

**Modeling & Simulation of Discrete Event Systems**  
**Dr. Pradeep K Jha**  
**Department of Mechanical and Industrial Engineering**  
**Indian Institute of Technology, Roorkee**

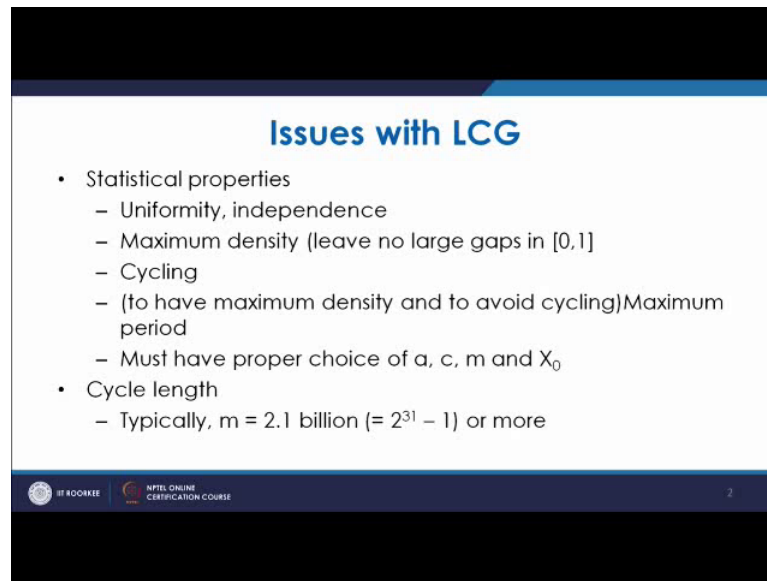
**Lecture – 17**  
**Issues and Challenges in Congruential Generators**

Welcome to the lecture on Issues and Challenges in Congruential Generators. In the last lecture we discussed about the linear congruential generator and we have seen that as the seed value was selected. So, we saw that after certain time that number is repeating. So, that basically gives a limitation to the number of random numbers which are generated before cycling. So, that is basically the period and what we have seen is that since it was 62, it; we saw that we got maximum of the 63 numbers, but many a times depending upon the different values of the parameters like a c or m you may get even less number of these generative random numbers.

So, these are the issues in case of LCGS Linear Congruential Generators or Congruential Generators that what should be the proper value of a c or m. So, we will discuss about it and then we also discuss about other requirements of the congruential generators or the number generated. So, what we saw that the properties which are required is the uniformity and independence. So, your uniformity means the probability should be same if you draw the plot between  $z_i$  and  $z_{i+1}$ , then it should be completely in the domain the point should be scattered in whole of the domain. So, that basically is a test which tells that the number generated is uniform.

Similarly, the density so density means there is very less gap between 0 and one we will see that many a times when we generate the numbers you have a large gap in between the numbers.

(Refer Slide Time: 02:23)



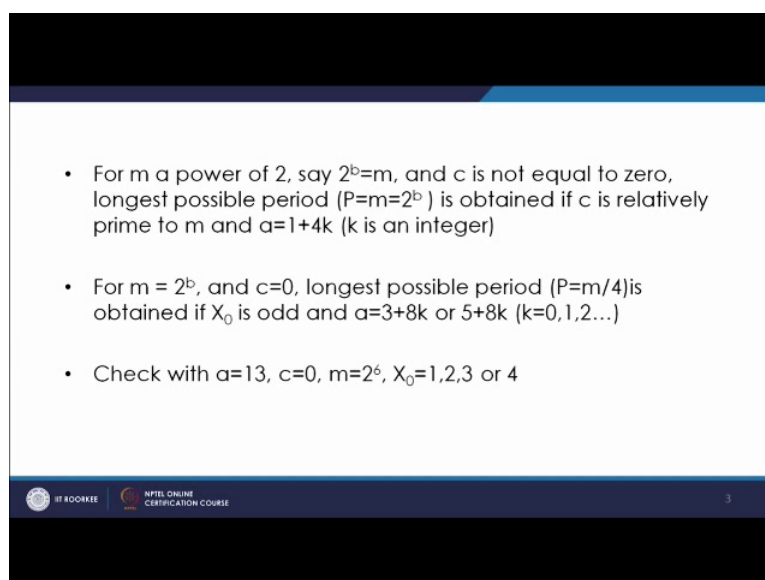
The slide is titled "Issues with LCG" in blue text. It contains a bulleted list of issues. The first bullet is "Statistical properties", which has three sub-bullets: "Uniformity, independence", "Maximum density (leave no large gaps in [0,1])", and "Cycling". The second bullet is "Cycle length", which has one sub-bullet: "Typically,  $m = 2^{31} - 1$  or more". At the bottom of the slide, there is a footer with two logos: "IIT ROORKEE" and "NPTEL ONLINE CERTIFICATION COURSE", and a page number "2" on the right.

- Statistical properties
  - Uniformity, independence
  - Maximum density (leave no large gaps in [0,1])
  - Cycling
  - (to have maximum density and to avoid cycling) Maximum period
  - Must have proper choice of  $a$ ,  $c$ ,  $m$  and  $X_0$
- Cycle length
  - Typically,  $m = 2^{31} - 1$  or more

So, basically that is the told to have less density you must have large density of the numbers. It should be independent as it is desired anywhere that numbers generated should be independent, cycling is there one issue because after certain time it will further start generating the similar numbers. So, this cycling is there and for having maximum density and to avoid this cycling which is cycling basically. So, to avoid this cycling you can say that you must have maximum period.

So, if the period is large it is going to give you the maximum density and it can also avoid the cycling. So, for that you have the requirement of proper choice of  $a$   $c$   $m$  and  $x$  naught that is seed. So, once you have the proper choice of these parameters then you can ensure that you have large number of random numbers generated before it is further recycled or further repeated. So, normally you are taking  $m$  as 2 raised to the power 31 or even more it is the requirement in today's world. So, you for such a large generation of random numbers you require proper you know selection of these parameters.

(Refer Slide Time: 04:09)



- For  $m$  a power of 2, say  $2^b=m$ , and  $c$  is not equal to zero, longest possible period ( $P=m=2^b$ ) is obtained if  $c$  is relatively prime to  $m$  and  $a=1+4k$  ( $k$  is an integer)
- For  $m = 2^b$ , and  $c=0$ , longest possible period ( $P=m/4$ ) is obtained if  $X_0$  is odd and  $a=3+8k$  or  $5+8k$  ( $k=0,1,2,\dots$ )
- Check with  $a=13$ ,  $c=0$ ,  $m=2^6$ ,  $X_0=1,2,3$  or  $4$

So, for  $m$  normally  $m$  is taken as a power of 2. So, for  $m$  to be a power of 2 say suppose  $2$  raise to the power  $b$ , that is taken as  $m$  and when  $c$  is not equal to  $0$  in that case. So, it is like a multiplicative type of generator, in that case the longest possible period you can get as  $2$  raise to the power  $b$ , when you get that  $c$  is relatively prime to  $m$  and  $a$  equal to one plus  $4k$ . So, this is one of the you know condition which is given that when you are taking  $m$  as  $2$  raise to the power  $b$  normally because we work on binary operation in computers. So, most of the numbers are represented in terms of you know  $1$  or  $2$ .

So, now if you take  $m$  as a power of  $2$  in that case suppose  $2$  raise to the power  $b$  and  $c$  is not equal to  $0$ . So, the maximum period as we have seen earlier. So, the maximum period would be which can we can achieve is  $2$  raise to the power  $b$ , but for that the condition is that  $c$  should be relatively prime to  $m$  and  $a$  will be equal to one plus  $4k$ . So, it means  $k$  is an integer. So, either it will be  $1$  or  $5$  or  $9$  or  $13$  like that. So, this way if that is the condition satisfied in that case you will get maximum possible period.

Again when  $m$  will be again  $2$  raise to the power  $b$  and  $c$  is equal to  $0$ , in that case the maximum possible period will be  $P$  equal to  $m$  by  $4$ . So,  $2$  raise to the power  $b$  minus  $2$ . So, if suppose you are taking  $2$  raise to the power  $6$  in 1 case  $64$  and if you take proper value of  $a$  and  $c$  is not equal to  $0$  and if you take proper value of  $a$  like  $1$   $5$  or  $9$  or  $13$ . So, in some case you can get the period of  $64$  whereas, if the  $c$  is not equal to  $0$  in that case period will be  $64$  by  $4$ . So, it will be  $16$ . So, period will come down to  $16$  and that  $16$

(Refer Time: 06:31) is also obtained when the seed value  $x_0$  is taken as the odd 1 and  $a$  is taken as 3 plus 8  $k$  or 5 plus 8  $k$ .

So,  $s$  would be either 3 plus 8  $k$  or 5 plus 8  $k$ . So,  $k$  will be an integer. So, when you will see that when you have this value of  $a$  taken in that case you get the period of 16. You can check with  $a$  equal to 13  $c$  equal to 0  $m$  equal to 2 raise to the power 6 and  $x_0$  1, 2, 3, or 4 let us check how it will look like.

(Refer Slide Time: 07:18)

$i$	$x_i$	$x_{i+1}$	$x_i$	$x_{i+1}$	$x_i$	$x_{i+1}$
0	1	13	2	26	4	52
1	13	41	26	18	52	36
2	41	21	18	42	36	20
3	21	17	42	34	20	4
4	17	29	34	58	4	52
5	29	57	58	50	52	36
6	57	37				

So, in this case you have  $a$  is 13 and  $c$  is 0,  $m$  is 2 raise to the power 6 and  $x_0$  you can take either 1 2 3 or 4. So, your formula is  $a X_i$  minus 1 and  $c$  is 0 and then it will be mod  $m$ . So, if you are  $a$  is 13 and you take 1. So, as you go at 0 your  $X_i$  is. So, if you are taking it is as 1 in that case your  $X_i$  plus 1  $X_i$  plus 1 will be.

So,  $X_i$ . So, this is  $X_y$  if it is  $X_i$  plus 1 here it will be  $X_i$  in that case  $X_i$  plus 1. So, in that case this number will be 1 into 13 and mod 64 so it is 13. Then it is 1. So,  $X_i$  is 13  $X_i$  plus 1 will be 13 into so  $a$  into  $X_i$  minus 1,  $a$  is 13 and 13 into 13 that is 169 so 169 mod 64. So, it will be 169 mod 64 so 64 into to 128 and that is 41. So, further you go here it is coming as 41 here 41 into 13 is 533. So, again 64 533 mod 64 so 64 into 8 is 512. So, it will be 21.

Similarly, you go to 3 21, 21 into 13, 21 into 13 is 273. So, again you will have 17. So, 17 then again 17 multiplied by 13 this way you will get 221. So, 64 into 31 92, it will be

29 and so on. So, this way it will go on further producing we can go for some more steps if you go further. So, you have 17 and this is 29 here. So, this 29 will come here again this 29 multiplied again 13. So, that is 377 and 377 will be 16 into 64 into 53 20. So, it will be 57 and 57 will come again in the 16th it will be 67; 67 into 13 that is further 741.

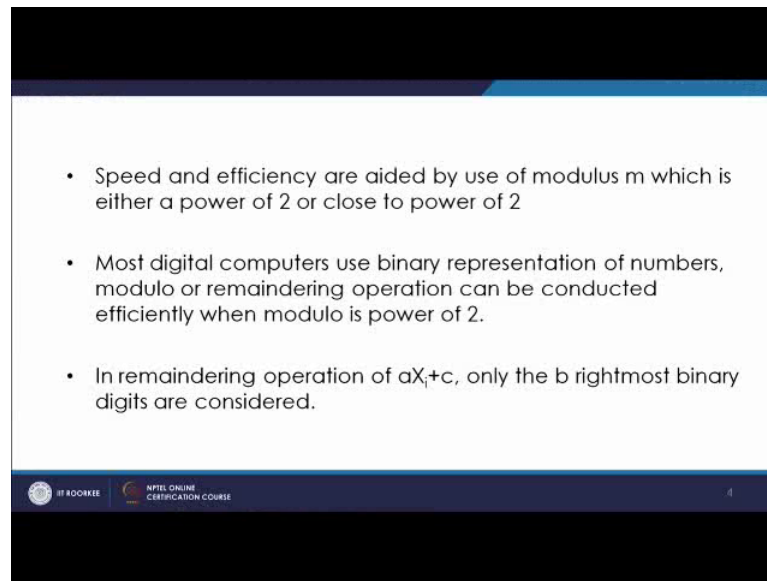
So, 741 again it will be 64 into 11 is 7 0 4. So, it will be 37 like that. So, it will go on increasing. Now it will go and it will have certain period and here as  $c$  is 0 you will see if you go proceed further you will see that the period is number coming number 2 16. Now let us check it for  $X_i$  as so we have taken this value as a odd number an odd number, now let us take with another number if we take  $X_i$  as initially as 2. So, what we get with 2 we get 26, 2 into 13 and then you are so a into  $X$  initial number so initially you are getting 2 say 26.

Now, if you take 26 here 26 into 13 338, now 338 will be 16 into 4 320. So, it will be 18 18 into 3 13 that is 234. So, it is 16 4 into 3 192 plus 42. So, it is 42 and 42 multiplied by 13 546. So, it will be 64 into 8 512. So, then it is remaining 34. Similarly 34 34 into 13 442 so it will be 64 into you know 6 384. So, it will be 58; 58 and multiplied by again if you take into 13. So, it will be 754.

So, 754 will be 64 into 11 704. So, it will be 5 like that it will move. Now let us see with 4 initial value as 4, if you take initial value as 4 then how long it goes 4 into 13 52 then 52 into 13. So, 52 into 13 will be 676. So, that will be 36 36 into 13 36 into 13 will be 468. So, it will be 464 into 7 448 plus 20. So, 20 and then 20 into 13 260 so 26 for that for 2 6 it will be 4. So, that we see that this 4 comes here what we see that this 4 is coming here from. So, this is coming here.

Now, after this 4 it will further go as 52 from 52 it will go to 36 like that. So, this is repeated what we see that in this case as we take the different value of the seed the cycle. So, here you see that only in 4 it is getting vanished. So, you have only 4 different random numbers generated whereas, in this case it will go maybe up to 8 or 16. So, in this case it will go if you check it will go to a larger value. So, this is what how the value of the seed can affect the cycle of the generated numbers.

(Refer Slide Time: 15:06)



- Speed and efficiency are aided by use of modulus  $m$  which is either a power of 2 or close to power of 2
- Most digital computers use binary representation of numbers, modulo or remaindering operation can be conducted efficiently when modulo is power of 2.
- In remaindering operation of  $aX_i + c$ , only the  $b$  rightmost binary digits are considered.

Now, speed and efficiency are aided by use of modulus  $m$  which is either a power of 2 or close to the power of 2.

So, basically we have the speed as well as the efficiency and what we see is that normally depending upon the proper selection of  $m$ , you have proper efficiency more and more number of numbers are generated and the period will be higher and higher. Now what we see that we have studied; so far that normally we take this modulo value as 2 raise to the power some integers. So, that is 1 preferred 1 while in normal practice we prefer to go for the decimal presentation of the numbers. So, in that we feel easier to take the remainder value. So, we as you have simply to take the last values. So, if suppose modulus value is something like 10 raise to the power 3 or so the last 3 digits are simply taken.

Suppose modulus value is 10 raise to the power 3 or 1000 and if you get the 4 digit number. So, only the last 3 digits are taken as the remainder ones. So, the same thing applies in the case of binary representation also, in case of binary representation also you have some number and once you have the modulus. So, modulus is 2 raise to the power  $b$  so the last  $b$  significant digits towards the right that is coming as the remainder. So, that can be checked and that we can see by an example. So, what we see that most digital computers use binary representation of numbers and remaindering operation can be conducted efficiently when the modulo is in power of 2. So, rightmost binary digits are

taken just like an example you can see if you have m raise to the power 2 a is 9 19 c is 0 and X 0 that is seed is 63.

(Refer Slide Time: 17:11)

$m=10^2, a=19, c=0, X_0=63$   
 $X_1 = (1197) \bmod 10^2 = 97$   
 $X_2 = \left( \frac{97 \times 19}{1843} \right) \bmod 10^2 = 43$   
 $X_3 = \left( \frac{43 \times 19}{817} \right) \bmod 100 = 17$   
Binary representation  
 $(173) \bmod 2^7 = 45 \quad \left[ 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 \right]$   
 $173 = 1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$   
 $= 10101101$   
 (last 6 digits or last 6 digits)

So, now in this case what you see is X 1 will be. So, this is a multiplicative type of generator you have 19 into 63. So, it is 1197 and it will be mod m. So, if mod is 10 raise to the power 2. So, mod you can write 10 raise to the power 2. So, what you do is only the; this 10 to the power b that is b 2 is b. So, only the last 2 digits are taken so it is 97. Similarly you go to X 2; X 2 will be 97 multiplied by 19 that is all and mod again 10 square. So, 97 multiplied by 19 it will be 1843. So, it will be 1843. So, in that case 43 will come because it is equal to 1843. So, only these 2 digits i is coming.

So, that will go on; similarly further X 3 will be 43 into 19. So, that is 817. So, 4 3 into 19 817 that is 817 mod 100, it will be last 2 digits 17 like that it will be coming. So, this way you see in the binary you know in the decimal representation we find it easier similar thing happens in the case of binary representation. So, in the case of binary representation suppose you have the number 173 and mod 2 raise to the power 6. So, if that is 64 and 173 comes now if you look at 173; 173 will be represented in what terms in the binary.

So, it will be 2. So, this 173 can be written as one into 2 raise to the power. So, we have to come to 128. So, 2 raised to the power 7 plus then 45 is remaining. So, 45 will be

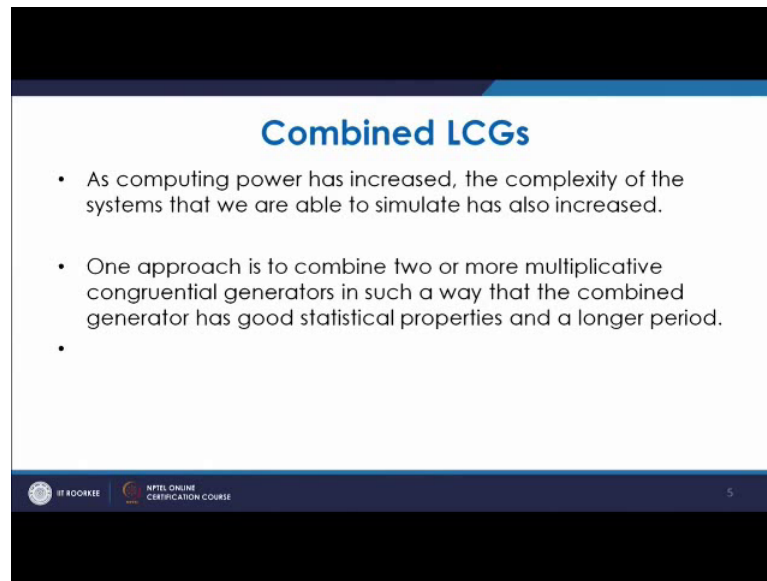
again  $2^5$  is 32. So, 6 will be  $0 \times 2^6$  is 0 here, then you have 1 into  $2^5$ . So, that is your 32, your 13 is remaining; 13 means  $2^4$  is 16. So, again that will not be there. So, 0 into  $2^4$  plus then you have 1 into  $2^3$ . So, that is 8. So, you have 5 remaining 1 into  $2^2$ , 4 remaining 1 remaining 1 into  $2^1$  plus 1 into  $2^0$ .

So, we write it as 1 0 1 0 1 1 0 1, now if we take the mod of this number with  $2^6$  means 64. So, that comes out to be 45, now what will be the 45 45 is nothing, but we are getting 45 from here onwards. So, if you try to find 45 for 45 it will be 1 into  $2^5$  32 plus 13. So, 0 into  $2^4$ , plus 1 into  $2^3$ , plus 1 into  $2^2$ , plus 0  $2^1$ , plus 1  $2^0$ . So, it is coming as 1 0 1 1 0 1. Now look at this is nothing, but this is the last b digits you have b equal to 6 this was this is  $2^b$ . So, b is 66 b in 6 your b is 6 here. So, what we see is this last 6 digits last b digits that is last 6 digits.

So, what we see is that the remaindering operation becomes quite easy when we take this system of binary digits and in that case you simply take the last b digits. So, either b is the decimal presentation or b is the value representation you can get this and since most of the computers are working on this binary representation. So, this becomes easier for the computer. Then most digital computers use so as we discuss that most of the digital computers are using this binary representation of numbers. So, modulo or remaindering operation can be conducted efficiently and you have to simply use the b rightmost digits. So, that you get the remainder numbers quite early quite efficiently.



(Refer Slide Time: 23:52)



The slide is titled "Combined LCGs" in blue text. It contains three bullet points. The first bullet point states that as computing power has increased, the complexity of the systems that we are able to simulate has also increased. The second bullet point states that one approach is to combine two or more multiplicative congruential generators in such a way that the combined generator has good statistical properties and a longer period. The third bullet point is a single dot. At the bottom of the slide, there is a footer with the NPTEL logo, the text "NPTEL ONLINE CERTIFICATION COURSE", and the number "5".

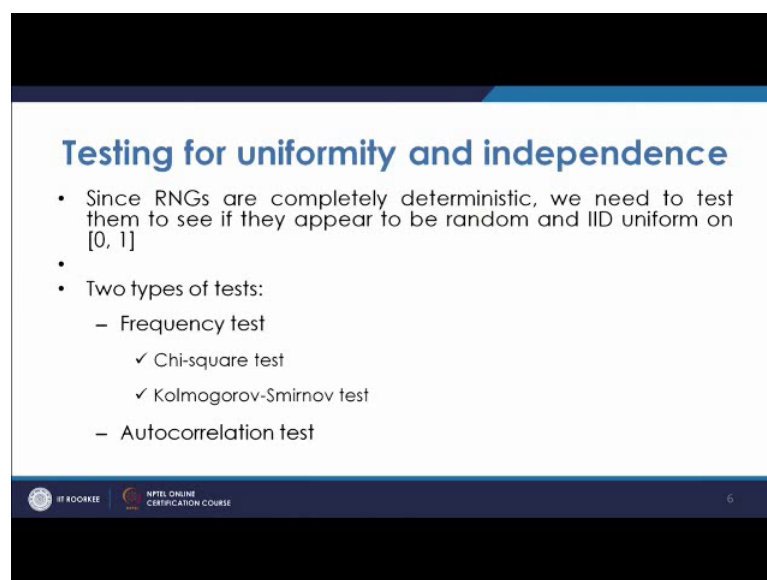
- As computing power has increased, the complexity of the systems that we are able to simulate has also increased.
- One approach is to combine two or more multiplicative congruential generators in such a way that the combined generator has good statistical properties and a longer period.
- 

As computing power has increased the complexity of the system that we are able to simulate has also increased now the thing is that many a times you require large number of these generators I mean number generated numbers. So, even these linear congruential generators are not sufficient to generate that many numbers where you can generate enough number of random numbers. So, another way is to combine the you know linear congruential generators. So, you have 2 streams and you are combining taking the  $a$  and  $m$  of both these streams and they are joining together. So, that is known as combined linear congruential generators.

So, by combining the 2 or more core multiplicative congruential generators, in such a way that the combined generator has good statistical properties and a longer period so you have composite generators you have combined congruential generators linear congruential generators, also by which you can generate the random numbers having larger you know periods.

So, that is used in the simulations nowadays where large and large number of similar numbers are and you can read them from the standard book of modeling and simulation. So, that you get used to those methods.

(Refer Slide Time: 25:27)



**Testing for uniformity and independence**

- Since RNGs are completely deterministic, we need to test them to see if they appear to be random and IID uniform on  $[0, 1]$
- 
- Two types of tests:
  - Frequency test
    - ✓ Chi-square test
    - ✓ Kolmogorov-Smirnov test
  - Autocorrelation test

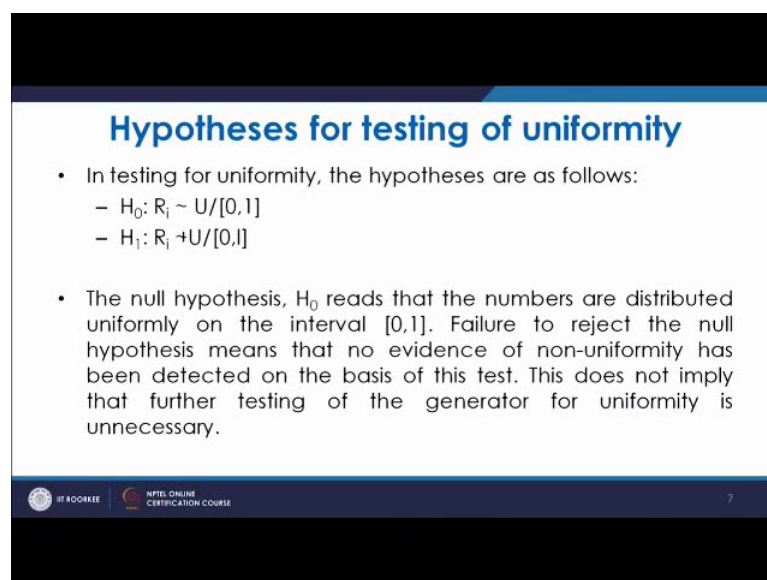
IF ROOKIE NPTEL ONLINE CERTIFICATION COURSE 6

Now, testing for uniformity and independence so once you have generated the random numbers then you need to test them. So, these random number generators are deterministic, we know that which of the random number is going to come because it is based on certain mathematical formulation. So, we need to see that these numbers which are generated they are really uniform and they are independent. So, for that we have certain type of tests which we carry on it, now the test which we carry on it are the frequency test and the autocorrelation test, now in the frequency test we basically. So, that is test for the uniformity.

Now, in that case there are 2 kinds of test which are done one is chi square test, another is called Kolmogorov Smirnov test. So, what it does is basically looks at. So, the main purpose as we had discussed that once you have  $X_i$  and  $X_{i+1}$   $X_i$  and  $X_{i+1}$  so be plotted against each other and we should see that the plot which we are getting, the points which we are getting this will be easy if it is completely you know spread in the domain it means it is completely uniform otherwise they may not be uniform. So, there may be you know non uniform here may be you know either they may be tilted towards one side or other side so mean as we had discussed many points that may be faced.

So, these tests we will try to discuss how we are going to test in the uniformity and independence. Now let us see let us discuss about the Kolmogorov and Smirnov test. So, in that test basically what we do is.

(Refer Slide Time: 27:38)



**Hypotheses for testing of uniformity**

- In testing for uniformity, the hypotheses are as follows:
  - $H_0: R_i \sim U/[0,1]$
  - $H_1: R_i \not\sim U/[0,1]$
- The null hypothesis,  $H_0$  reads that the numbers are distributed uniformly on the interval  $[0,1]$ . Failure to reject the null hypothesis means that no evidence of non-uniformity has been detected on the basis of this test. This does not imply that further testing of the generator for uniformity is unnecessary.

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 7

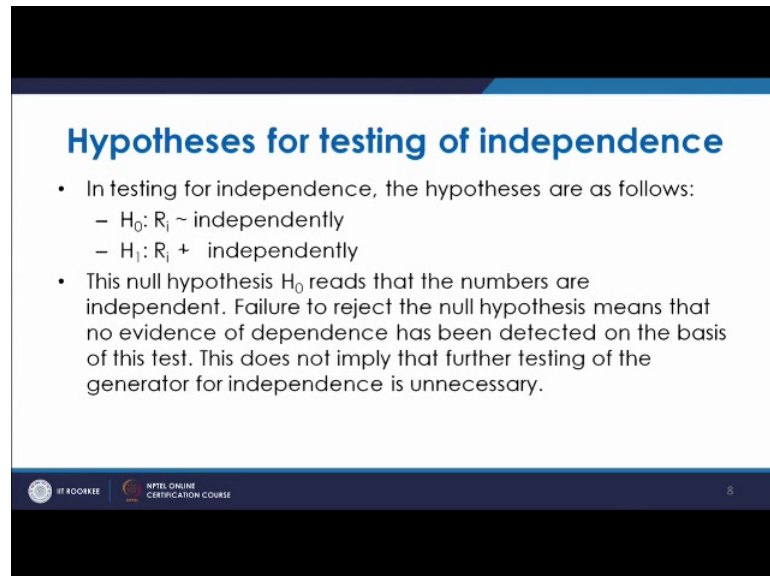
Now, first of all there is certain hypothesis for the testing of the uniformity, what is that hypothesis the hypothesis is telling like this is the null hypothesis  $H_0$  and it reads that if the numbers are distributed uniformly on the interval  $[0,1]$ , then failure to reject the null hypothesis means that there is no evidence of non uniformity which has been detected there is certain hypothesis you have certain testing methods and you have some standard data.

So, once the condition which is basically set if it meets that condition then we can say. So, that null hypothesis if it is. So, if it is said that it is satisfying that condition, in that case the; we tell that the you do not have much of the evidence to reject that it is not uniformly distributed. So, failure to reject the null hypothesis means that there is no proper evidence of non uniformity, which has been detected on the basis of the test and it does not also imply that further testing of the generator for uniformity is unnecessary. So, you can further go testing the uniformity and you can see and you cannot say that it is unnecessary. So, you can see the necessity of further testing.

So, that is known as null hypothesis for testing of uniformity similarly you have hypothesis for testing of independence again here also you have the null hypothesis,  $H_0$  so if it may be independent it may not be. So, if it is satisfying the condition it you can say that it is independently distributed. Independently generated and if not so the null hypothesis  $H_0$  reads that the numbers are independent, if they are meeting with that

condition you are generating certain numbers based on the ideal conditions and if the that condition is satisfied, we are telling that the numbers are independent. Failure to reject the null hypothesis means that no evidence of independence has been detected on the basis of the test this does not imply that further testing of the generator for independence is unnecessary.

(Refer Slide Time: 30:03)



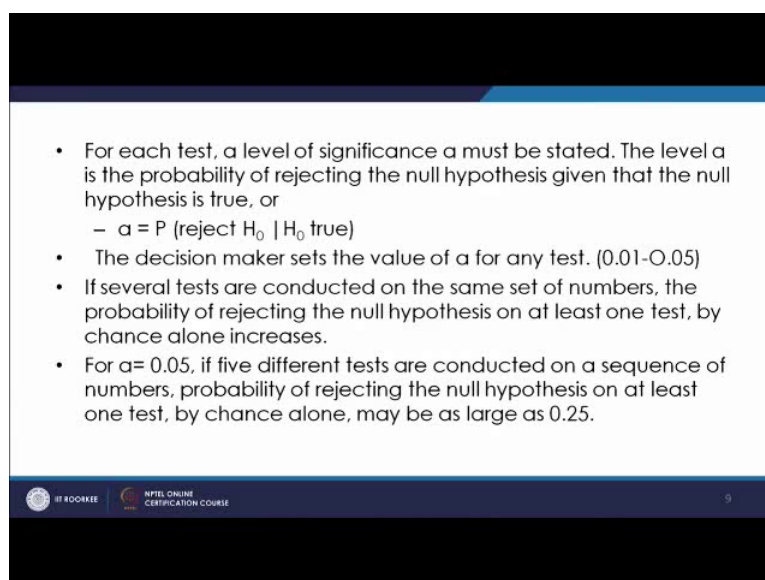
**Hypotheses for testing of independence**

- In testing for independence, the hypotheses are as follows:
  - $H_0: R_i \sim$  independently
  - $H_1: R_i \nmid$  independently
- This null hypothesis  $H_0$  reads that the numbers are independent. Failure to reject the null hypothesis means that no evidence of dependence has been detected on the basis of this test. This does not imply that further testing of the generator for independence is unnecessary.

IT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 8

So, in these cases what happens that you have set of numbers which are to be tested, these sets are basically going through certain mathematical computation and then we are finding certain statistical parameters and then we are seeing whether this parameter can work I mean whatever you get whether this will be or not.

(Refer Slide Time: 30:43)



- For each test, a level of significance  $\alpha$  must be stated. The level  $\alpha$  is the probability of rejecting the null hypothesis given that the null hypothesis is true, or
  - $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true})$
- The decision maker sets the value of  $\alpha$  for any test. (0.01-0.05)
- If several tests are conducted on the same set of numbers, the probability of rejecting the null hypothesis on at least one test, by chance alone increases.
- For  $\alpha = 0.05$ , if five different tests are conducted on a sequence of numbers, probability of rejecting the null hypothesis on at least one test, by chance alone, may be as large as 0.25.

IT ROOKIE NPTEL ONLINE CERTIFICATION COURSE 9

So, for each test a level of significance  $\alpha$  is stated now when we are thinking of having any test you are signifying a level  $\alpha$  like it is  $\alpha$  is 0.05 or 0.1 or 0.06 or. So, that the level  $\alpha$  is the probability of rejecting the null hypothesis given that null hypothesis is true.

So, that basically tells you the probability of rejecting the hypothesis. So, its value may be something like 0.01 to 0.05. So, it tells you the probability of rejecting the null hypothesis, the another significance of this  $\alpha$  is that when we are doing more number of tests. So, if you are doing one number of tests it is if you are to doing this several number of tests on the same set of numbers, then probability of rejecting the null hypothesis on at least one test by chance alone is increasing. So, basically if you are setting this  $\alpha$  as 0.05 it means for 1 test there is 5 percent chance that for 1 test that it chance that it may be rejected.

Similarly, if  $\alpha$  is 0.05 and if there are 5 different tests conducted on a sequence of numbers, in that case probability of rejecting the null hypothesis on at least one test will be increasing. So, it will be increase by 5 times. So, it means if it is once then it will be 5 percent and if it is done suppose  $n$  number of times, then this one of the one of the time it basically rejects the null hypothesis test that basically increases to 5 times whatever the  $\alpha$  value is it. So, it will increase to 5 times  $\alpha$  that is 0.05 that is 0.25 5 times 0.05 so 0.25. So, that probability is increasing. So, what happens in normal cases you are

given certain set of numbers and I mean you generate these numbers and once you generate these numbers then you are basically converting it into you know variates between 0 and 1 and by dividing it with  $m$  and once you divide get that value then that value is subject to this testing.

Now, for testing you need to specify the value of  $\alpha$ . So, once you have the  $\alpha$  now for that you have certain you know methodology how to go further. So, we will see in our next lecture that if you have been given the set of numbers. So, first of all the numbers in the Kolmogorov Smirnov test, what we do is the numbers which are given first of all they are arranged in certain orders and then some parameters are calculated and these parameters on the basis of these parameters you have another parameter which will be coming from the ideal you know table and if this value is lesser than this ideal value from the table we say that we cannot reject it and if it is more than we reject it.

Similar is the case in the case of chi square test. So, in that also you find a parameter based on the interval numbers and then the you know uniform ideal realistic uniform distribution you know samples and based on that you get another number is chi square is specified and again that value being less than the uniform distributed value in that case we say that we cannot reject it. So, that passes the condition of uniformity. So, this way the condition of this condition of uniformity are you know tested. So, may be in the next lecture once we get time. So, we have in the next lecture we will discuss about this uniformity and independence test for such type of you know general random number generators and we can check it.

Thank you very much.