Numerical Methods and Simulation Techniques for Scientists And Engineers

Saurabh Basu
Department of Physics
Indian Institute of Technology- Guwahati

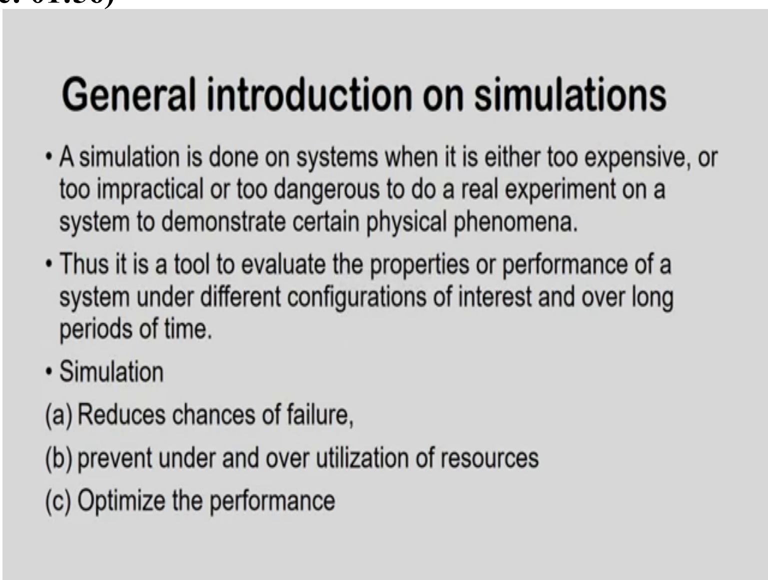
Lecture 18
Introduction to Monte Carlo technique

So, having looked at a number of numerical methods for solving problems related to science engineering and most of the time we have done problems that are of academic interest which are algebraic solutions or computing integrals etcetera of various functions. We now turn our attention to simulations in science and engineering. So, there are these common simulation methods the most popular ones in fact we are going to discuss.

So, just as a brief recap of the methods that numerical methods that we have seen one of them being the interpolation method we began with that and we also have looked at solution of non-linear equations. And then we have gone and calculated derivatives and then calculated the integrals integrations performed integrations. And then we have spent a good chunk of time dealing with the differential equations and the initial and the boundary value problems.

And now it is time for us to talk about the simulations, so simulation techniques. So, before we start with the simulation techniques let us see some common or rather a general introduction to simulations and when we do that and how we do that etcetera.

(Refer Slide Time: 01:56)



So, as a general interaction introduction to the problem simulation is done on systems when it is either too expensive, expensive in terms of either money or in terms of computational rather you know time consuming, so it could be either in terms of resources or time or it becomes too impractical or sometimes even too dangerous because if you are simulating a diffusion of bomb then it is a easier problem than actually trying to diffuse it with you know various parameters that are there.

So, it could be dangerous as well to do real experiment on a system to demonstrate certain physical phenomena so we want to understand certain physical properties and phenomena and that becomes a little difficult if we could be it could become very difficult if we want to do real experiments to understand that rather simulations provide easier ways to understand them. It is of course then becomes a tool to evaluate the properties or the performance of systems under different configurations of interest and it could be done over long periods of time so that we understand that there is whether a steady state has been achieved or not.

So, sometimes if we do not perform an experiment over a large period of time then this fact remains unknown whether the system has achieved steady state and is we are not looking at any transient phenomena. So, it is very important to run an experiment for a large you know amount of time or over long durations and so on. Simulation thereby reduces the chances of failure in actual experiments okay.

So if we know what are the precautions that need to be taken, so it becomes easier for us to understand the bottlenecks a priori before we actually perform the real experiment? And it prevents you know under and over utilization of resources sometimes without understanding we use too much of resources which may not be needed for this particular problem and this becomes clear while doing a simulation.

And of course we can optimize the performance of these exercise or the simulation and then there by you know directing the real experiments to optimize the parameters that are related to the best performance of the system.

(Refer Slide Time: 04:47)

Steps involved in a simulation process

- Identify the problem
- Formulate the problem
- Collect and process real system data
- Formulate and develop a model
- Validate the model
- Document model for future use
- Select appropriate experimental design
- Establish experimental condition for runs
- Perform simulation runs
- Interpret and present results
- Recommend further course of action

So, just a very quick recap of the steps that are involved in a simulation process, we need to identify if I the problem, we need to have a good idea of what we want to understand or what properties do you want to understand whether we want to see a phase transition, we want to see a scattering phenomena, we want to see a traffic flow and have a very good sort of understanding that what we are looking for what kind of properties or what kind of specific phenomena that we are looking for has to be understood.

Then it follows the formulation of the problem follows after that it is important that we formulate the problem correctly to have a you know a sort of unambiguous result and result that we has a merit. And then we have to collect and process real system data ok. So, these are important in the sense that the correct input has to be given and for that we actually can collect those data or sample those points either from a random distribution or a deterministic distribution or anything for that matter.

It is important that we sample these data input data properly then formulate and develop a model ok this is important that we know that the model that we are taking for a particular system for demonstrating a particular phenomena is a correct model. So, it is important to have the correct model for that we need to validate that model and this is of course important of validation of the model.

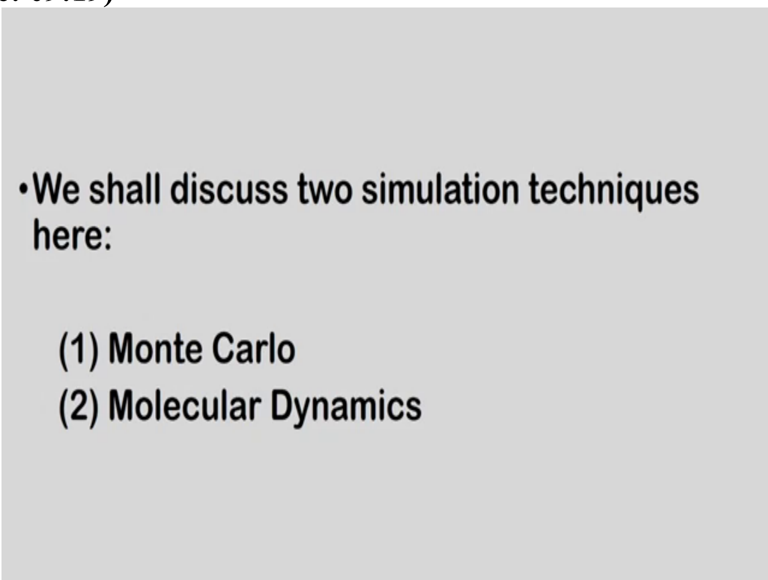
And then the documentation is very important for future use. So, sometimes when you write programs and which you may have to look after say year or so you have very little recollection of why you have done a particular thing or rather why you have done a particular step in a computer if you write documents if you write notes commenting those notes and then it is easier for not only you but someone else who's looking at the program to understand that what steps have been followed exactly.

Then select appropriate experimental design that design has to be selected which gives the right you know kind of system. Establish experimental condition for runs these are the parameters which would govern the final result. So, it is very important for us to know that what are the parameters that we the values that we set in order to get meaningful results. Suppose you are talking about a traffic flow in an urban city say like Delhi or Bombay and then you need to understand that which are the hot spots right, well a lot lots of vehicle actually move in a particular interval of time.

So those data will certainly differ from smaller cities where there are not too many vehicles flying or at least not the amount that that deploying Bombay or Delhi. So, these are the experimental data or rather the conditions that have to be said you know supplied for giving this simulation runs and this perform this simulation etc and interpret and present results. This is one of the most key points that simulations will give results finally if the coding etcetera if the software's etcetera all are correct they will give some data.

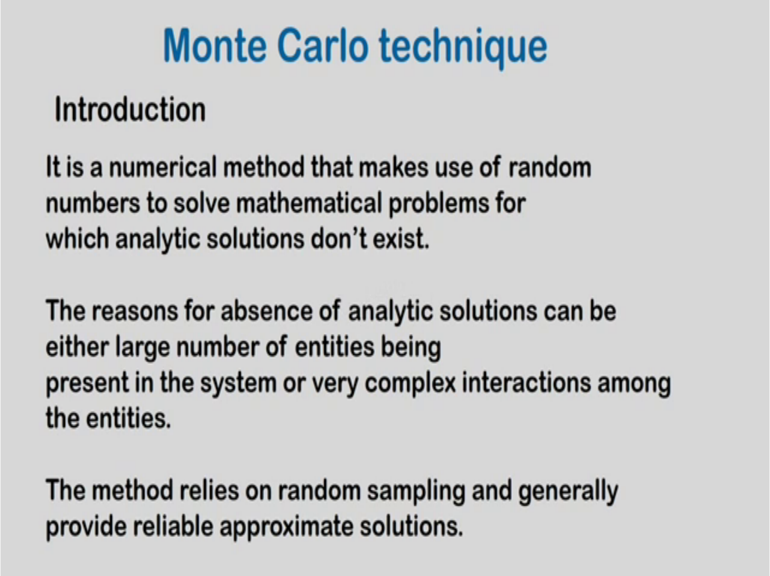
And these data have to be analyzed and without the analysis we would not understand that what they mean and these have to be presented by you know in the correct perspectives keeping in mind about the objectives that this program wants to deliver and hence you know a recommend for their course of action okay. So, these are the various steps that are involved in the simulation process.

(Refer Slide Time: 09:19)



So, we shall discuss mainly to simulation techniques I said that these are the most important ones or one is called as a Monte Carlo and the other is a molecular dynamics. So, these are the two techniques that we are going to use and so let us look at the first one of them.

(Refer Slide Time: 09:39)

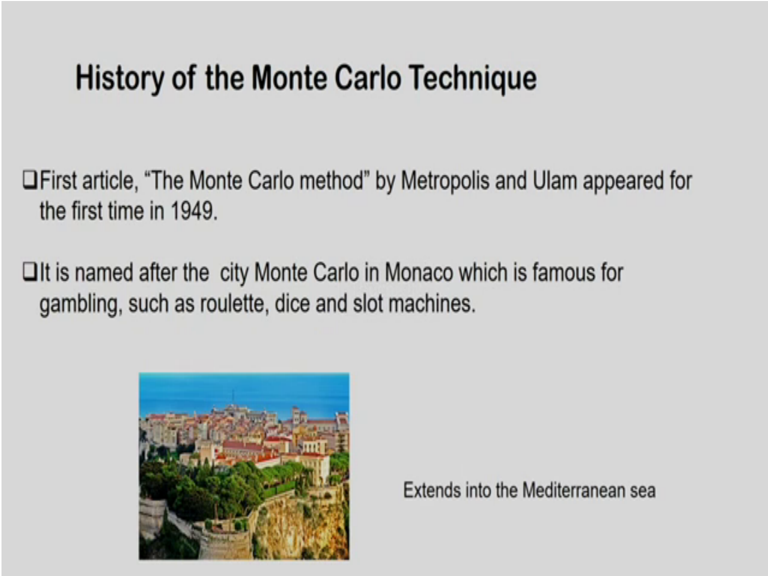


That is the Monte Carlo so let us talk about the Monte Carlo technique and it is got a nice you know introduction very interesting that the name itself came from the name of a city where gambling is important I will just come to that in a while. So, it is a numerical method so Monte Carlo technique is a numerical method that makes use of random numbers to solve mathematical problems for which analytic solutions to not exist.

This we had of course make sure that analytic solution to a particular problem does not exist because if it does then of course doing this thing would be not serving the purpose that it is intended for. And the reasons for the absence of analytic solutions can be can either be large number of entities being present in the system that is a large number of particles of CNG gas a real gas say for example or very complex interaction among the particles or among the entities.

So these are the things that inhibit having a analytics having an analytic solution for this problem and the simulation techniques become more and more important in these particular cases. This, particular method the Monte Carlo method relies on random sampling and generally provide reliable approximate solutions these solutions approach though they are approximate solutions yet you can rely upon them.

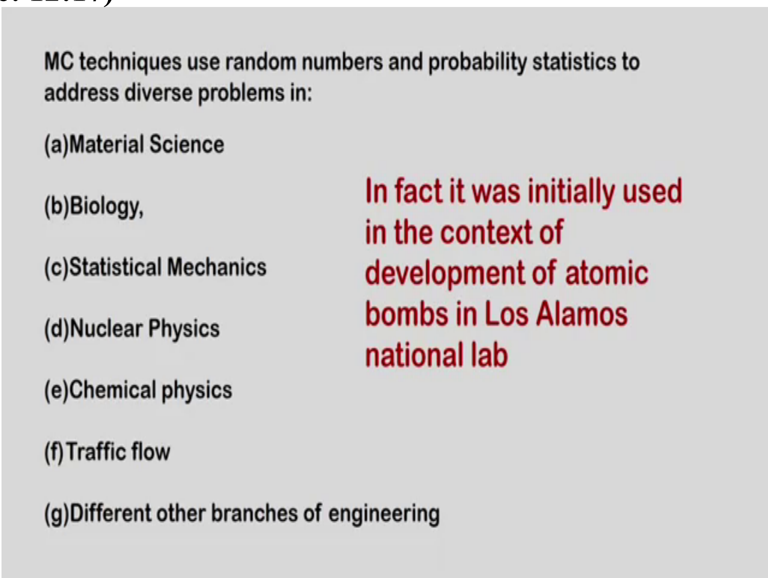
(Refer Slide Time: 11:15)



So, just a little bit on the history so the first article the Monte Carlo method by two people called metropolitan Coulomb appeared for the first time in 1949 and it is named after the city Monte Carlo in Monaco which is famous for gambling and all kinds of gambling such as because roulette dies slot machines etc some of you may be familiar in the context of a casino so these all these things are there.

It is a very nice City that it extends into the Mediterranean Sea and as these gambling are gambling's are quite you know this city is known for gambling so that is why this method because it is based on random you know inputs or we will see that as a random number generation so this is actually named after that city.

(Refer Slide Time: 12:17)



So Monte Carlo are techniques use random numbers and probability statistics to adverse, diverse problems in material science say a crystal growth etcetera. They have been used in order to understand that how a crystal when they are depositing particular say a thin film or something using some molecular beam epitaxy method or some other method. How the thickness of the film grows from a perfectly two-dimensional to a quasi two-dimensional and these are phenomena that will have to be you know seen as a function of time.

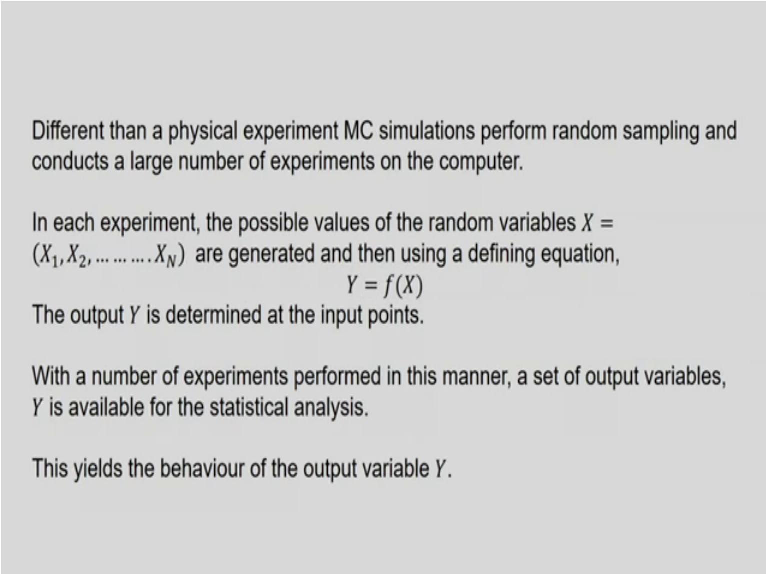
And so these are some of the important problems in material science in biology say in colony of bacteria how they are growing and you know you know how they can be contained in a in an effective manner. Statistical mechanics of the most you know common example is Icing model which we may discuss if we have time. How the phase transition in Icing model occurs and nuclear physics it is about particle collisions.

So, calculating this scattering cross-section of these collisions and thereby predicting the you know out coming particles etcetera. Chemical physics say in the petroleum industry and things like that so there are usage of this Monte Carlo method. As I told about traffic flow so one uses it in order to get a an outcome of the traffic you know as time progresses and how the traffic can be diffused effectively in urban cities.

So, basically all these; the traffic engineering which is a branch of civil engineering. Say for example that continuously you know evolves and tries to find out that what would cause less traffic jam in a city when there are very large number of traffic's vehicles applying especially during the office hours the peak office hours. So, these are and there are a large number of cases where Monte Carlo is used the list is of course not exhaustive in just a listing a few of them trying to give you an idea that why I mean how Monte Carlo is useful in various you know branches of science.

So, in fact it is initially used in the context of development of atomic bomb in the Los Alamos National Lab this was during the Second World War and this was one of the first applications.

(Refer Slide Time: 14:58)

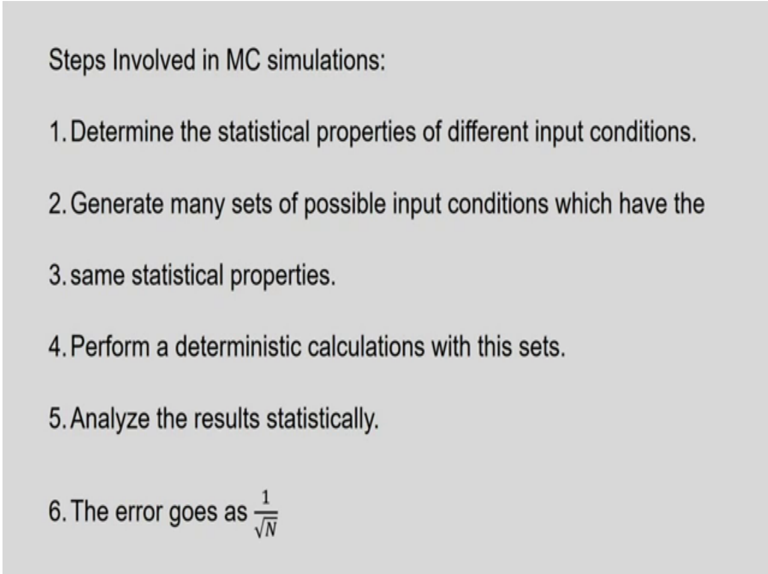


So, how does it differ from physical you know experiment. So, it differs from the physical experiment in the following sense that the MC simulations Monte Carlo is MC, MC simulations performed random sampling and conducts a large number of experiments in the computer. These experiments when there have to be repeated a number of times in the real sense it could be extremely expensive in terms of as I said in terms of both resources and time.

Whereas doing it in the computer really does not hurt much in the sense that not as much as with regard to doing real experiments, so in each experiments the possible values of the random numbers are given by the variable x which has which can take random values x1 x2 and so on and till xn are generated. And then using a defining equation y equal to f of x the output y is determined at all these input points y x 1 x 2 and so on.

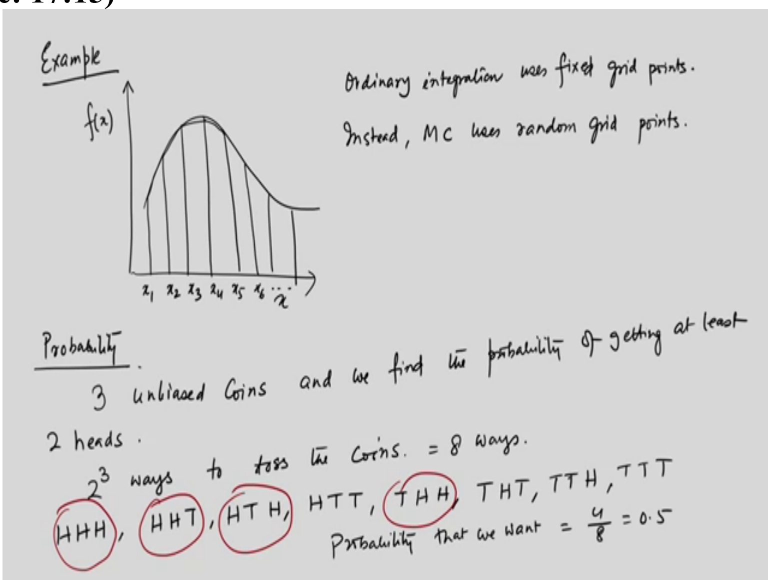
So with the number of experiments performed large number of experiments in fact performed in this manner a set of output variables y is available for the statistical analysis okay. So, e y is also you get a set of y values corresponding to these different sets of the input variables and once when we analyze it we get the behaviour of the output variable y. And this is what we want to understand this thing.

(Refer Slide Time: 16:41)



So, let us let us look at some of the steps involved in the MC simulations so determine the statistical properties of different input conditions, generate many sets of possible input conditions which have the same status same statistical properties is not. Number three they perform a deterministic calculation with this sets and analyze the results statistically and the error is found to go as 1 by root n.

(Refer Slide Time: 17:15)



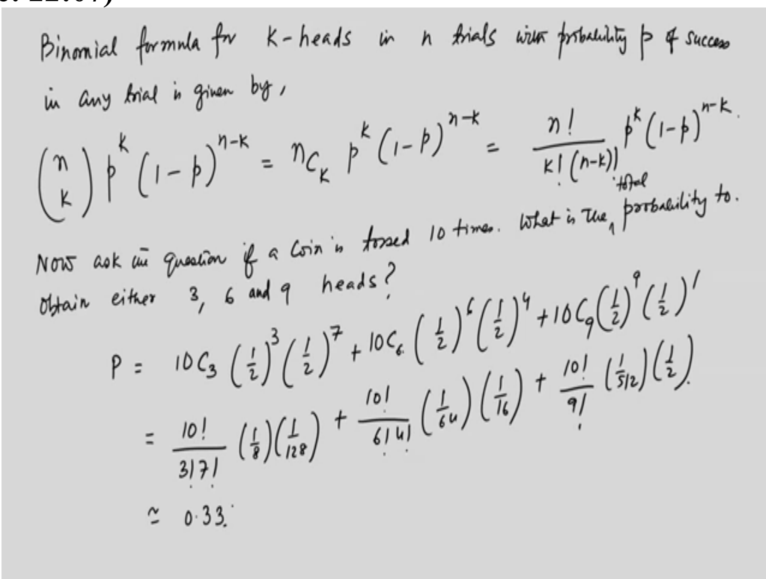
So, let me tell you giving some examples that what exactly is this Monte Carlo method. So, if we have a function like this it is mainly used Monte Carlo is mainly used to do integration. So, this is f of x and x if you have a function like this which you have to integrate which means that you have to find the area under the curve. What normal you know integration procedure does is that it converts it into some kind of rectangles and then calculate the area of each of the rectangle or the rhombus and then one can get the full area under the curve.

By summing up the areas of all these rectangles or this rhombus and there are various approximation that we have seen there. But in so what it does is that it actually creates fixed grid size so this is say x 1 and this is x 2 and this x 3 x 4 x 5 x 6 and so on okay. So, these are fixed grid points so ordinary integration so instead of that MC uses random grid points. Let us give so these are not chosen you know systematically and in a homogenous manner but there are random grid points that are chosen and the integral is carried over that.

And the probabilistic concepts that come in let us try to understand the probabilistic concepts okay. Suppose we have 3 unbiased coins and we have to find the probability of getting at least two heads. So, now you understand because there are three coins and there are 2 possibilities so there are two to the power three ways of doing this to toss the coins okay so which means equal to 8 ways what are the 8 ways all 3 coins could yield heads or both could two could yield H and one T then HTH HTT THH THT TTH and TTT okay.

So, there are these 1 2 3 4 5 6 7 8 now how many of them have at least 2 heads this one has this one has, this one has, this one has and so on. So, the probability that we asked for equal to 4 divided by 8 is equal to 0.5. So, there is a half probability of getting 2 heads for 3 unbiased coins okay.

(Refer Slide Time: 22:07)

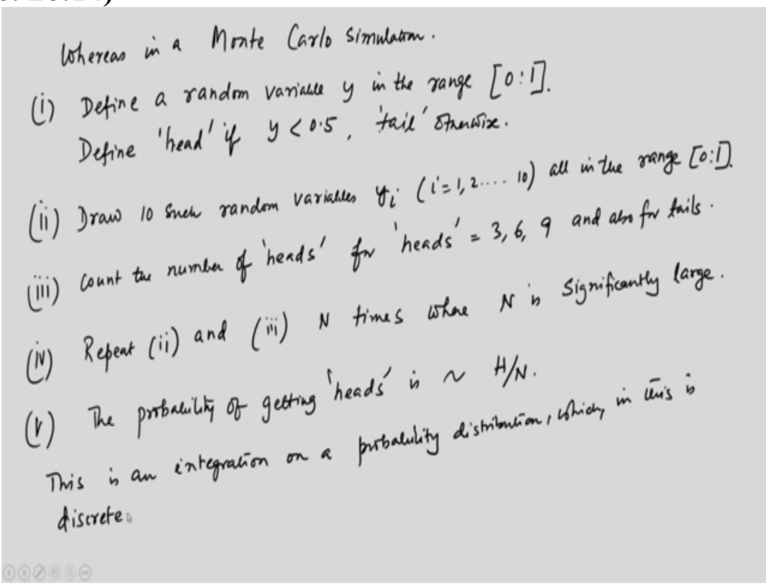


So, if you need to you know get into details of this problem so we need to understand a binomial formula. So, a binomial formula for K heads in n trials in trials means we are tossing it n times with probability P, P of success in any trial is given by n k p to the power k 1 - p to the power n minus k this is n C k P to the power k and 1 - P to the power n - k which is nothing but n factorial divided by k factorial n - k factorial and P to the power k 1 - P n – k.

Now ask the question if a coin is tossed 10 times what is the probability to obtain either 3 6 or 9 heads okay. So, the first 1 4 3 4 3 heads we have the probability equal to a 10 C 3 half cube and half to the power 7 + 10 C 5 okay. So, this is for the for the total probability. Let is calculate the total probability instead of calculating. So it is and 10 C 6 and 1/2 to the power 6 and 1/2 to the power 4 + 10 C 9 1/2 to the power 9 and 1/2 to the power 1. So, if you simplify this becomes equal to 10 factorial divided by 3 factorial 7 factorial this is 1 over 8 and 1 over 128 + 10 factorial divided by 6 factorial.

So we asked what is the total probability and 1 over 64 and 1 over 16 + 10 factorial by 9 factorial 1 by 5 1 to 1/2 and so if you simplify this, this comes out around 0.33. So, it is a 1/3 probability that is a 1/3 probability. So, this is we are very familiar in the class maybe + 2 level probability and statistics but what Monte Carlo does is the following.

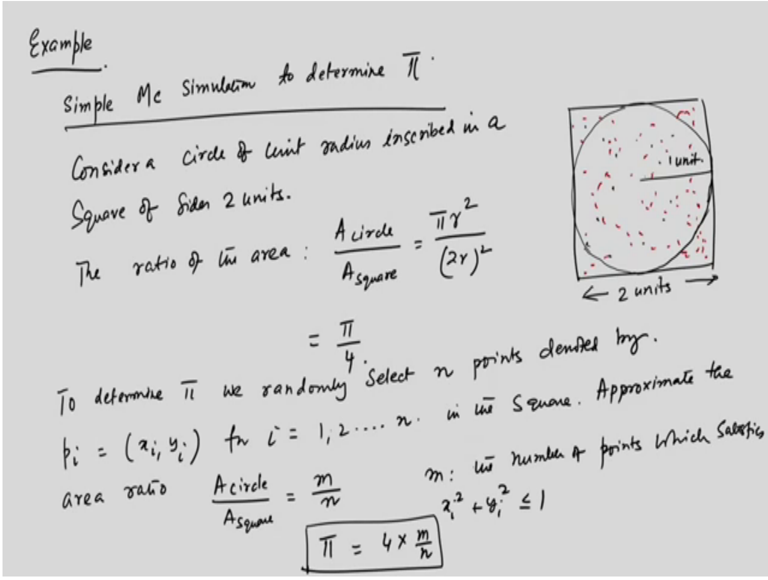
(Refer Slide Time: 26:14)



Whereas in Monte Carlo simulation one define a random variable y in the range 0 and 1 so then define so you pick this random variable from a random number generator which are abundantly available in any computer. So, define head if y is less than 0.5 tail otherwise. So, then what you do is draw 10 such random variables y i i equal to 1 to 10 all in the range 0 1 count the number of heads 4 heads equal to 3, 6 and 9 and also 4 tails.

If you like then as a fourth step repeat 2 and 3, n times where n is significantly large. The probability for getting heads is approximately H over N okay. So, this is an integration of probability that is a summation that we have written down in the last slide integration on a probability distribution which is which in this case is discreet. Let me make more concrete show more concrete example here.

(Refer Slide Time: 30:26)



And that will make things easier for you to understand, simple MC simulation to determine pi the value of pi okay which you know as 3.15 14 and all that. So, consider a circle of unit radius inscribed in a square of sides 2 units. So, this is the square of sides two units and excuse my freehand drawing of circle so this is 1 unit and this is 2 units okay. So, it is completely inscribed here so this is 1 unit.

So the interesting thing is that the ratio of the area of the circle and the square so A circle and A square so this is PI r square and this is 2r whole square. So, this becomes equal to PI r is equal to 1 so it is a PI over 4 so this ratio is PI over 4. Now you use a trial and error of throwing in dots on this square ok. So, we have let us draw the dots with a red ink. So, you have you throw a very large number of dots and keep a track of those dots that have been thrown here okay.

So count the number of dots so this to determine pi we are randomly select n points denoted by PI equal to x i y i for i equal to 1 to n of course n is large in the square and this approximate the area ratio, ratio of the area so A circle by A square equal to m by n this ratio would converge as root over n, but m is the number of points which lie within the circle which means we satisfies x i square + y i square is equal to less than equal to 1.

So, these M are the points that actually lie on the circle and m n is the total number of points M is only within the circle and n is the total number of points so pi can be evaluated using 4 into m by n now this number we know m and n both these numbers we know just multiplied by 4 which will give you the value of PI ok. We will see more examples of this MC technique.