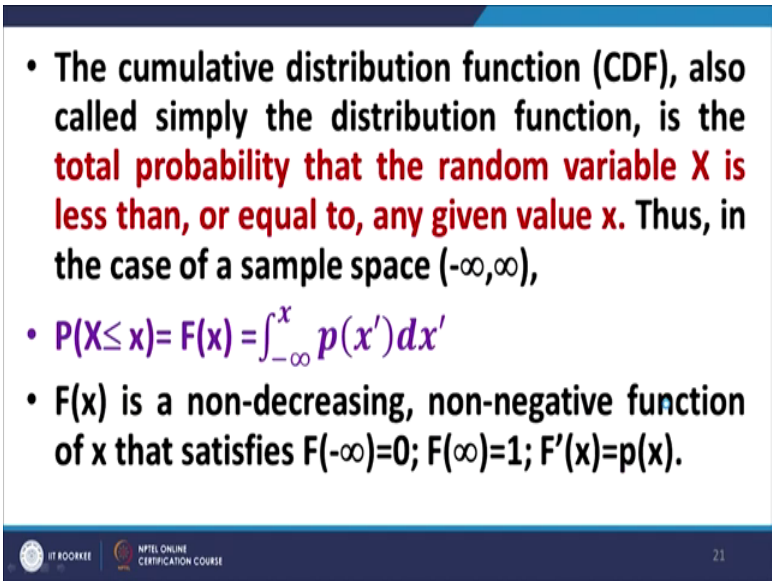
(Refer Slide Time: 13:15)




Cumulative probability distribution is an extension of the concept of probability density function. In this case, what we are concerned about is the probability of the random variable taking values below a certain pre assigned value. You are given a certain pre assigned value, let us say X equal to small x; then what are, what is the probability of capital X, that is the random variable taking values below that is value small x; that constitutes the probability, the cumulative probability distribution.

So, this is the definition of cumulative probability distribution; P; F of x is equal to P capital X is less than equal to x. So, whatever is the lower bound; whatever is the minimum value up to the pre assigned value, what is the probability of the random variable lying in that range that constitutes the cumulative probability distribution; obviously the following follow immediately.

(Refer Slide Time: 14:21)

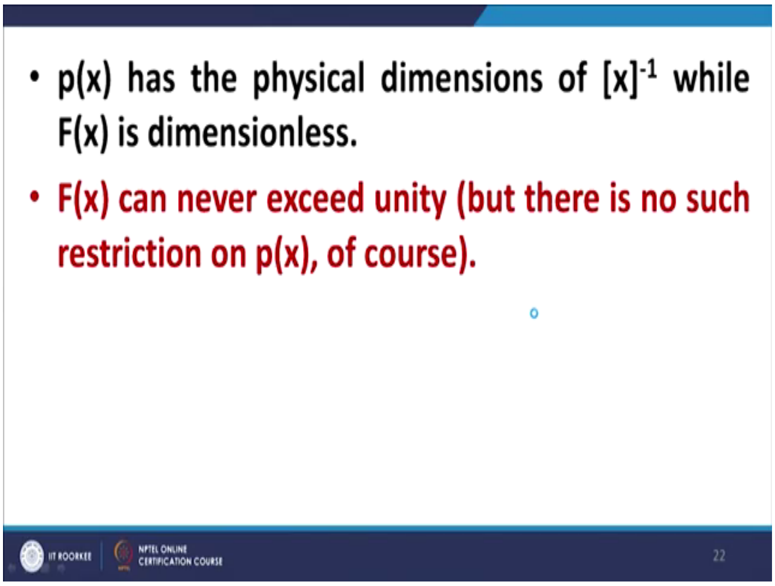



The cumulative probability distribution at minus infinity has to be 0 because there cannot be any value lower than that; of X which is lower than that. The cumulative probability distribution at X equal to infinity has to be 1 is because there cannot be any value above that.

And the differential of F; x with respect X constitutes the probability density function and further more in the case of continuous distributions, you can define; you can define, the cumulative distribution function by this equation F. Capital X is less then equal to x is equal to integral minus infinity up to the pre assigned value, let us take it as small x and p; x dash and dx; that means, it is giving you the total probability of the random variable capital X, lying upto from the lower bound; up to capital X, I am sorry small x.

The probability density function has dimensions of X inverse; however, the cumulative density function or the cumulative distribution function is dimensionless that you can verify; these two conditions are there.

(Refer Slide Time: 15:47)

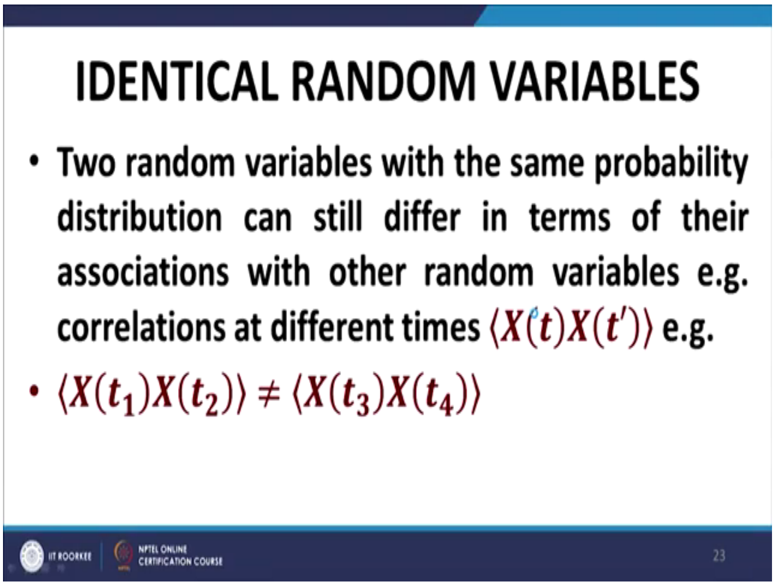


And secondly, capital F; x that is the cumulative distribution function, evaluated at any point can never exceed 1. Maximum value is 1, where all the values of the random variable are covered within the range of summation or the range of integration.

Now, identical random variable; two random variables which are identical in distribution need not necessarily; in other words if they have a same distribution; let us say they both are normally distributed with the same mean and the same variant, but they can still be different in terms of their association with other random variables.

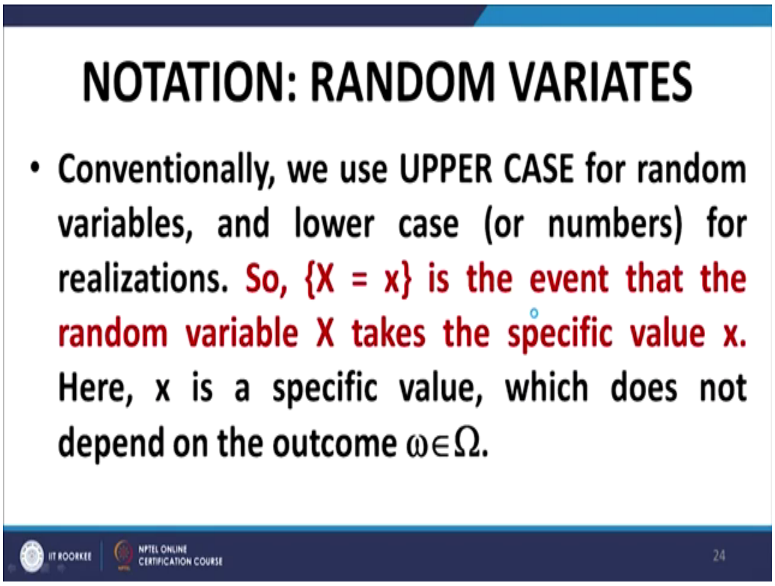
Take a simple example, if I workout the co variance or the co relations between random variable at; between a particular property represented by a random variable of a physical system at X is equal to t 1 and the same property at later point in time X equal to t 2, then it is given by the expected value of X; t 1, X; t 2 and this need not be precisely the same as X; t 3 and X; t 4 corresponding to a different pair of time points t 3 and t 4.

(Refer Slide Time: 17:07)

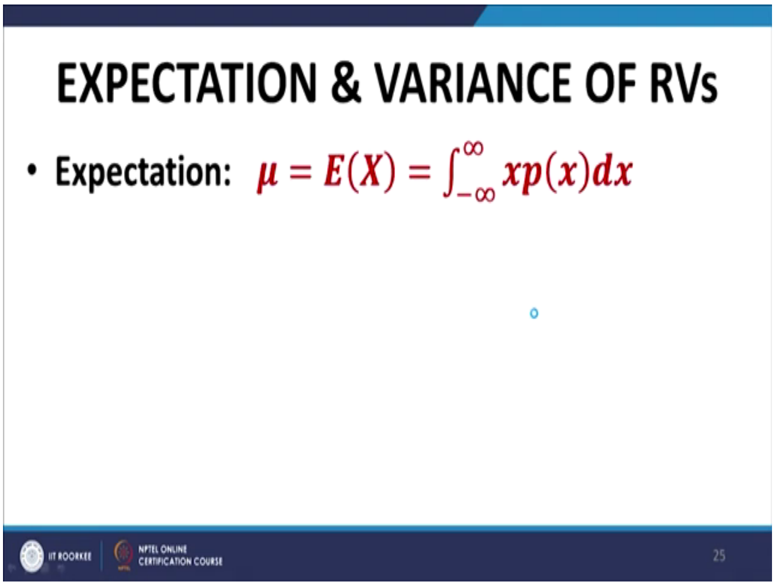



So, it is not necessarily true that random variables which have identical distributions; do have identical associations with all other random variables.

(Refer Slide Time: 17:21)



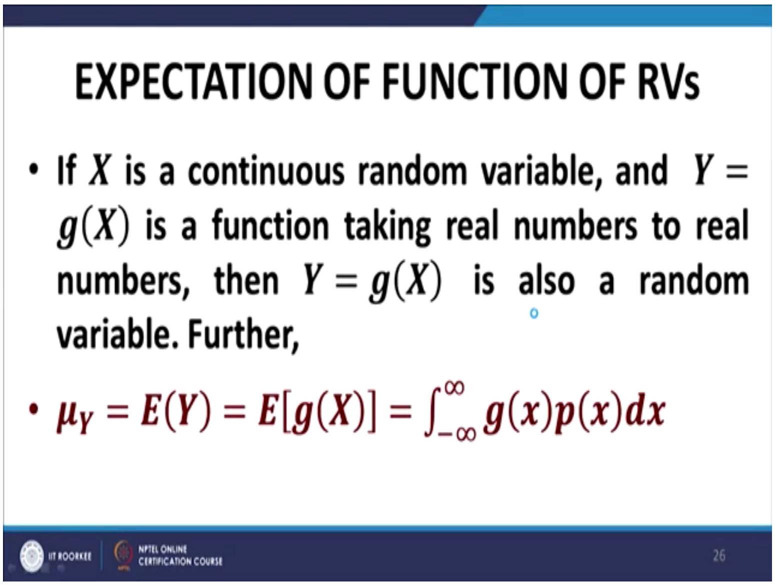
(Refer Slide Time: 17:25)




Notation, I have already explained; the concept of random variates. Now, expectation and variance of random variables; expectation is very common term, expectation is given by the mean value.

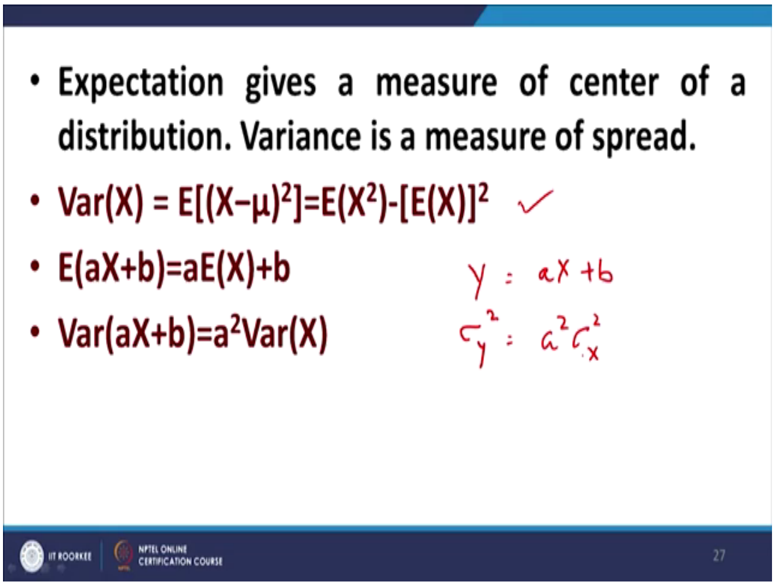
The mean value in the case of; in the case of a discrete random variable is given by sigma X into p x; the summation done over all possible values that the X can the variable X can take. In the case of continuous distributions, it is given by the integral over all possible values that the random variable can take x into p x; dx.

(Refer Slide Time: 18:03)




And, if you want to work, if suppose you are given a random variable X and you are given a function of the random variable X. Let us say g; X is a function of X, then the expectation value of g; X is given by; you simply substitute g; X in value of X and you the expectation is given by g; x, p; x, dx.

(Refer Slide Time: 18:34)

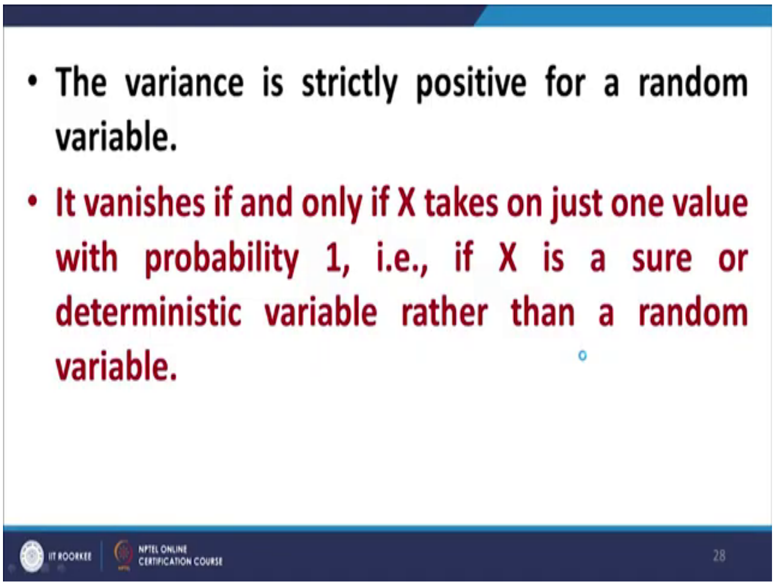



For example, if you want to work out the expected value of X square; instead of X; p; x, it will be X square p; x. These are certain identities which are followed by the expectation measure and the variance measure. The variance of X is given by the expectation of X square minus the expectation of X whole squared. I repeat, this is a very important identity; variance of X is equal to expectation, it is the second moment about the mean; variance of X is equal to E of X square, expectation of X square minus expectation of X whole squared.

So, that is an important thing and the expectation of a random variable; suppose you transform the random variable X to another random variable Y, Y is given by let us say Y is given by a of X plus b. Then, the expectation of Y is given by a into expectation of X plus b; however, when you talk about the variance of Y; sigma square of Y, then this is given by a square sigma square of X.

This; these three identities are very important, very simple, but never the less very important, first is variance of X is equal to E of X square minus E X whole square where expectation value of aX plus b is equal to a; expectation of b, expectation of X; I am sorry plus b and variance of aX plus b is equal to a square into variance of X; variance is strictly positive, that is quiet obvious.

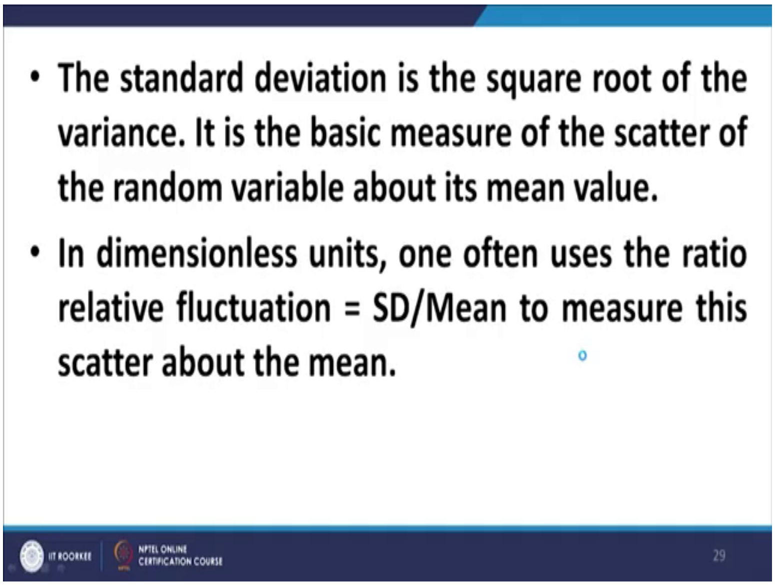
(Refer Slide Time: 20:29)




The important thing is; the variance will vanish only if the random variable takes only one value and if the random variable take only one value, then it is said to be a deterministic variable. It no longer remains random because once a particular variable is going to take only one value, there is no element of randomness; that means, you see it; it implies that the value of that random variable can be perfectly predicted and once it can be perfectly predicted, it no longer remains random.

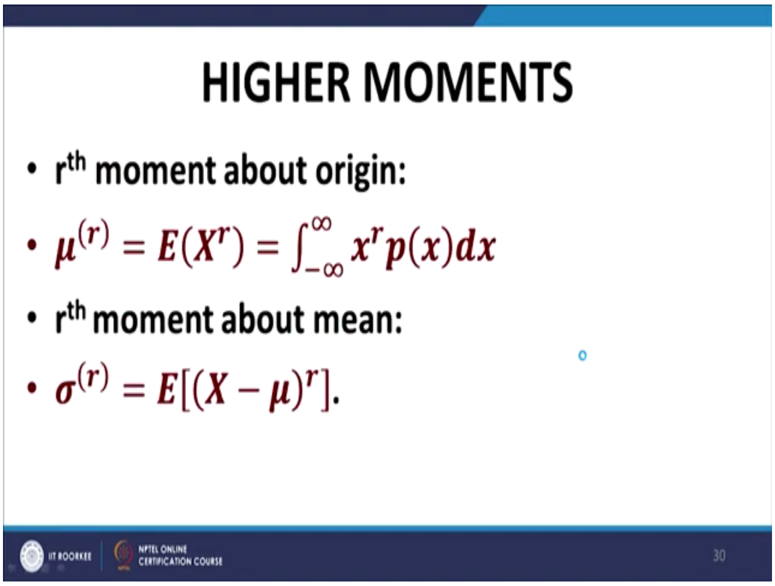
So, that randomness is lost therefore, in another words; the inferential that in the case of deterministic variables or this deterministic evolution, you can say the variance is 0 and this operates both ways; if the variance is 0, then the variable is; if variance of X is 0, then X is deterministic conversely, if X is deterministic; the variance of X is 0.

(Refer Slide Time: 21:42)




The standard deviation which is more commonly used is the square root of variance and usually we have another quantity; which we use as the ratio of the standard deviation and the mean so that it becomes dimensionless and this is relative fluctuation.

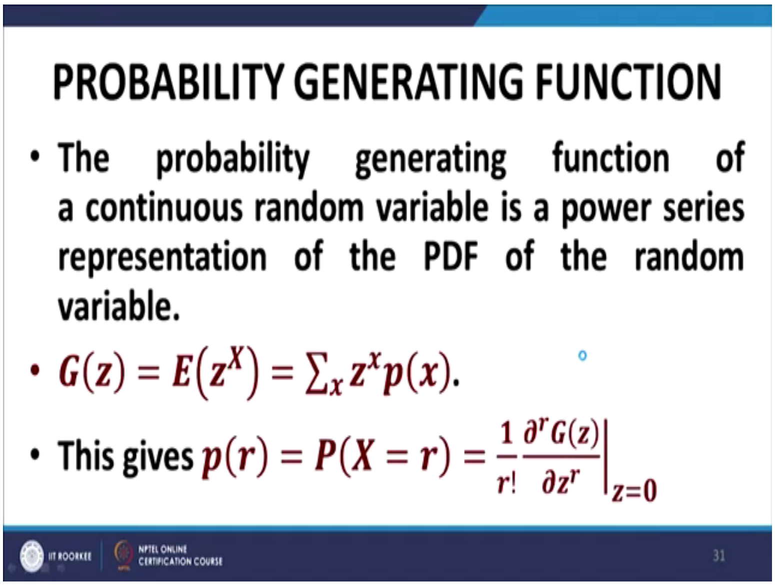
(Refer Slide Time: 21:58)




Just like mean and variance, we can define higher moments of a distribution; higher moments of a probability distribution, the rth moment about the origin. Origin means X equal to 0, the rth moment about the origin is defined by the expectation value of X to the power r, that is integral X to the power r; p r into x right. Integration, again over all possible values that X can take; the whatever values the random variable can take the integration or the summation, as the case may be depending on whether it is a continuous random variable or a discrete random variable; it has to be done over all possible value.

Similarly, we define the rth moment about the mean as expectation value of X minus mu to the power r.

(Refer Slide Time: 22:55)




Now, we come to a very interesting concept probability generating function. We define the probability generating function G of z; given a continuous random variable, we define G of z as E of the expectation value of z to the power X, where X is the random variable, z is simply a parameter; ok. So, that is equal to summation of x; z to the power x, p; x because we are taking the expected value of z to the power x. Now, z to the power x is a function of x.

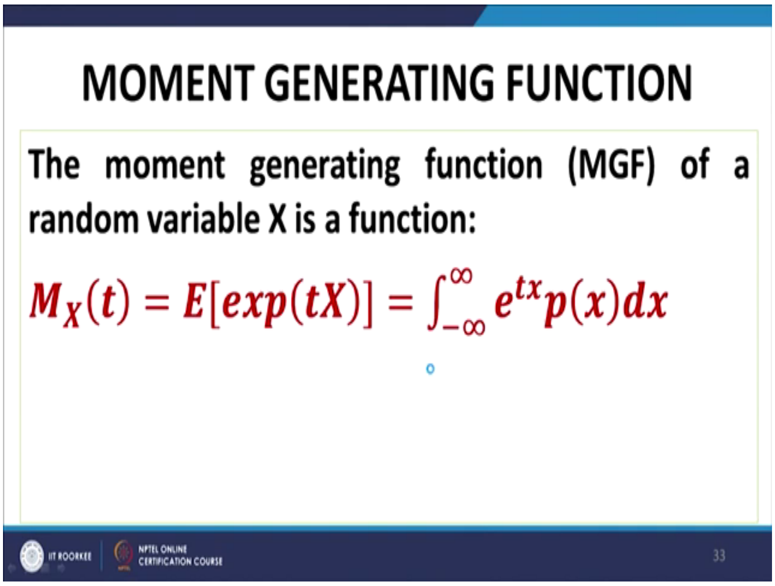
So, we use that rule that I told you earlier; we have summation of x; z to the power x into p; x. Now, if you differentiate this expression with respect; to with respect to z and put z equal to 0; first in fact, if you expand this simply; put z equal to 0, you get the value p 0. Then, if you differentiate it first time; you workout d 1 x and d 1; G; z upon d; z a and then put z equal to 0, here you get the value p 1.

And, similarly if you differentiate a second time; you get the value of g 2 with the correction for the factorial. So, in other words knowing this particular G; z, we can workout the probability corresponding to any value of X or any value of the outcome of the random variable. This concept of probability of generating functions is a very fundamental concept which is going to crop up again and again, when we talk about path integrals.

In that context of course, we will be; we will be using the concept to generating functionals rather than generating functions. But generating functions gives you a start, a feel about what exactly we mean by generating functions or functionals. So, to repeat G; z is equal to E of z to the power X, which in the summation form can be written as because we are taking the expectation values of it, writing it like this and then when we differentiate it one by one, we get the various values of p 0, p 1, p 2 and p 3 on putting z equal to 0.

Now, I will just explained the concept of moments; the first moment is the mean, the second moment is given by E of X square and third moment E of X cube and so on. Now, if we know the; all the moments of a particular distribution, we literally know the distribution itself. The distribution is determined by its moments or it can be viewed as other way around as well, but the inference is that if we know the moments of a distribution, the distribution can be uniquely identified.

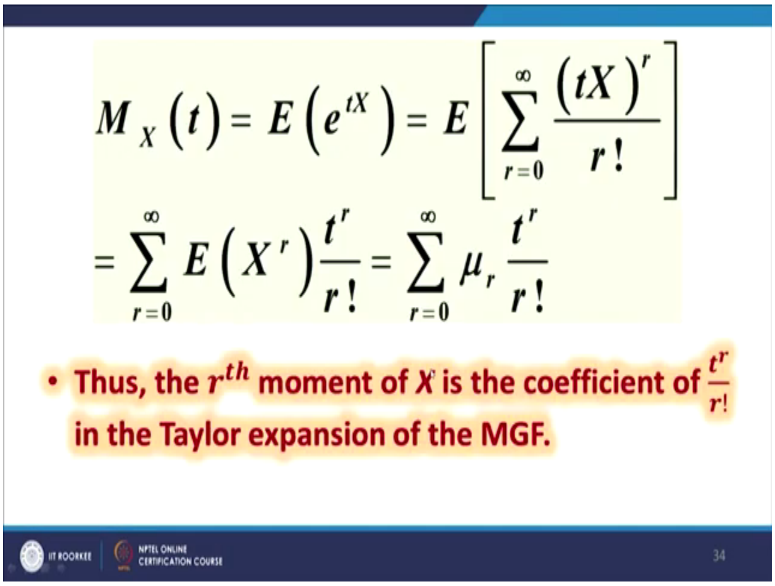
(Refer Slide Time: 26:06)




Now, therefore, the concept of moments is very important; just like we had the probability generating function, we have the moment generating function where by a similar trick, we are able to generate the moments of any distribution. For example, the moment generating function is defined by M X of t is equal to the expected value of E to the power; that is the exponential tX, where X is the random variable, t is parameter.

In the case of continuous distribution, it will take the form of an integral minus infinity to infinity; E to the power tx, p; x and dx.

(Refer Slide Time: 26:50)

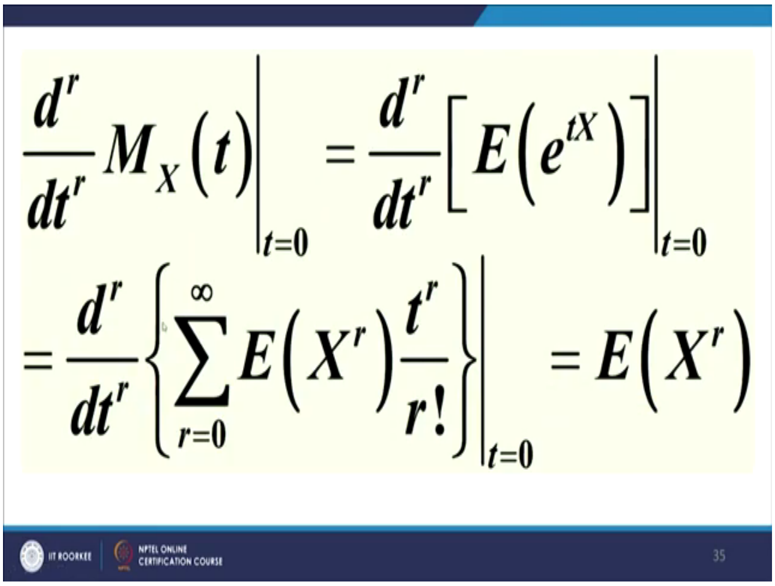



Now, if I expand this M; X to the power t, e to the power tX, if I expand the exponential as an exponential series, 1 plus X plus X square upon 2 factorial and so on, I get this expression that is within the square brackets.

And because the expectation of a sum is equal to the sum of the expectations, I can write this; in this form summation of; I can take the summation outside the expectation or other way or equivalently, I can take the expectation inside the summation and I can write it as summation E; X to the power r because t is a parameter, t is not a random variable; so t has nothing to do with the expectation.

So, t to the power r; r factorial remain as it is; they do not contribute in any form to the expectation and E; X to the power r is nothing, but mu r, where this is the rth moment about the origin. Therefore, I can say that the rth moment is the coefficient of t to the power r upon r factorial in the Taylor expansion of the Moment Generating Function.

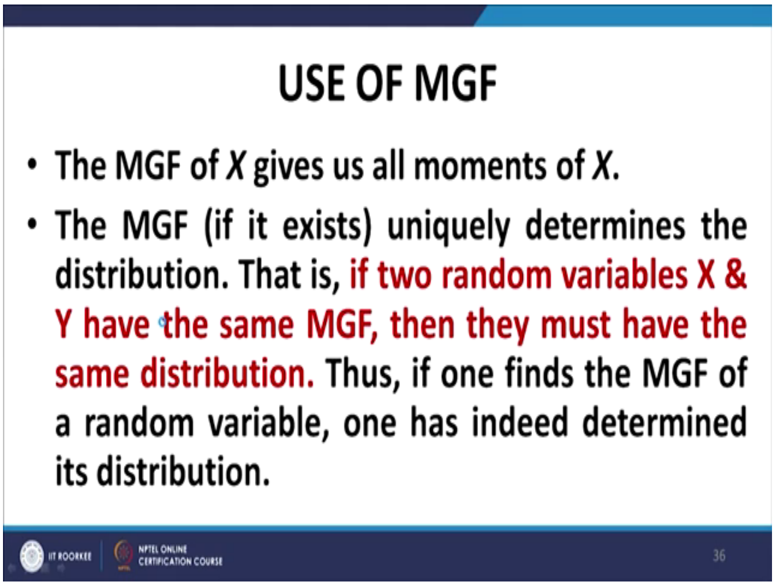
(Refer Slide Time: 28:06)




There is another approach to this, you can workout the derivative, as we did in the case of probability generating function workout the rth derivative of M; M X; t with respect to t, substitute t equal to 0 and you end up with E; X to the power r, as you can see on the slide.

Simply straight forward differentiation, after you are expanding the; after performing the steps, we had on the previous slide; you differentiate it, you get rid of the; the more the number of times you differentiate, the more the number of tr's that become constant or become 0. And for the remaining t's, when you substitute t equal to 0, they also go out. So, there is only one term that will remain and that will be E; X to the power r.

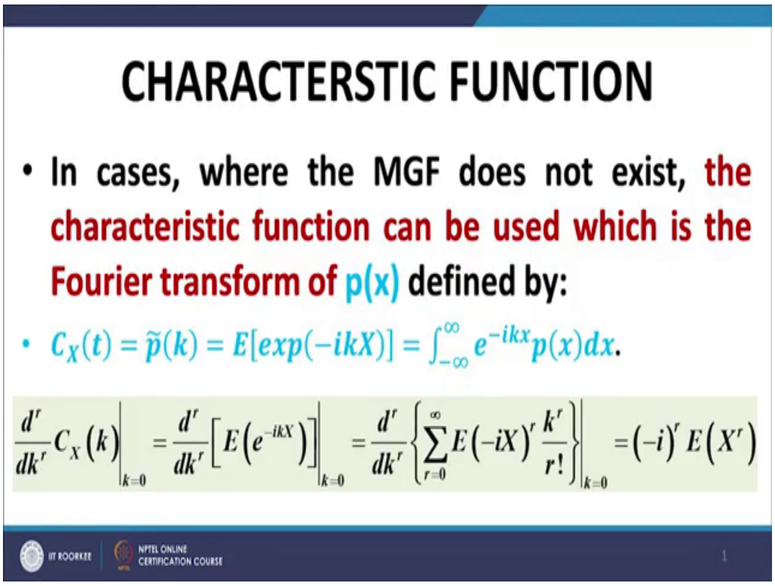
(Refer Slide Time: 29:11)




Now, applications of moment generating function it gives us; obviously, it gives us all the moments of X and then as I mentioned, the moment generating function uniquely determines the distribution.

Therefore, if two random variables X and Y; if two random variables X and Y have the same moment generating functions, then they are equal as far as the distributions are concerned. In other words, they have the same distributions; obviously, the equality or the similarity is confined to the distributions. As I mentioned earlier, two random variables can have the same distribution, but they can have different associations with other random variables.

(Refer Slide Time: 29:51)

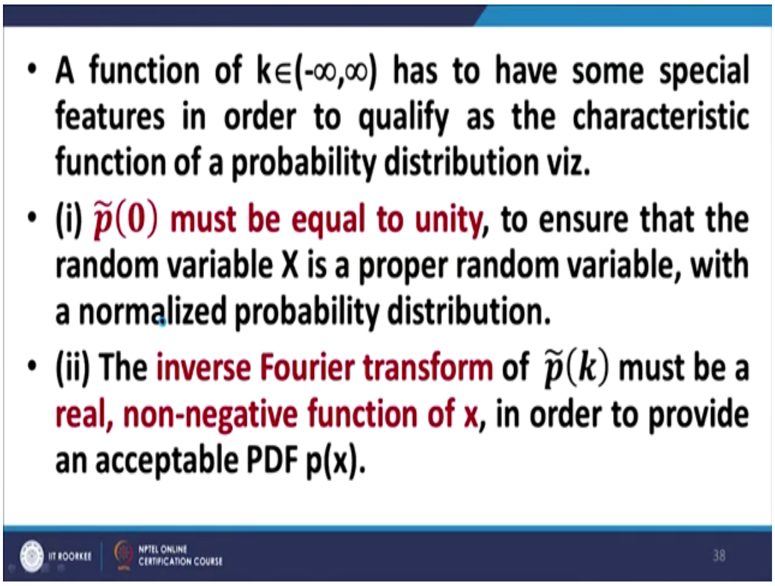



Now, concept which is parallel to the moment generating function is called the characteristic function. It is simply Fourier transform. And therefore, it is given by P tilde of k is equal to exponential minus ikX. Remember, when we talked about the moment generating function; it was the expected value of E to the E to the power t X.

Here, we are having exponential minus itX; so the characteristic function is the Fourier transform and the rest is absolutely same ah. Because of this minus i, this additional factor minus i erupts in when we workout the; when we workout the derivatives and we equate k equal to 0, we have to take care of this additional factor of minus i; otherwise, it is the same thing.

So, the rth derivative of the characteristic function will give me minus i to the power r; E; X to the power r, this additional factor minus i to the power r needs to be taken care of. Now, there are certain fundamental properties that need to be satisfied by a Fourier transform, in order that it may be regarded as a characteristic function of a probability distribution.

(Refer Slide Time: 31:26)




First of all, p tilde of 0; k equal to 0 must be equal to unity to ensure normalization. In order that we are able to define a normalized probability distribution p tilde of 0 has to be 0; has to be equal to 1, I am sorry; p tilde of 0 has to be equal to 1. And the second thing is that the inverse Fourier transform of p tilde k, must be real non negative function of x.

And therefore, when we workout the inverse of this Fourier transform; we need to have or we need to return quantities which can be identified with probability distribution. And therefore, they must have real; must be real non-negative functions of x. We will continue from here after the break.

Thank you.