Chapter 7
Random Number Generation

Banks, Carson, Nelson & Nicol

Discrete-Event System Simulation

Properties of Random Numbers

- Two important statistical properties:
  - Uniformity
  - Independence.
- Random Number, *Ri*, must be independently drawn from a uniform distribution with pdf, given by f(x)
- The expected value of each Ri is given by E(R)
- The variance is given by V(R) = $\int_0^1 x^2 dx - [E(R)]^2$ = 1/12

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$E(R) = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

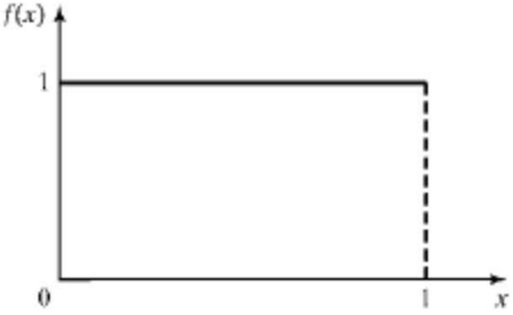


Figure: pdf for random numbers

Generation of Random Numbers

- Important considerations in RN routines:
  - Fast
  - Portable to different computers and different programming languages
  - Have sufficiently long cycle, before the numbers start repeating.
  - Replicable. If a starting point of RNs or a condition is given, we should be able to generate the same set of RNs (for debugging)
  - Closely approximate the ideal statistical properties of uniformity and independence.

Generation of pseudo-random numbers

- “Pseudo” means false. Generating random numbers using a known method removes the potential for true randomness.
- If the method is known the random numbers can be replicated.
- Goal: To produce a sequence of numbers in [*0,1*] that simulates, or imitates, the ideal properties of random numbers (RN).

Characteristics of a Good Generator

- Maximum Density - Such that the values assumed by *Ri, i = 1,2,…,* leave no large gaps on *[0,1].* They have to be uniformly distributed.
- The random numbers may be discrete instead of continuous. Each *Ri* is discrete.
  - □ Solution: a very large integer for modulus m
- The mean of generated numbers might be too high or too low.
- The variance of generated numbers might be too high or too low.
- Maximum density should be achieved and cycling should be avoided.

Characteristics of a Good Generator (contd.)

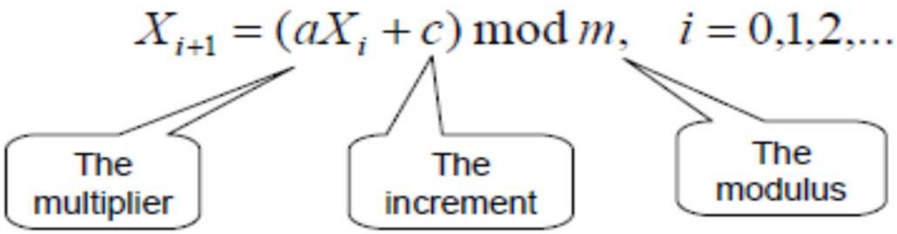
- Speed and efficiency are aided by a modulus, *m*, to be (or close to) a power of *2*.
- For the above point, Maximum Period is very important.
  - Achieved by: proper choice of *a*, *c*, *m*, and *X0.*
  - *If m is very large,* $m = 2^{31} - 1$ *or* $m = 2^{48}$
- Dependence can exist :
  - Autocorrelation b/w numbers
  - Numbers successively higher or lower than adjacent numbers.
  - Several numbers above the mean followed by several numbers below the mean

Techniques for Generating Random Numbers

- Linear Congruential Method (LCM). (most widely used)
- Combined Linear Congruential Generators (CLCG).
- Random-Number Streams.

Linear Congruential Method

- To produce a sequence of integers, *X1, X2, …* between *0* and *m-1* by following a recursive relationship:



$$X_{i+1} = (aX_i + c) \bmod m, \quad i = 0,1,2,\ldots$$

The multiplier

The increment

The modulus

- The selection of the values for *a*, *c*, *m*, and X_0 *(seed)* drastically affects the statistical properties and the cycle length.
- If c = 0, the form is multiplicative congruential method
- If c != 0, the form is mixed congruential method
- The random integers are being generated [*0,m-1*], and to convert the integers to random numbers:

Ri = Xi / m (I = ,1,2 ….)

Example

- Use $X_0 = 27$, $a = 17$, $c = 43$, and $m = 100$.
- The Xi and Ri values are:

$X1 = (17*27+43) \bmod 100$

$= 502 \bmod 100$

$= 2,$

$R1 = 0.02;$

$X2 = (17*2+43) \bmod 100$

$= 77$

$R2 = 0.77;$

$X3 = (17*77+43) \bmod 100$

$= 52$

$R3 = 0.52;$

Example

- Use $X_0 = 7$, $a = 5$, $c = 3$, and $m = 16$.
- The *Xi* and *Ri* values are:

$X1 = (5*7+3) \bmod 16$

$= 38 \bmod 16$

$= 6,$

$R1 = 0.375;$

$X2 = (5*6+3) \bmod 16$

$= 1$

$R2 = 0.063;$

$X3 = (5*1+3) \bmod 16$

$= 8$

$R3 = 0.5;$

Example

- Use $a = 13$, $c = 0$, and $m = 64$
- The period of the generator is very low
- Seed X_0 influences the sequence

For X0=1, 3 the period is 16
For X0= 2, period is 8
For X0= 4, period is 4

	X_i	X_i	X_i	X_i
i	X_0=1	X_0=2	X_0=3	X_0=4
0	1	2	3	4
1	13	26	39	52
2	41	18	59	36
3	21	42	63	20
4	17	34	51	4
5	29	58	23	
6	57	50	43	
7	37	10	47	
8	33	2	35	
9	45		7	
10	9		27	
11	53		31	
12	49		19	
13	61		55	
14	25		11	
15	5		15	
16	1		3	

Linear Congruential Method [Techniques]

- Maximum Density
  - The values assumed by R_i, $i=1,2,\dots$ leave no large gaps on [0,1]
  - Problem: Instead of continuous, each R_i is discrete
  - Solution: a very large integer for modulus m
    - Approximation appears to be of little consequence

- Maximum Period
  - To achieve maximum density and avoid cycling
  - Achieved by proper choice of a, c, m, and X_0

- Most digital computers use a binary representation of numbers
  - Speed and efficiency are aided by a modulus, m, to be (or close to) a power of 2.

Linear Congruential Method

- For m a power 2, $m=2^b$, and $c\neq0$
  - Longest possible period $P=m=2^b$ is achieved if c is relative prime to m and $a=1+4k$, where k is an integer
- For m a power 2, $m=2^b$, and $c=0$
  - Longest possible period $P=m/4=2^{b-2}$ is achieved if the seed X_0 is odd and $a=3+8k$ or $a=5+8k$, for $k=0,1,\ldots$
- For m a prime and $c=0$
  - Longest possible period $P=m\text{-}1$ is achieved if the multiplier a has property that smallest integer k such that $a^k\text{-}1$ is divisible by m is $k = \mathrm{m\text{-}1}$

Combined Linear Congruential Generators

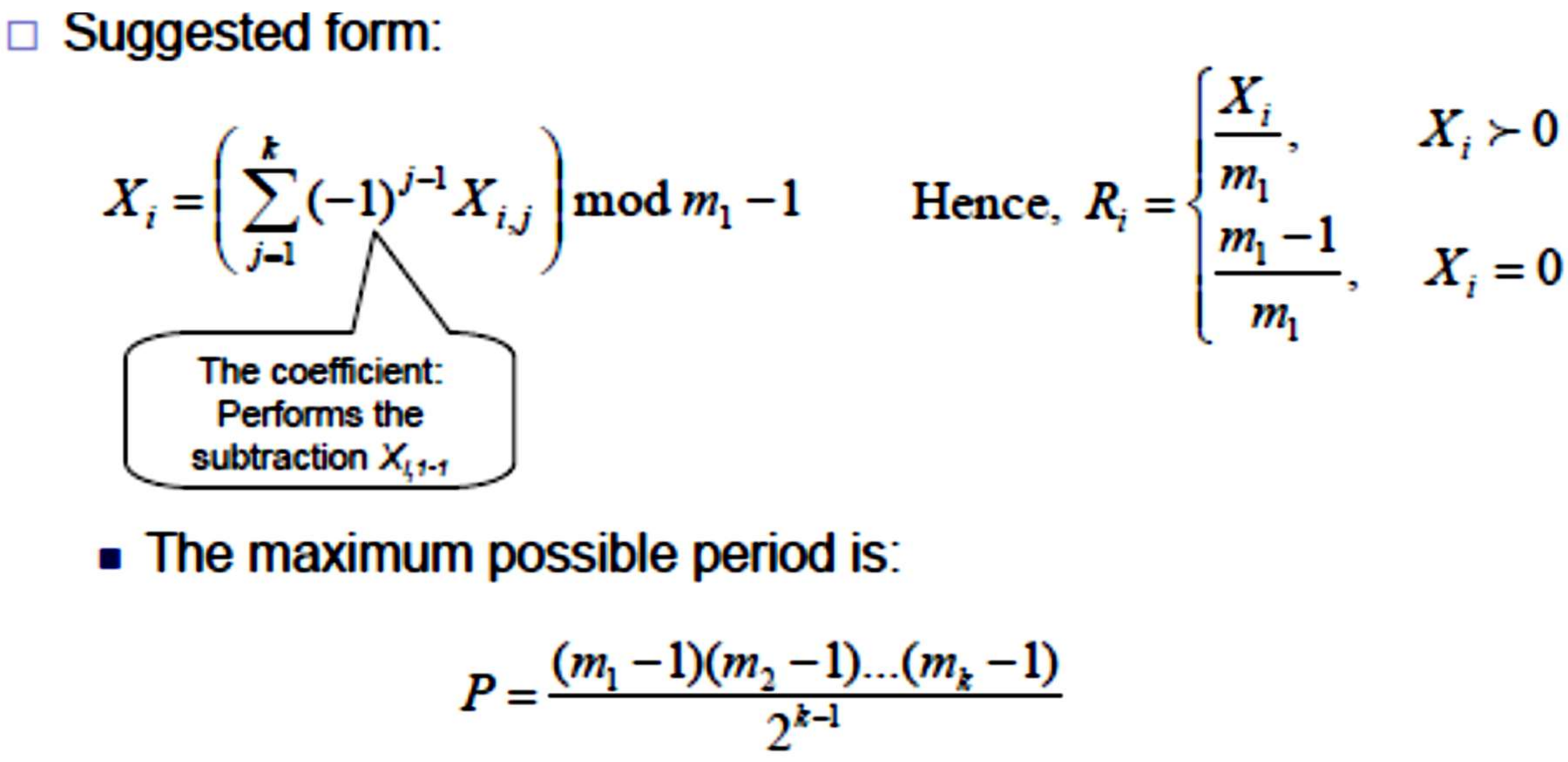
- Reason: Longer period generator is needed because of the increasing complexity of stimulated systems.
- Approach: Combine two or more multiplicative congruential generators.
- L'Ecuyer suggests that :
  - □ If $W_{i,1}$, $W_{i,2}$, $W_{i,3}$ …. $W_{i,k}$ are any independent, discrete valued variables, $W_{i,k}$ is uniformly distributed on the integer from 0 to m_1-2, then,

$$W_i = \left[\sum_{j-1}^{k} W_{i,j} \right] \bmod m_1 - 1$$

Combined Linear Congruential Generators

- Let $X_{i,1}$, $X_{i,2}$, …, $X_{i,k}$, be the *i*th output from *k* different multiplicative congruential generators
- The jth generator:
  - □ Has prime modulus m_j & multiplier a_j *is chosen such that* period is m_j*-1*
  - □ Produces integers $X_{i,j}$ is approx ~ Uniform on integers in [*1,*m_j*-1*]
  - □ $W_{i,j} = X_{i,j}$ *-1* is approx ~ Uniform on integers in [*1,* m_j*-2*]

Combined Linear Congruential Generators



- Suggested form:

$$X_i = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) \bmod m_1 - 1 \qquad \text{Hence, } R_i = \begin{cases} \dfrac{X_i}{m_1}, & X_i \succ 0 \\ \dfrac{m_1 - 1}{m_1}, & X_i = 0 \end{cases}$$

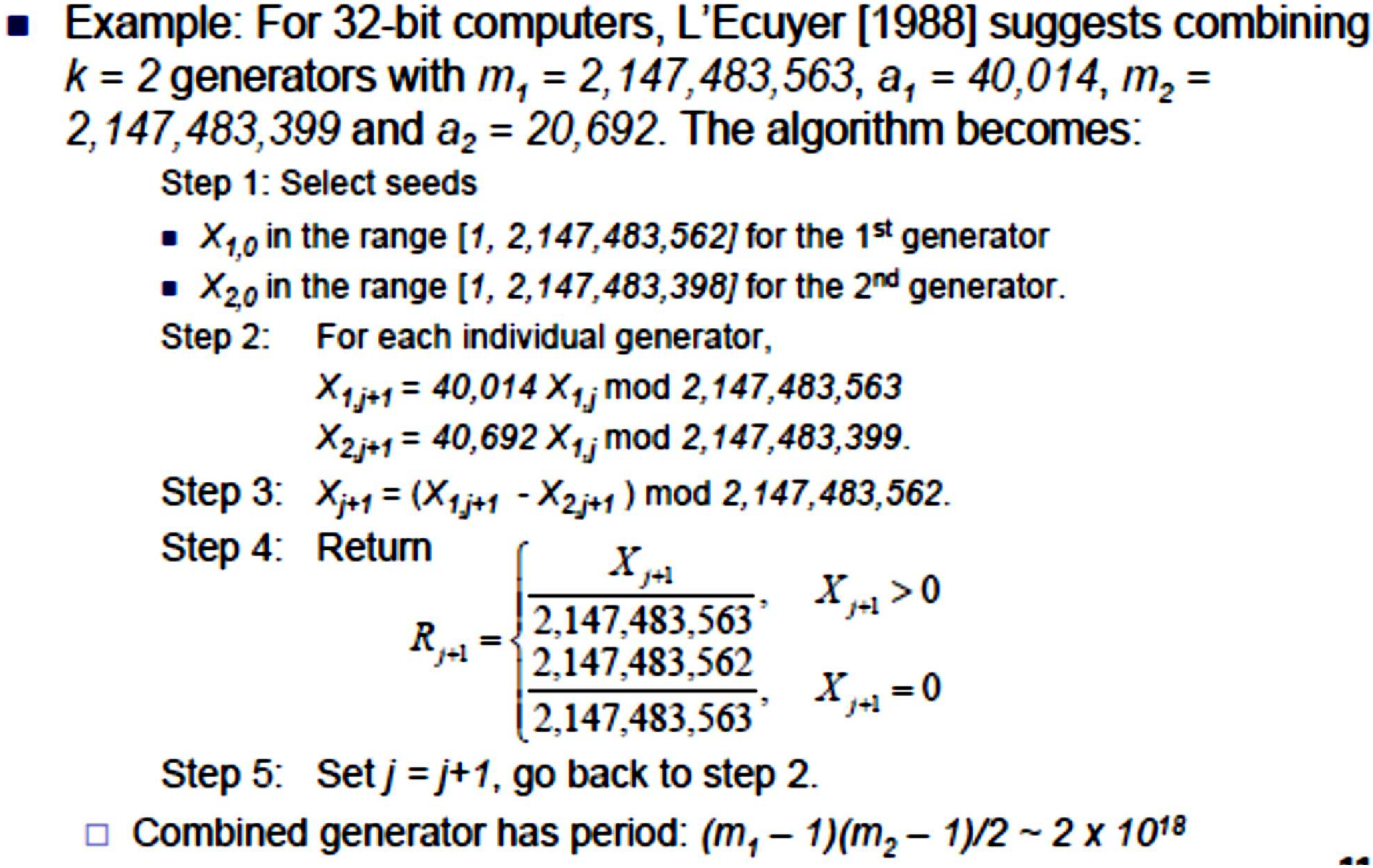
The coefficient: Performs the subtraction $X_{i,1-1}$

  - The maximum possible period is:

$$P = \frac{(m_1 - 1)(m_2 - 1)\ldots(m_k - 1)}{2^{k-1}}$$

The combined generators have periods as long as 2^{191} approx = 3×10^{57}

Combined Linear Congruential Generators



- Example: For 32-bit computers, L'Ecuyer [1988] suggests combining $k = 2$ generators with $m_1 = 2,147,483,563$, $a_1 = 40,014$, $m_2 = 2,147,483,399$ and $a_2 = 20,692$. The algorithm becomes:

  Step 1: Select seeds

  - $X_{1,0}$ in the range $[1,\ 2,147,483,562]$ for the 1st generator
  - $X_{2,0}$ in the range $[1,\ 2,147,483,398]$ for the 2nd generator.

  Step 2: For each individual generator,

  $X_{1,j+1} = 40,014\ X_{1,j}$ mod $2,147,483,563$

  $X_{2,j+1} = 40,692\ X_{1,j}$ mod $2,147,483,399$.

  Step 3: $X_{j+1} = (X_{1,j+1} - X_{2,j+1})$ mod $2,147,483,562$.

  Step 4: Return

  $$R_{j+1} = \begin{cases} \dfrac{X_{j+1}}{2{,}147{,}483{,}563}, & X_{j+1} > 0 \\ \dfrac{2{,}147{,}483{,}562}{2{,}147{,}483{,}563}, & X_{j+1} = 0 \end{cases}$$

  Step 5: Set $j = j+1$, go back to step 2.

  □ Combined generator has period: $(m_1 - 1)(m_2 - 1)/2 \sim 2 \times 10^{18}$

Random Number Streams

- The seed for a linear congruential random-number generator:
  - Is the integer value X_0 that initializes the random-number sequence.
  - Any value in the sequence can be used to “seed” the generator.
- A random-number stream:
  - Refers to a starting seed taken from the sequence $X_0, X_1, \ldots, X_{P.}$
  - If the streams are b values apart, then stream i could defined by starting seed: $S_i = X_{b(i-1)}$
  - Older generators: $b = 10^5$; Newer generators: $b = 10^{37}$.
- A single random-number generator with k streams can act like k distinct virtual random-number generators
- To compare two or more alternative systems.
  - Advantageous to dedicate portions of the pseudo-random number sequence to the same purpose in each of the simulated systems.

Tests for Random Numbers

- The desirable properties of random numbers is uniformity and independence.
- Tests are performed to check whether the two properties are satisfied by the sequence of random numbers generated.
- Two types of tests are performed
  - Frequency test
  - Autocorrelation test

Tests for Random Numbers (contd.)

- Testing for **uniformity:**

  H_0: *Ri ~ U[0,1]*

  H_1: *Ri !~ 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, H_0, means that evidence of non-uniformity has not been detected by this test.
- Testing for **independence:**

  H_0: *Ri* ~ independently

  H_1: *Ri !~* independently

- This null hypothesis H_0, reads that the numbers are independent.
- Failure to reject the null hypothesis, H_0, means that evidence of dependence has not been detected by this test.

Tests for Random Numbers (contd.)

- Testing for the above to properties and proving that the hypothesis is not rejected does not mean that further tests for the generator are not required.
- For each test, a level of significance is α stated.
- Level of significance α, the probability of rejecting H0 when it is true:

$$\alpha = P(\textit{reject } H0 | H0 \textit{ is true})$$

- *The decision maker sets the value of* α *for any test*
- *Frequently* α *is set to 0.01 or 0.05*

Tests for Random Numbers (contd.)

- If $\alpha = 0.05$ and 5 different tests are conducted on a set of numbers, then, the probability of rejecting the null hypothesis on atleast one test, by chance alone, could be as large as 0.25
- If 100 numbers were subject to a test, with $\alpha = 0.05$, it would be expected that 5 of those tests would be rejected.
- If the number of rejections in 100 tests is close to 100α then, the generator should not be rejected.

- When to use these tests:
  - If a well-known simulation language or random-number generator is used, it is probably unnecessary to test
- When not to use these tests:
  - If the generator is not explicitly known or documented
  - When spreadsheet programs, symbolic/numerical calculators are used,
  - When random number generators are added to software which are not developed for simulation,

- Types of tests:
  - Theoretical tests: evaluate the choices of m, a, and c without actually generating any numbers
  - Empirical tests: applied to actual sequences of numbers produced, which is our emphasis in this learning.

Frequency Tests

- Test of uniformity
- Two different methods:
  - Kolmogorov-Smirnov test
  - Chi-square test
- Both these tests measure the degree of agreement b/w the distribution of a sample of generated random nos. and the theoretical uniform distribution.
- Both tests are based on the null hypothesis of no significant difference b/w the sample distribution theoretical distribution.

Kolmogorov-Smirnov Test

- Compares the continuous cdf, *F(x)*, of the uniform distribution with the empirical cdf, *SN(x),* of the *N* sample observations.
- We know that : F(x) = x,   0 <= x <= 1
- If the sample from the RN generator is

  $$R_1, R_2, \ldots, R_n,$$

  then the empirical cdf, $S_N(x)$ is:

  $$S_N(x) = \frac{\text{number of } R_1, R_2, \ldots, R_n \text{ which are} \le x}{N}$$

- As N becomes larger, S_N(x) should become a better approximation to F(x), provided that the null hypothesis is true.

Kolmogorov-Smirnov Test

- It is based on the largest absolute deviation between $S_N(x)$ and $F(x)$ over the range of the random variable.
- It is based on the statistics :

$$D = \max \left| F(x) - S_N(x) \right|$$

- Sampling distribution of *D* is known (a function of *N*, tabulated in Table A.8.)

Kolmogorov-Smirnov Test

- For performing Kolmogorov test against a uniform cdf, follow the steps :

1. Rank the data from the smallest to largest
2. Compute

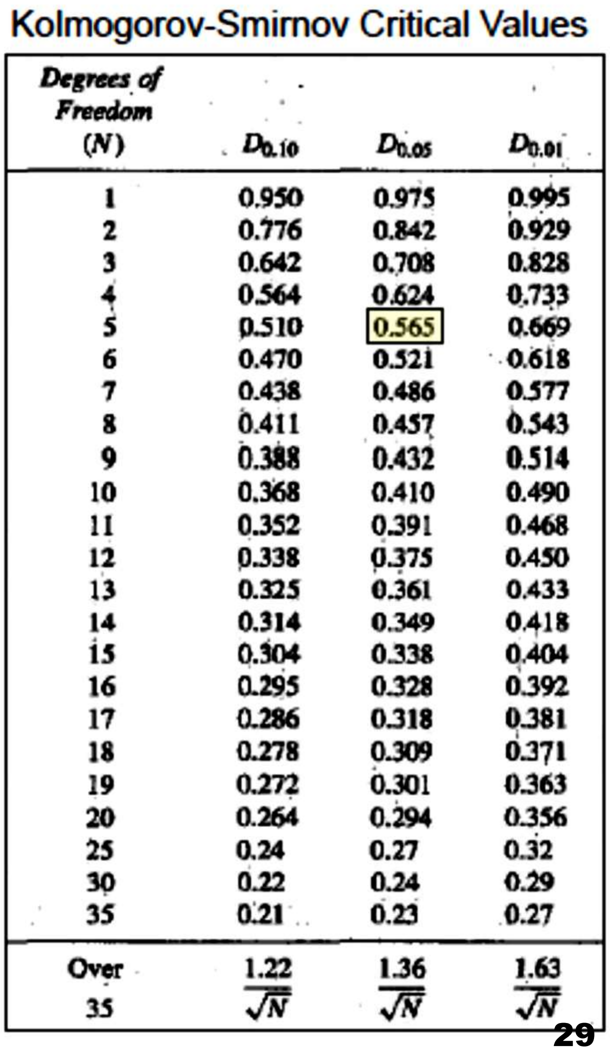
   $$D^+ = \max \{ i/N – R_i \}$$

   $$D^- = \max \{ R_i – (i-1)/N \}$$

3. Compute $D = \max (D^+ , D^-)$
4. Locate the critical value of D_α for the significant level α and the given sample N in table A.8
5. If $D > D_\alpha$, then, the null hypothesis H0 is rejected.
6. If $D <= D_\alpha$, then, there is no difference b/w the true distribution of random numbers generated and the uniform distribution. Hence the null hypothesis is not rejected.

Kolmogorov-Smirnov Test

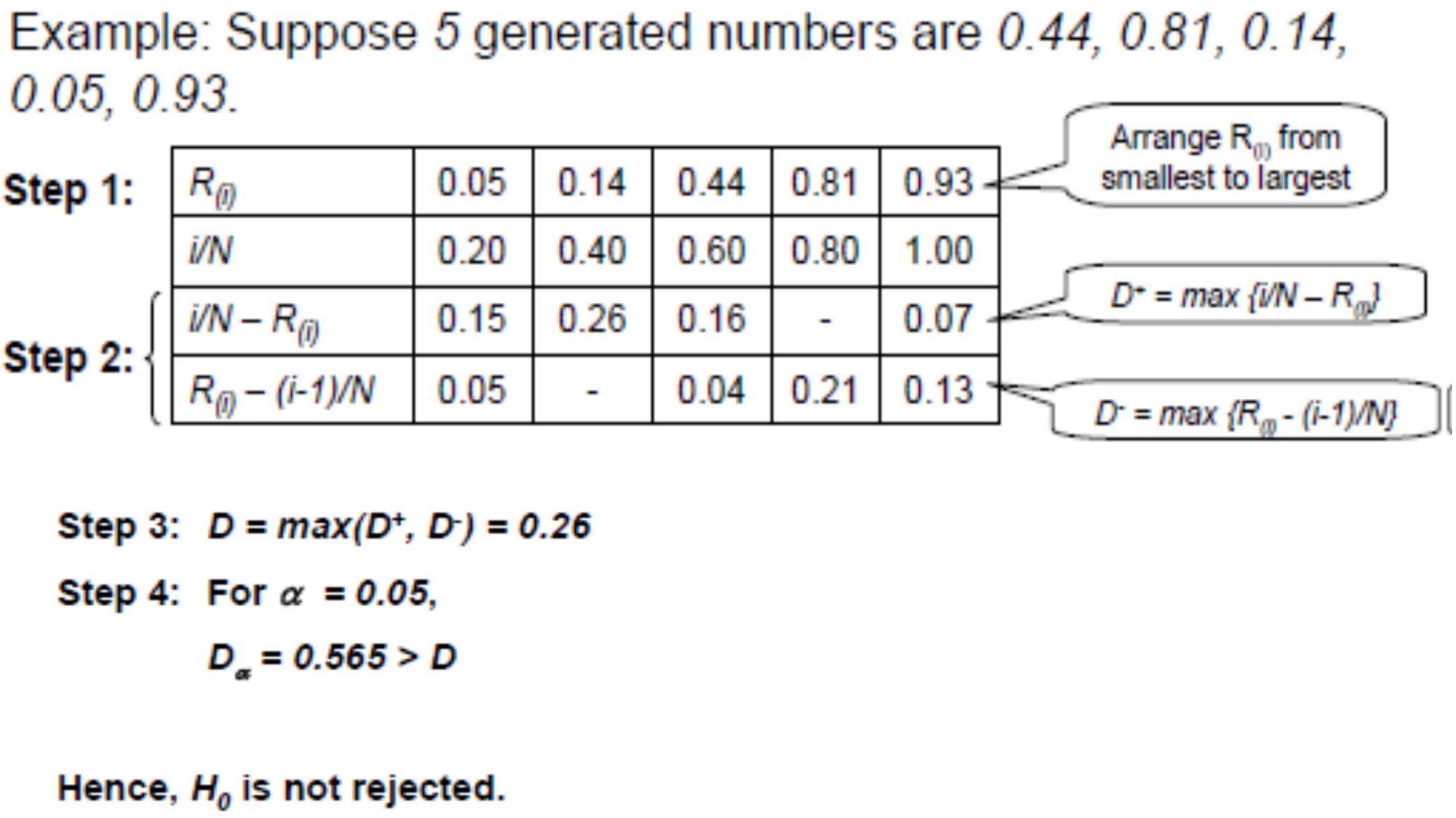
- The test consists of the following steps
  - **Step 1:** Rank the data from smallest to largest
    $R_{(1)} \le R_{(2)} \le \ldots \le R_{(N)}$
  - **Step 2:** Compute
    $$D^{+} = \max_{1 \le i \le N} \left\{ \frac{i}{N} - R_{(i)} \right\}$$
    $$D^{-} = \max_{1 \le i \le N} \left\{ R_{(i)} - \frac{i-1}{N} \right\}$$
  - **Step 3:** Compute $D = \max(D^{+}, D^{-})$
  - **Step 4:** Get D_{α} for the significance level α
  - **Step 5:** If $D \le D_{\alpha}$ accept, otherwise reject H_0



Kolmogorov-Smirnov Critical Values

Degrees of Freedom (N)	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29
35	0.21	0.23	0.27
Over 35	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

Kolmogorov-Smirnov Test



Example: Suppose 5 generated numbers are *0.44, 0.81, 0.14, 0.05, 0.93.*

Step 1:	$R_{(i)}$	0.05	0.14	0.44	0.81	0.93
	i/N	0.20	0.40	0.60	0.80	1.00
Step 2:	$i/N - R_{(i)}$	0.15	0.26	0.16	-	0.07
	$R_{(i)} - (i-1)/N$	0.05	-	0.04	0.21	0.13

Arrange $R_{(i)}$ from smallest to largest

$D^{+} = max\ \{i/N - R_{(i)}\}$

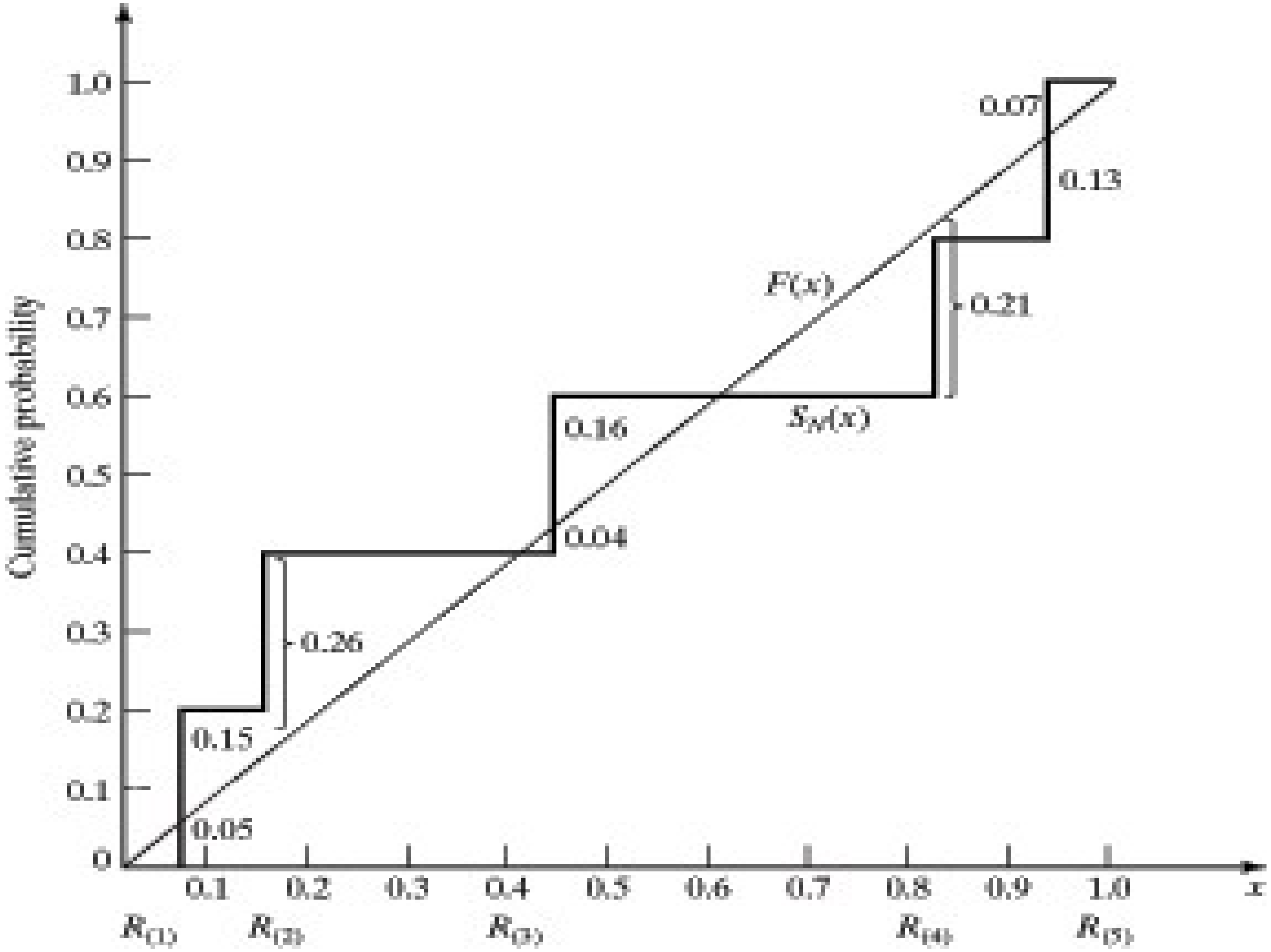
$D^{-} = max\ \{R_{(i)} - (i-1)/N\}$

Step 3: $D = max(D^{+}, D^{-}) = 0.26$

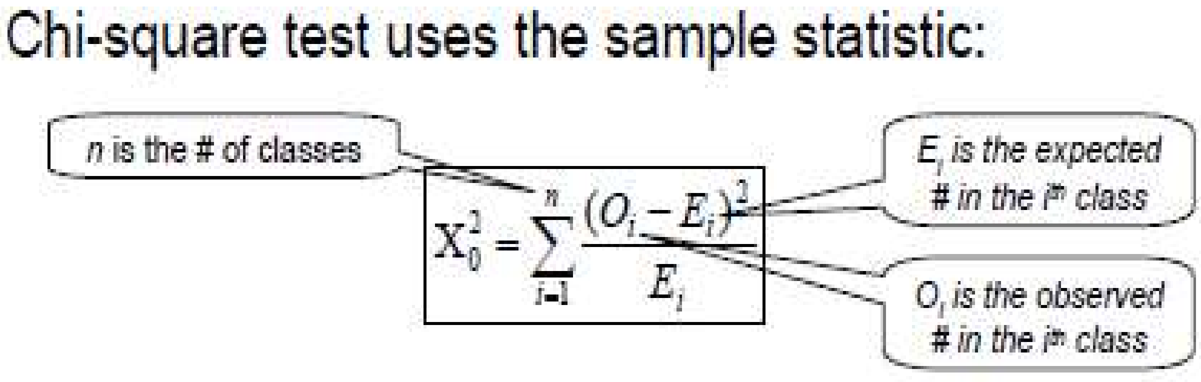
Step 4: For $\alpha\ = 0.05$,

$D_{\alpha} = 0.565 > D$

Hence, H_0 is not rejected.


Cumulative probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.07
0.13
F(x)
0.21
0.16
$S_N(x)$
0.04
0.26
0.15
0.05
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
x
$R_{(1)}$
$R_{(2)}$
$R_{(3)}$
$R_{(4)}$
$R_{(5)}$

The Chi-Square Test



Chi-square test uses the sample statistic:

$$X_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

n is the # of classes

E_i is the expected # in the i^{th} class

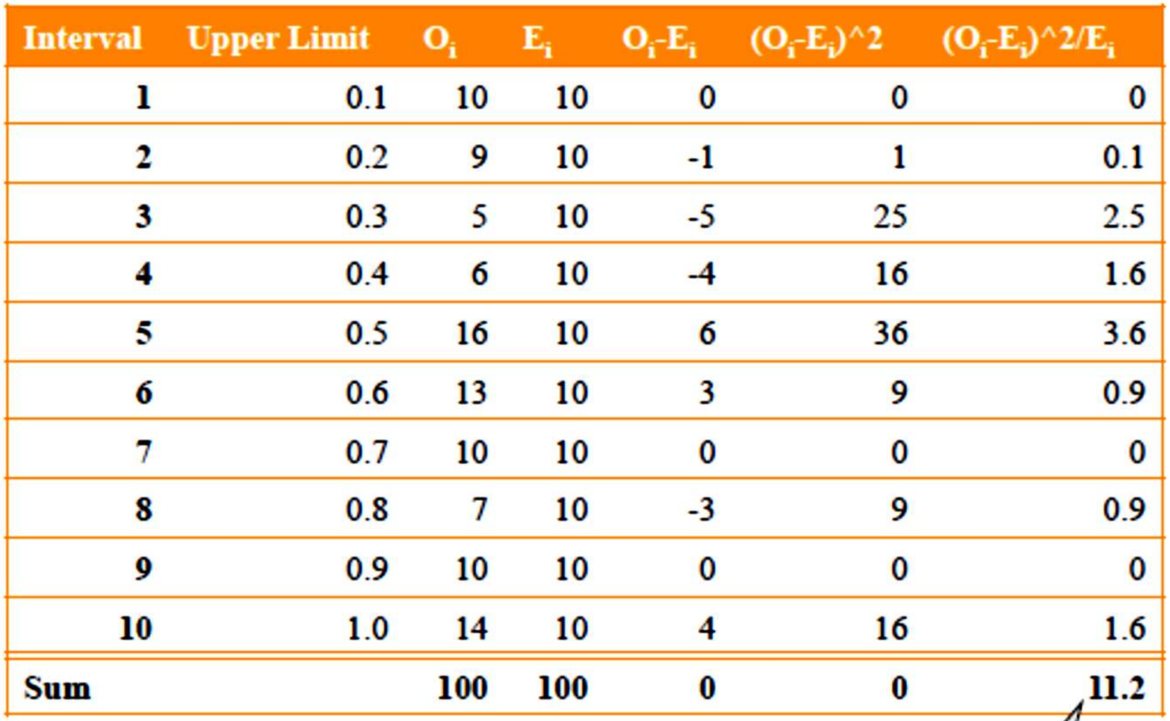
O_i is the observed # in the i^{th} class

- N is the total number of observations.
- O_i is the observed number in the i^{th} class,
- n is the number of classes,
- E_i is the expected number in each class,

  given by : $E_i = N/n$ for equally spaced classes.
- It can be shown that the sampling distribution X^2 is approximately the chi-square distribution with n-1 degrees of freedom.

Chi-square test ; Example

- Example with 100 numbers from [0,1], α=0.05
- 10 intervals
- $\chi^2_{0.05,9} = 16.9$
- Accept, since
  - X^2_0=11.2 < $\chi^2_{0.05,9}$



Interval	Upper Limit	O_i	E_i	O_i-E_i	(O_i-E_i)^2	(O_i-E_i)^2/E_i
1	0.1	10	10	0	0	0
2	0.2	9	10	-1	1	0.1
3	0.3	5	10	-5	25	2.5
4	0.4	6	10	-4	16	1.6
5	0.5	16	10	6	36	3.6
6	0.6	13	10	3	9	0.9
7	0.7	10	10	0	0	0
8	0.8	7	10	-3	9	0.9
9	0.9	10	10	0	0	0
10	1.0	14	10	4	16	1.6
Sum		100	100	0	0	11.2

X^2_0=11.2

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Tests for Autocorrelation

- The tests for autocorrelation are concerned with the dependence b/w numbers in a sequence.
- Consider the sequence of numbers

0.12 0.01 0.23 0.28 0.89 0.31 0.64 0.28 0.83 0.93
0.99 0.15 0.33 0.35 0.91 0.41 0.60 0.27 0.75 0.88
0.68 0.49 0.05 0.43 0.95 0.58 0.19 0.36 0.69 0.87

- It appears that every 5th, 10th, 15th number is a large no.
- 20 – 30 nos. are too small to reject a random no generator, but the notion is that the numbers might be related.
- The autocorrelation ρ_{im} *of* interest between numbers :

$$R_i, R_{i+m}, R_{i+2m}, \ldots., R_{i+(M+1)m}$$

- *M* is the largest integer such that

$$i + (M + 1)m <= N,$$

where N is the total number of values in that sequence.

Tests for Autocorrelation

- Hypothesis:

  H_0*:* ρ$_{i,m}$ *= 0 if numbers are independent*

  H_1*:* ρ$_{i,m}$ *!= 0 if numbers are dependent*

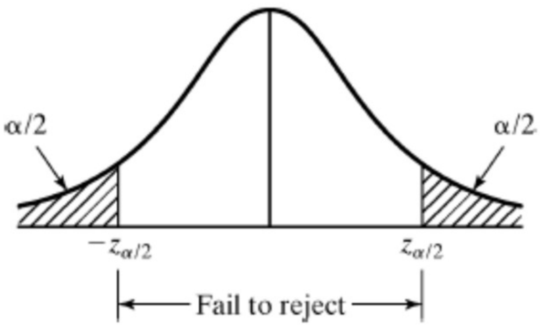
- If the values are uncorrelated:
- For large values of M, the distribution of the estimator of ρ$_{im}$, denoted by (cap)ρ$_{im}$ is approximately normal, if, R_i, R_{i+m}, R_{i+2m}, …., $R_{i+(M+1)m}$ *are uncorrelated.*
- The test statistics is : $$Z_0 = \frac{\hat{\rho}_{im}}{\hat{\sigma}_{\hat{\rho}_{im}}}$$
- *Z0 is distributed normally with mean = 0 and variance =1*

Tests for Autocorrelation

- The formula for $\hat{\rho}_{im}$ in slightly a different form. And the standard deviation of the estimator, given by $\hat{\sigma}_{\rho_{im}}$ Schmidt and Taylor are as follows :

$$\hat{\rho}_{im} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km} R_{i+(k+1)m}\right] - 0.25$$

$$\hat{\sigma}_{\rho_{im}} = \frac{\sqrt{13M+7}}{12(M+1)}$$



- After computing Z_0, do not reject the null hypothesis of independence, if,

$$-Z_{\alpha/2} <= Z_0 <= Z_{\alpha/2}$$

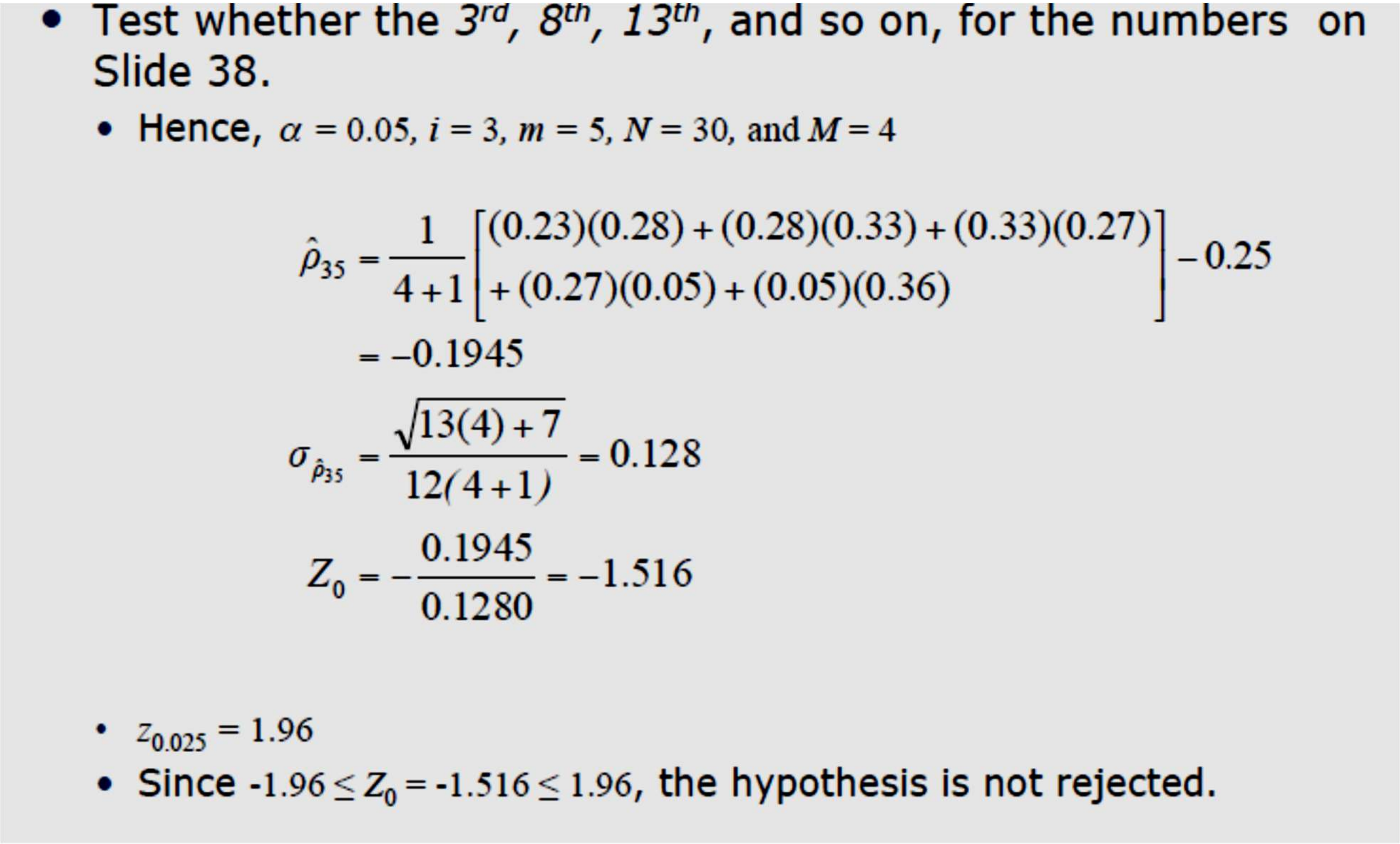
Where α is the level of significance and is obtained from the table A.3. The fig 7.3 illustrates this test.

Tests for Autocorrelation

- **If $\rho_{i,m} > 0$, the subsequence has positive autocorrelation**
  - High random numbers tend to be followed by high ones, and vice versa.

- **If $\rho_{i,m} < 0$, the subsequence has negative autocorrelation**
  - Low random numbers tend to be followed by high ones, and vice versa.

Tests for Autocorrelation



- Test whether the *3rd*, *8th*, *13th*, and so on, for the numbers on Slide 38.
  - Hence, $\alpha = 0.05$, $i = 3$, $m = 5$, $N = 30$, and $M = 4$

$$\hat{\rho}_{35} = \frac{1}{4+1}\begin{bmatrix}(0.23)(0.28)+(0.28)(0.33)+(0.33)(0.27)\\ +(0.27)(0.05)+(0.05)(0.36)\end{bmatrix} - 0.25$$

$$= -0.1945$$

$$\sigma_{\hat{\rho}_{35}} = \frac{\sqrt{13(4)+7}}{12(4+1)} = 0.128$$

$$Z_0 = -\frac{0.1945}{0.1280} = -1.516$$

  - $z_{0.025} = 1.96$
  - Since $-1.96 \le Z_0 = -1.516 \le 1.96$, the hypothesis is not rejected.

Shortcomings

- The test is not very sensitive for small values of M, particularly when the numbers being tested are on the low side.
- Problem when “fishing” for autocorrelation by performing numerous tests:
  - If $\alpha = 0.05$, there is a probability of 0.05 of rejecting a true hypothesis.
  - If 10 independence sequences are examined:
    - The probability of finding no significant autocorrelation, by chance alone, is $0.95^{10} = 0.60$.
    - Hence, the probability of detecting significant autocorrelation when it does not exist = 40%