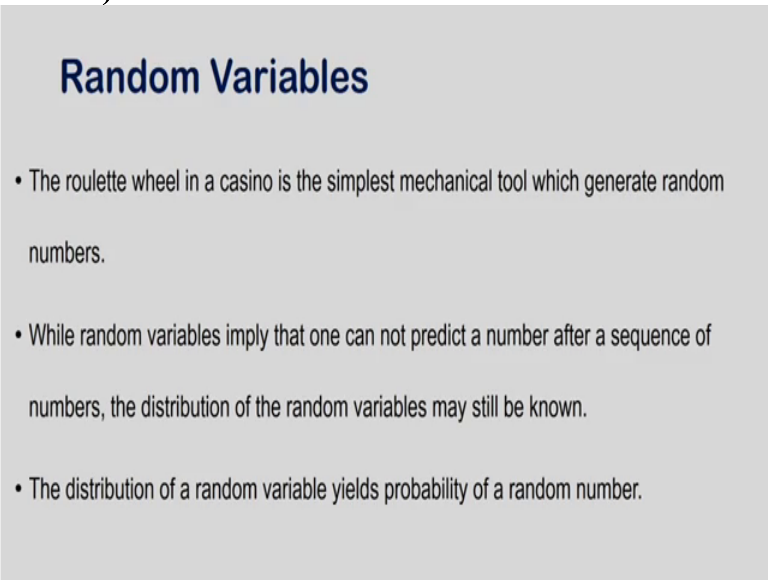
Numerical Methods and Simulation
Techniques for Scientists And Engineers
Saurabh Basu
Department of Physics
Indian Institute of Technology- Guwahati

Lecture 19
Details of the Monte Carlo method

So let us get ahead with this Monte Carlo technique that we have been discussing. And there are a number of mathematical things or concepts that we want to review to begin with. And this will form the initial part of our discussion today. So, we need to learn about random variables. So what are random variables, if we want sequence of numbers, where the predictability of you know, the next number would not be there is called a random variable set of random variables.

So, that if we want to choose a number, or rather want to predict a number after a given sequence of numbers, that is, that cannot be done, so, it is completely random.

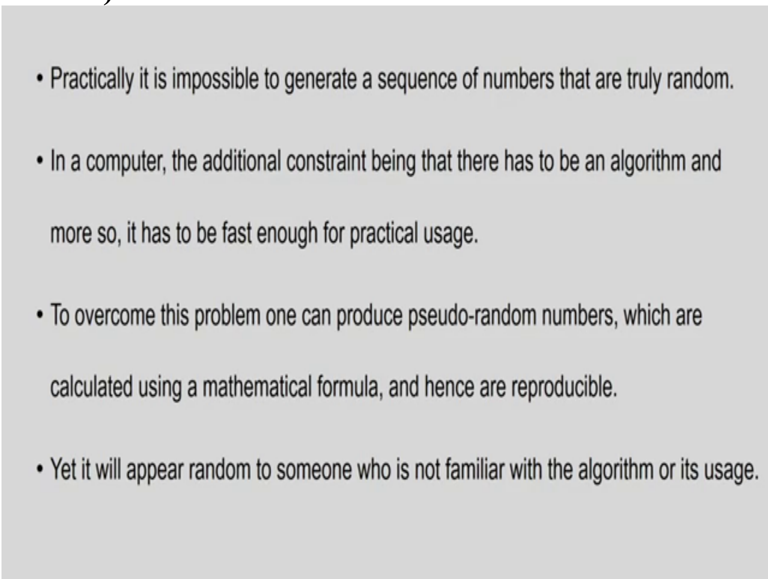
(Refer Slide Time: 01:23)



So, we know, the roulette wheel, the wheel that you have seen in casinos, is the simplest mechanical tool that can generate random numbers, okay. And so, this will be completely random, with each turn of the wheel, 1 would get a different number. And while of course, we know that the random variables imply that 1 cannot predict a number after a sequence of numbers, the distribution of the random variables may still be known, ok. So the distribution the region or the way they are distributed from some A to some B, or from 0 to 1, that could still be known.

And so there is a property of the random variable that the distributions are often known and so the distribution of random variable eels the probability of a random number, okay, so, if you want to know what is the probability of getting a random numbers, that information is supplied by the distribution of the random variable?

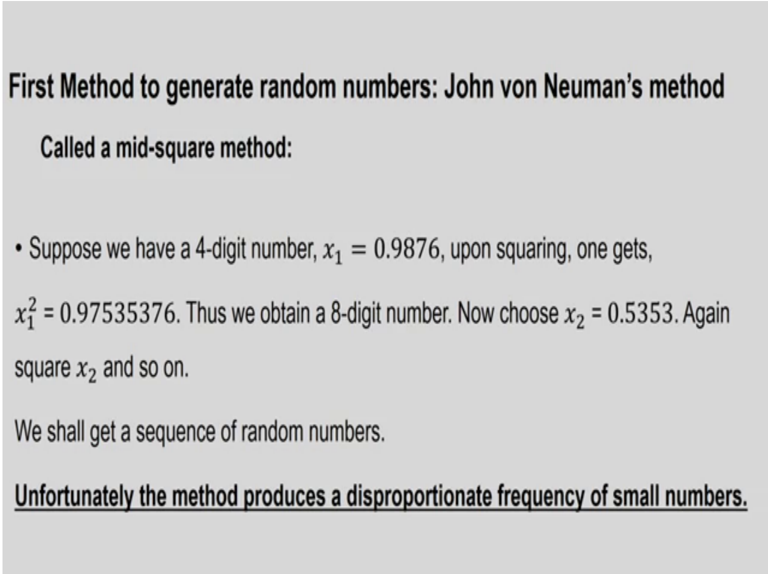
(Refer Slide Time: 02:36)



So, we understand that practically, it is not possible to generate a sequence of run numbers, which are completely random, which are truly random. The reason being that, that in computer there are certain algorithms which produce these random numbers, okay. And these things are often you know, follow since they follow a certain a lot of them, they cannot predict it. And on top of that, the additional constraint would be the usage has to be fast enough for you know, practical applications, okay.

So they have to churn out the numbers random numbers fast enough, so that the algorithm should not be too complicated. So to overcome this problem, one can actually produce pseudo random numbers, which are calculated using a mathematical formula. And since they are coming out emerging from a mathematical formula, they are bound to be reproducible. But of course, it will appear random to someone who is either not familiar with the algorithm, or that it has been used at all, for, you know, him or her, it will appear as a random variable, or completely random.

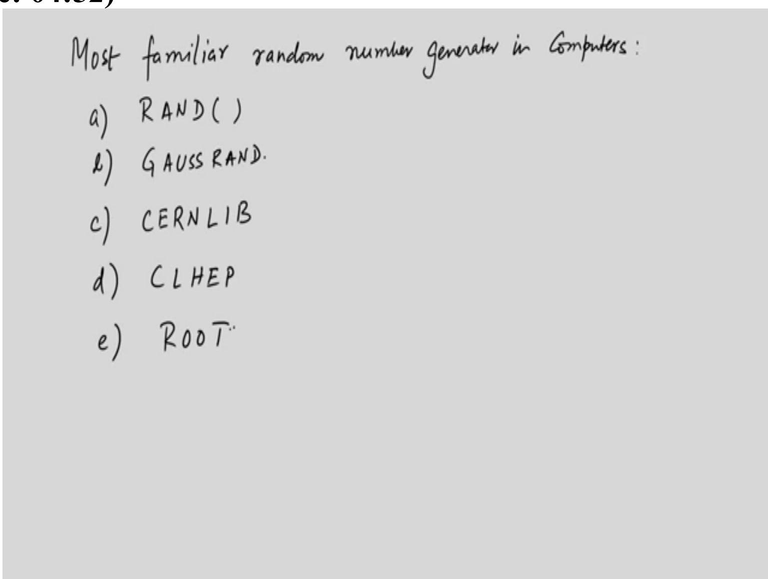
(Refer Slide Time: 03:57)



So the first method to generate random numbers was by this John Von Norman, who we are familiar with the boundary condition. So this call is a mid square method. And how it is done is that suppose we have a 4 digit number, say it is equal to x 1 equal to 0.9876 so upon, squaring one gets x 1 square equal to 97535376, that is from a 4 digit number, we actually come to an 8 digit number, we arrive at an 8 digit number.

And now take the middle 4 digits, that is 5353 and then again, square x 2, and so on. So we will get a sequence of random numbers. So unfortunately, what happens is that this method produces a disproportionate frequency of smaller numbers. Okay.

(Refer Slide Time: 04:52)

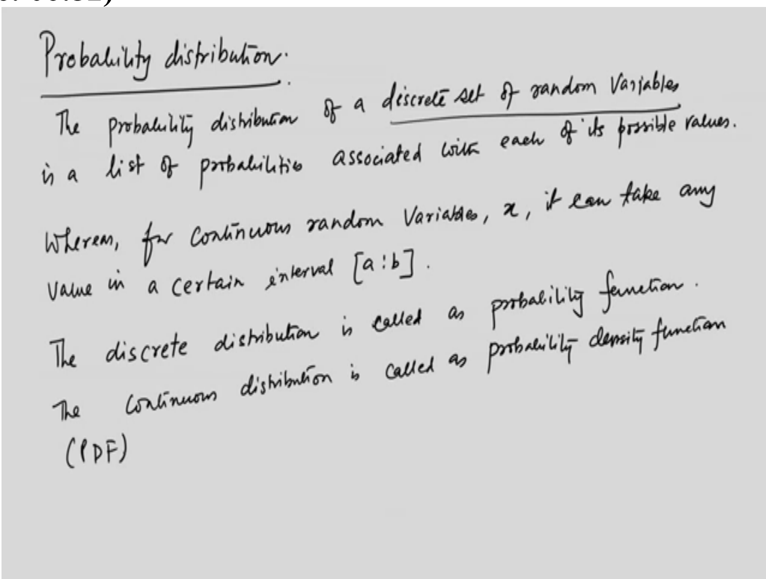


In computer, there are a number of algorithms. So the familiar most familiar random number generator in computers, they are a number of them are there, they are called Rand. So these are library, which, you know, usually comes with a bracket. So these, these are libraries which churn out random numbers, you can give the start point, and then point you can give the variants and things like that will learn what the variants etc are.

Then there is something called the Gauss Rand. Then from the CERN lab is called CERN, labs, which is the CERN library, then there is something called a CLHEP, HEP is for high energy physics. And then there are roots and so on. Okay. So most of them give you random numbers with varying degree of speed, and of course, random properties. Okay. So these are a preliminary discussion of how 1 can get random numbers.

Let us now talk about the probability distribution. As we said, that, even if a number is a random number one can the probability of those random numbers are or the probability distribution of those random numbers can be known.

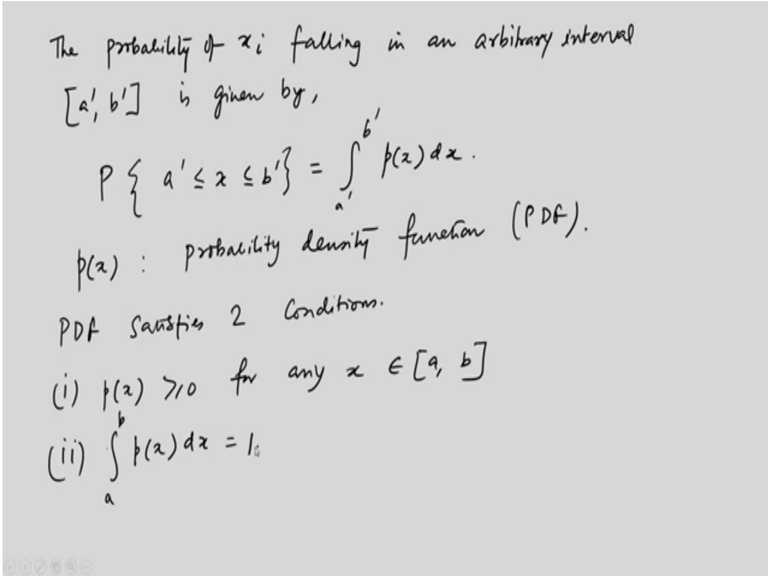
(Refer Slide Time: 06:52)



So let us talk about the probability distribution okay. The probability distribution of discrete set of random variables is a list of probabilities associated with each of its possible values okay. So, this is for a discrete set of random variables. Whereas for a continuous set of random variables or a continuous say x, it can take it can take any value in a certain interval, a, b the discrete distribution, which is this above 1 is called the probability function.

And; whereas, the continuous distribution, which is often going to be used by us as probability density function, or in short, we can call it a PDF, the probability of x i, let us write it in the front page.

(Refer Slide Time: 10:17)

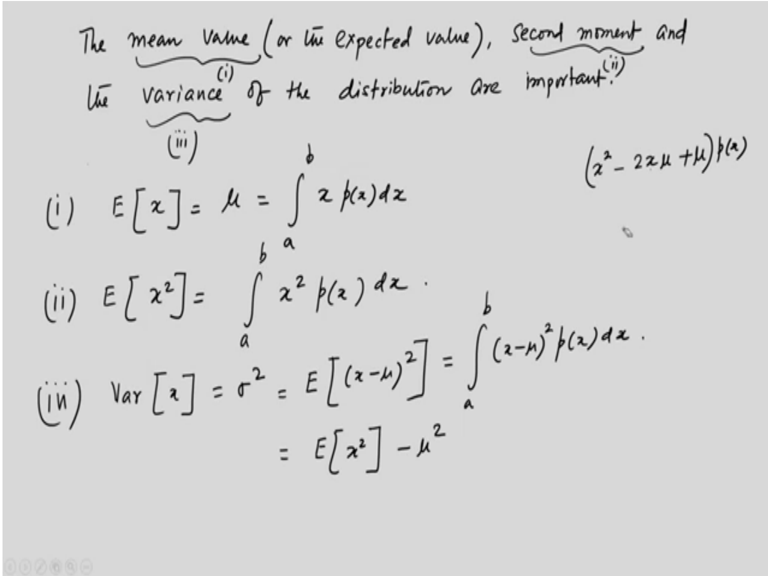


The probability of xi falling in an arbitrary interval, a prime and b prime, so, for example, is given by P, which is a prime less than x less than b prime, it is equal to a prime b prime pxdx, where p of x is called the, this is the probability density function. That is the PDF. Okay. All right. So, the PDF's satisfies 2 conditions. One is that it is positive definite for any x in the range, ab. And second is that it is a normalized district distribution. Okay, so these are going to be a little mathematical.

But what it says is that we are interested in a random number sequence of random numbers, set of random numbers. And these random numbers, they have a distribution. And for a continuous case, which we mostly be interested in this distribution is called as the probability density function or it is called a PDF. And this PDF has a property that it is equal to a positive definite in an interval for x to be x is a random number to be in in an interval a and b.

And it is also normalized, which is the integral of pxdx between a and b should become equal to one. Now, the various things that are interesting in this distribution and which are use will be used in your Monte Carlo technique or the simulation.

(Refer Slide Time: 13:00)



So, the first one is called as a mean value. And more often than not, it is called as the expectation value or expected value. Okay, the second moment of the distribution will define what that is, and the variance. So these are important quantities of this PDF. So once again, just to remind you, the PDF is the distribution of the random variable x, that is called px. So, because the first one so, let us this one, let us call this as the first one, this or this, this is the second one and this is the third one.

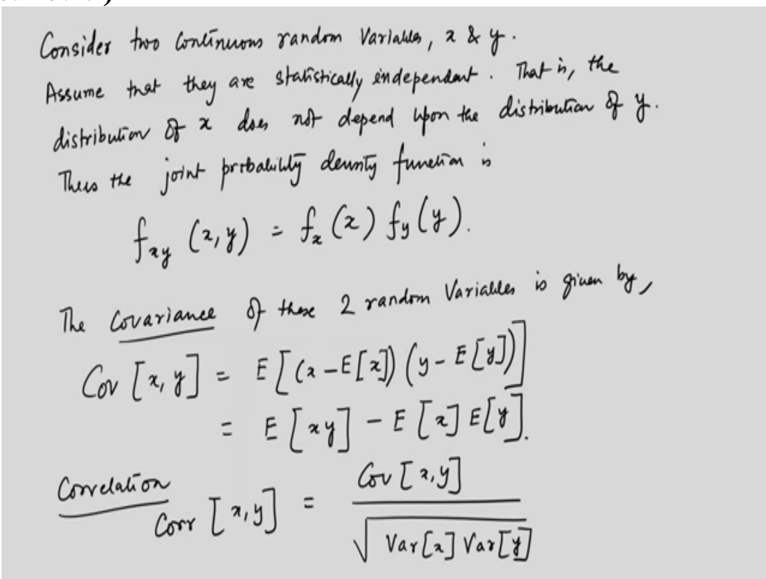
So, these are going to be defined by us now. So, the first one is written as this expected value of x is called as the mean and that is equal to a to b x px dx those who know waiting or calculation of the center of mass you are familiar with any way in your classical mechanics or some elementary physics scores, this is the way one actually calculates according to the weight and x is the distance from some chosen origin of given mass and then of course, this has to be divided by some quantity which is the total mass of the system.

But this is the first moment or here it is called assemble value or the expected value of the distribution. Similarly, the second one is called a second moment which is equal to the so this is equal to a to b and you have x square p of x dx and the third one which is called as a variance will write it with a ver of x, which is denoted by sigma square this is equal to this and then you have x - mu whole square Mu being the mean.

And this is defined as x - mu whole square p of x dx and so on. And this can be simplified as E of x square - mu square because, this term which is if you expand this, you will get a term which is x square then there is a term which is 2x Mu and then there is a term which is Mu square okay. So, these terms when you take you know, when you multiply it by the by the px, so, the first term is excess square px and then the term which is here will give you another x square here.

Because this x px and then the multiplied by the Mu and so, that is equal to your Mu square, and then there is a Mu here we cancel and then we will get a - Mu squared here okay. So, this is how you can expand it yourself and see this and this is equal to so, E expected value of x squared - Mu square and sigma is called as a standard deviation or the variance as I told here I mean the sigma is called as a standard deviation and square of that is called as a variance.

(Refer Slide Time: 17:29)

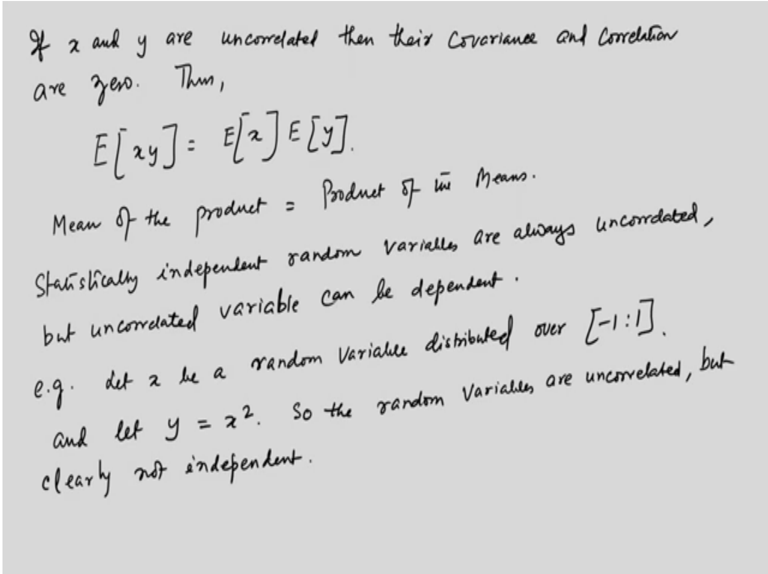


Moving ahead with the discussion, so considered to contain continuous random variables x and y okay assume that they are statistically independent. So, what we mean by that is so, that is the distribution of x does not depend upon the distribution of y okay and of course, vice versa, okay. Same with the distribution of y also does not depend upon the distribution of x. So, does the joint probability density function is fxy xy equal to fx x fy y okay.

So, these are 2 PDFs probably 2 distribution functions and the covariance of these 2 random variables is given by; so, we write it with the cov xy it is equal to E x - E of x and this will be there later. So, it is y - E then it is a y this and then this okay. So, this is the definition of covariance. So, this is equal to E xy – Ex Ey okay. So, this is the meaning of covariance and also the correlation between these 2 random variables.

So, this is given by So, corr xy which is equal to covariance xy divided by the variance of x and the variance of y okay. So, these are the definitions of the 2 random variables and the joint probability density function is given by this.

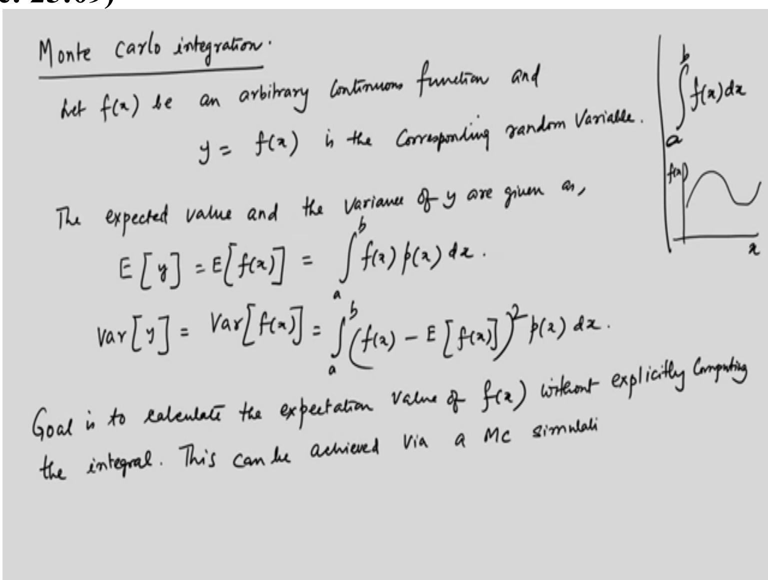
(Refer Slide Time: 21:38)



And of course, as we said if x and y are uncorrelated then their convergence or rather their covariance convergence that covariance and the correlation are automatically 0 which gives e xy it is equal to Ex as I said E is the expected value and E y, so, what it means is that a mean of the product mean or the expected value of the product equal to product of the mean okay. So, statistically independent random variables are always uncorrelated on correlated but uncorrelated variables random variables that is can still be dependent okay.

So, to give you an example here is that let x be a random variable distributed over - 1 and 1 and let y be another random variable such that y is equal to x square. So, the random variables are uncorrelated but clearly not independent they are of course dependent by this relation y equal to x square okay. So, these are some of the properties of this random variable and the definitions of these aspects value and the variants and so on. And let us now talk about the AMC integration and what are they its relationship with these.

(Refer Slide Time: 25:09)



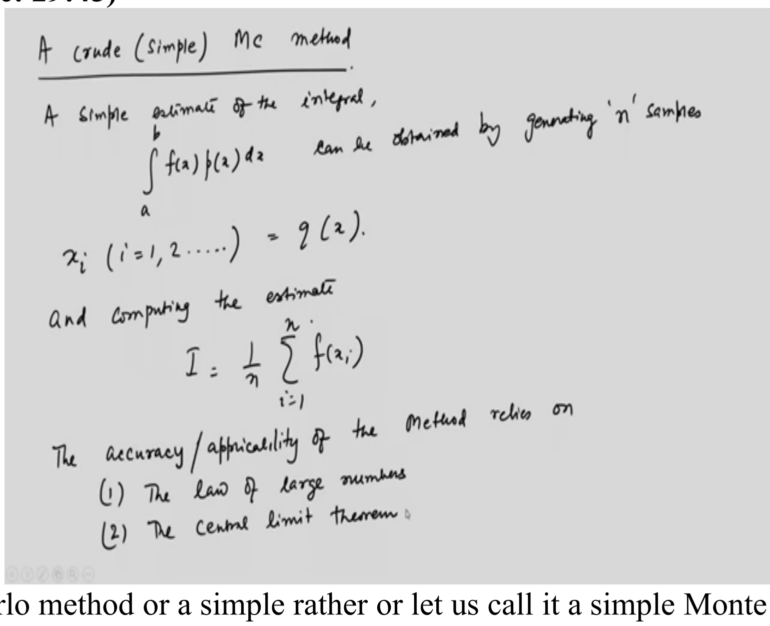
So, let fx be an arbitrary continuous function and y equal to fx is the corresponding random variable. We have seen that the expected value and the variance of y are given as Ey equal to E f of x equal to a to b f of x px dx of course, we have defined that px is the probability density function and the variance of y is a variance of f of x which is equal to a to b fx - E fx square px dx okay. So, this is we already know about it just cast it in a slightly different form. So, our goal is to calculate the expectation value of fx without explicitly computing the integral okay.

This is important to note that we also have done integration in which we have a function we need to integrate this function fx dx between a to b and are some BTQ and things like that okay. So, this is that function that we had said this is the function f of x and this is x and we need to actually find out the area under this curve. And in order to find the area under the curve we have divided the entire region that we have to integrate over into various grades of equal size.

And we have calculated the area of all those grades and have some them over according to certain formula okay. And we have seen that you know the Newton's or other The Simpsons 1/3 rule and the Simpsons 3/8 rules etc or there are Rumba formula and other formula which have with very degree accuracy of computed the integral. Here instead of doing that, we are taking random points and these random the distribution corresponding to these random points of the random variables are being used in order to compute the value of the integral.

And then we are of course, we are saying that will calculate the expectation value without explicitly performing the integral and this can be achieved via MC simulation how we let us see that.

(Refer Slide Time: 29:45)

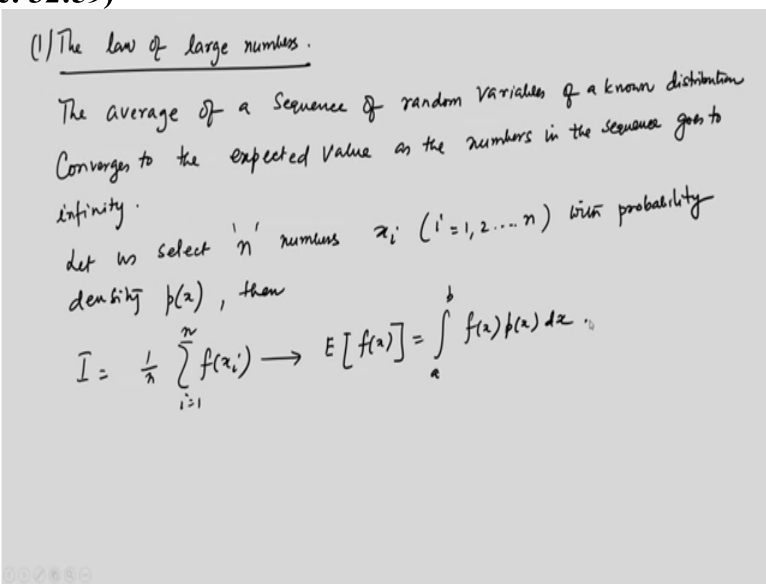


Crude Monte Carlo method or a simple rather or let us call it a simple Monte Carlo method, we call it MC in short. So, simple estimate of the integral which we want to perform is a to b fx px dx can be obtained by generating n samples such that you know xi from i equal to 1 to 2 to n and this is xi equal to say q of x and computing the estimate okay. So, what we are saying is that take some random variables or random points between a and b and calculate the values at those you know random points.

And then some all of them up and divided by the number of points that you have taken and then these will give you the value of the integral. So, we are not calculating the expected value of f of x, but using this summation we are going to get the value of the integrity. And this accuracy of the method or the utility of the method at depends upon 2 important concepts in mathematics and which we are going to discuss now.

So, whether you say accuracy or you say applicability. So, of the method or this claim for understanding that we will have to talk about 2 things one is called is the law of large numbers. And the second thing is called as a central limit theorem will mostly discuss the second one, but the first one is also important.

(Refer Slide Time: 32:59)

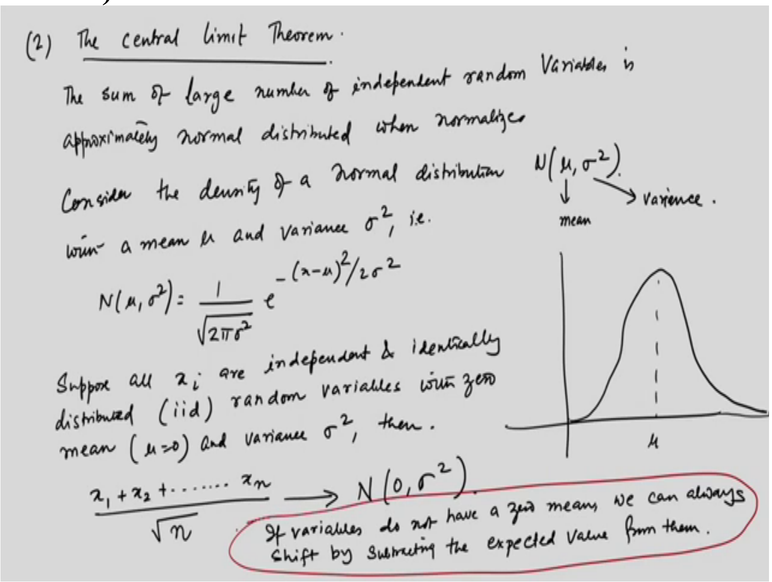


So, you let us write down the law of large numbers. So, what it says is that, so, there is the one. So, the average or the mean of a sequence of random variables of a known distribution converges to the expected value as the numbers in the sequence goes to infinity. So, let us select the numbers and numbers say xi equal to 1 to 2 to n with probability density px then i equal 1 by n f of x i equal to 1 to n this tends to the expected value f of x this is equal to a to b fx px dx this is exactly what is written in the simple or the crude estimate that we have written.

So, what it means is that if you have a number of or a sequence of random variables the average of that so, by taking the average means, we sort of sum them up all of them up and divided by the number of the numbers in that set this is equal to the expected value which means, the expected value is defined as the number or the you know the distribution multiplied by this function that we are talking about.

So, this in the limit of large n so, this is just simply calculating So, take a sequence of random numbers calculate the function at those you know random numbers and then take an algebraic sum. And these algebraic sum will converge to the expected value which we have defined with the probability density function of those random variables. And this will happen in the limit of when n is very large okay.

(Refer Slide Time: 36:47)



And importantly the second one which is called as the central limit theorem okay, the sum of large number of independent random variables is approximately normally distributed when normalized, let me tell you what it means is that. So, consider the density of normal distribution N mu sigma squared, this is how it is defined. So, the mu is the mean and this is the variance.

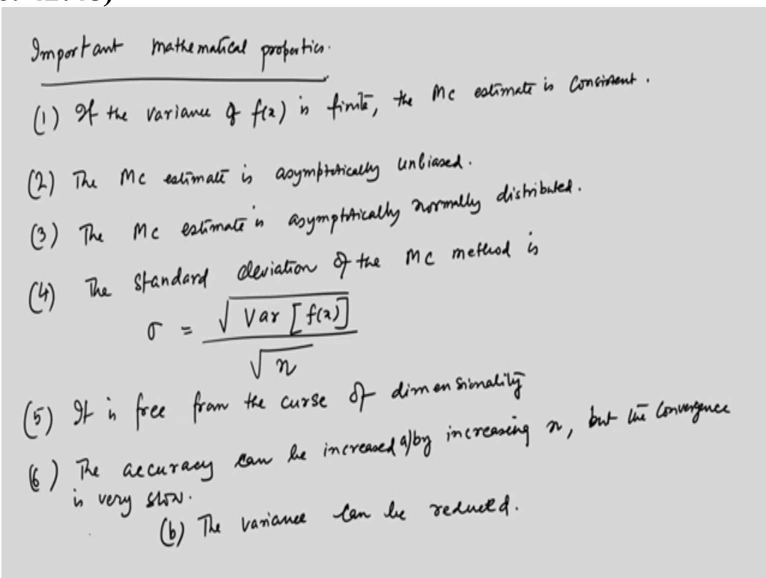
This is how normal distribution so, a normal distribution looks like this okay. So, this is your normal distribution with this you know being the Mu and full with that half maximum will give you a measure of the variance. So, this one with mean Mu and variance sigma square that is and Mu sigma square it is equal to a 1 by 2 pi sigma square exponential x - Mu whole square by 2 sigma square. So, suppose all x i's are independent all these random variables are independent and identically distributed.

So, they are called IID's identical independent and identical a distributed okay so, this was the sentence does not make any sense distributed random variables with 0 mean so, that is Mu equal to 0 does not matter, but you can additionally impose this constraint that Mu equal to 0 which means this shifting the normalized this normal distribution on the along the x axis and this is and of course, variance sigma squared.

Then what it means the central limit theorem is that x 1 + x 2 + xn this divided by 2 power of n this is it a convergence into the normal distribution for N to be large as 0 and sigma squared. So, this is the definition of or rather the statement of the central limit theorem. So, all these x's are random variables. So, the average of all these are some of all these random variables divided by this root over n is a normal distribution for him to be large, okay, with of course 0 mean but it does not matter.

I mean, if these variables do not have 0 mean we can always shift by subtracting by subtracting the expected value from them, which is Mu of x which is the mean as you know. So, it just says that it is not important to have a 0 mean but just that we for the ease of definition, we can assume that.

(Refer Slide Time: 42:48)



So, the important mathematical properties that this MC technique has okay. So, if the variance of f of x is finite, the MC estimate is consistent, okay. 2 the MC estimate is asymptotically unbiased. The MC estimate is a asymptotically, normally distributed. The standard deviation of the MC method is given by so this is a variation variants of f of x divided by 2 over of n, it is free from the, let us say the curse of dimensionality that it does not matter which dimension you are talking about.

Six, the accuracy can be increased 2 ways, of course, one is increasing in the number of sample points by increasing n, but the convergence or the success is very slow. Convergence by increasing the sample slow, rather, it is more, you know, convenient to decrease. So, by so this is a and let us not call it 7, but let us call it as b with the variance can be reduced okay. So, these methods are called as the variance reducing techniques and this is what we are going to see in the next discussion.

So, these just to summarize very quickly is that, instead of choosing specific data points, in order to do an integral, we have chosen a set of random variables and these random the properties of these random variables are discussed. But even if the variables are random or pseudo random, does not matter, I mean, the computer will ultimately give you a pseudo random according to an algorithm.

But even with that pseudo random distribution, that distribution can be actually well known the distribution of those pseudo random numbers and these distribution is called as the probability density function. And we can calculate the expected value of these random numbers by you know, sort of waiting it with the probability function or we can calculate the variance and so on. So, if you have a function f of x, you can calculate the value of the function at all the discrete point's xi where xi are chosen randomly chosen from a given distribution.

And then 1 can actually sum all these things up and divided by the number of points that will give you the simple estimate of the value of the integral. So this is MC method. And then of course, we have gone and navigated around certain mathematical properties and they are of course, law of large numbers and the central limit theorem, which says that these are the properties of these random variables.

So, the law of large numbers of course, say that you know, that you can the approximation that we have taken that is the value of the integral is when the limit of large numbers would converge into the expected value of the function are the function calculated that those discrete points are these random points. And the central limit theorem says something very important, which says that if you have a set of random variables, the average of all these random variables they are approximately normally distributed.

So let us correct it is normally distributed when they are of course normalized, normalized means we have a normalized distribution. And basically, based on that, these important mathematical properties arise such as if the variance is finite, then the MC estimate is consistent. It is it is unbiased, it is free from dimensional consideration. It is asymptotically normally distributed standard deviation can be found.

And we want the standard deviation to be low of this of this method and for doing the standard deviation to be low, one can either increase end which is seem to have very slow convergence whereas the variance can be reduced and there are various techniques that we are going to discuss on that.