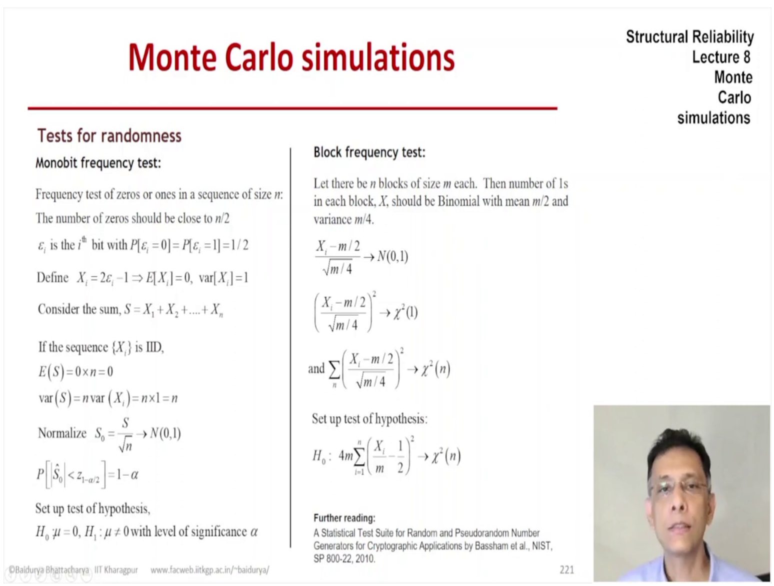
Structural Reliability
Prof. Baidurya Bhattacharya
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture –73
Monte Carlo Simulations (Part - 05)

Once a random number generator has been selected the user may want to test the generator for various properties because the accuracy of Monte Carlo simulations depends on the supply of a large number of IID random samples. So, um here we discuss few very basic and general tests of randomness they are by no means complete.

(Refer Slide Time: 00:59)



So, let us start with the idea that there is a sequence of zeros and ones these might be the outcome of some physical process like tossing a fair coin many times or one of the pseudo-random number generators working on a computer. So, if we have such a sequence we expect that the number of zeros or number of one's should be close to n by 2. Now let us formalize this. So, let epsilon i be the ith bit uh.

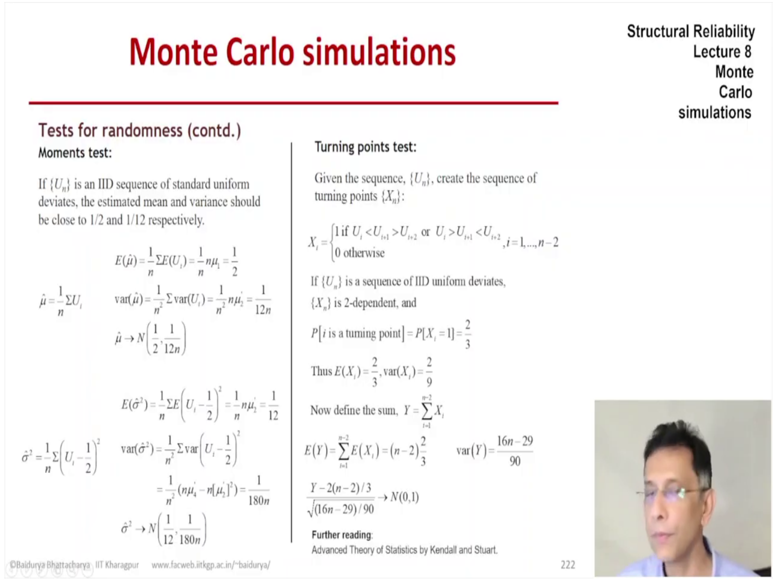
So, the probability of zero and 1 is each equal to half and then let us define a new random variable X which is twice epsilon minus 1. So, the mean of X you can work out is 0 and the variance of X is 1. Now let us consider the sum the sum of these n X's, n being the length of the sequence that we have and this sum S is also a random variable and um its mean is zero and if the sequence if all the X i's are independent then the variance of S would be the sum of all the variances.

So, the variance of S is n. So, um we can. Now set up a test of hypothesis. So, we normalize the sum by square root of n which reaches which approaches the standard normal distribution by central limit theorem and then we could set up the test that the estimated value of S naught would be between plus Z 1 minus alpha 2 n - Z 1 minus alpha 2 that would be our test of hypothesis with the level of significance alpha.

We could extend this idea to blocks of size m. So, if we have several such blocks each of size m. So, in each block the number of zeros of the number of ones if the epsilons are equally likely to be 0 and 1 would be the mean would be m by 2 and the variance would be m by 4. So, if if I define X to be the number of such zeros in each block then again by central limit theorem X normalized by mean and the standard deviation should reach the standard normal.

And here let us square each of these. So, I get the chi-square the chi-squared distribution with degree of freedom 1 and if I then sum all of them I get a chi-square random variable with n degrees of freedom and just like I did for the monobit frequency test I could also set up a hypothesis test for the block frequency. And there is a very nice report from NIST which discusses many such tests if you like you might want to refer to it.

(Refer Slide Time: 04:45)



We now take these ideas to the sequence of uniform deviates. So, they are no longer bits of zeros and ones that we are talking about we are talking about uniform deviates between 0 and 1. So, these are an IID sequence U 1 up to U n. So, one of the most intuitive tests would be that if this is indeed uniform then the mean of the sequence should be close to half and the variance should be 1 over 12 as we know.

So, we can estimate the mean and we can estimate the variance and each of these actually is a random variable. So, the mean of the estimated mean is half the variance of the estimated mean is 1 over 12 n as you see. So, we could set up the test of whether mu had the estimated mean is normal with mean half and variance 1 by 12 n. Likewise the estimated variance is also a random variable its mean is 1 over 12. It is variance can be computed to be 1 over 180 n and we could likewise set up another test of hypothesis for the estimated variance as a normal random variable with mean of 1 over 12 and a variance of 1 over 180 n.

Now obviously you probably sense a problem with this test because even if we have a true random a good random sequence which passes this tests of moments the mean and the variance. If I just rank ordered them that sequence would obviously not be random anymore by any stretch of imagination but they would still equally well satisfy the moment's test. So, obviously we need to see whether the sequence itself has the appearance all the required appearance of randomness.

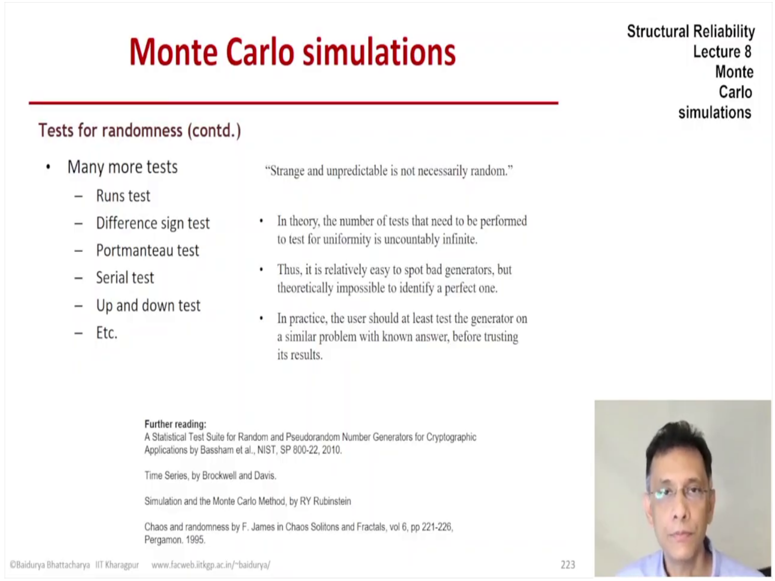
So, one very um interesting test is the turning point test. So, given the sequence U n we can see if there is a turning point or not. So, turning point means if the sequence first rises and then falls at that point or falls and then rises. So, the point i is a turning point X i equals one if U i + 1 is greater than both the number before it and after it or it is less than the number before it or after it.

So, we have n - 2 such points turning points from a sequence of size n and if this original sequence U is an IID sequence of uniform deviates then this turning point sequence is actually no longer an independent sequence it is 2 dependent and we can work out that that X i is a turning point it is the probability is 2 over 3. It basically comes from the property that if you if there are 6 possible ways of having a sequence of 3 and then the 4 of them are turning points and 2 of them are not.

So, that is how that 2 over 3 comes they are all equally likely and then. So, that the mean of X is 2 3rds and the variance is 2-3rds times one-third. So, that is 2 over 9 and if we define the sum of all these turning points all these n - 2 turning points then the mean of the sum the mean of y is n - 2 times 2 3rds the probability that X is one turning point the variance is little less than what would happen if they were all independent but as you know as I said the the X's are 2 dependent.

So the variance of Y the sum is 16n - 29 over 90.. So, it is it is less than n - 2 times 2 over 9 which would be if the sequences were completely independent. So, again invoking the central limit theorem the normalized value should approach the standard normal distribution and one could set up a test of hypothesis. A very nice a derivation in all of this is given in Kendall and Stewart's book.

(Refer Slide Time: 10:16)



So, there are obviously infinitely possible tests that one could do for randomness. Here we you have just a few mentioned on your screen the thing to remember is just because something is unpredictable it doesn't mean that they are necessarily random. So, one needs to look deeper and in theory it is much easier to test a bad generator a sequence that is not random than to satisfy oneself completely that a sequence is indeed random it is an IID sequence.

So, what is important is that for the application at hand one should test the generator and its output for accuracy. And preferably test it against some known solutions there are many excellent resources on this subject and I have listed a few of them and this they should give you a lot more information on tests of randomness.