RANDOM NUMBER

UNIT 5

Random Number

- Random numbers are characterized by the fact that ***their value can not be predicted.***
- Or, in other words, if one constructs a sequence of random numbers, ***the probability distribution of the following random numbers have to be completely independent of all the other generated numbers.***
- Random numbers are samples drawn from a uniformly distributed random variable between some satisfied intervals, ***they have equal probability of occurrence.***

Random Number

- Random numbers are a necessary basic ingredient (element) in the simulation of almost all discrete systems.
- Most computer languages have a subroutine, object, or function that will generate a random number.
- Similarly simulation languages generate random numbers that are used to generate event limes and other random variables.

Random Number Tables

- ***A table of numbers generated in an unpredictable, haphazard (hit-or-miss) that are uniformly distributed within certain interval are called random number table.***
- The random number in random number table exactly obey two random number properties: **uniformity** and **independence** so random number generated form table also called true random numbers

Random Number Tables

- Table of random numbers are used to create a Radom sample.
- A random number table is also called ***random sample table.***
- There are many physical devices or process that can be used to generate a sequence of uniformly distributed random numbers i.e. true random numbers.
- For example: An electrical pulse generator can be made to drive a counter cycling from 0 to 9.
- Using an electronic noise generator or radioactive source the pulse can be generated as random numbers.

Properties of Random Numbers

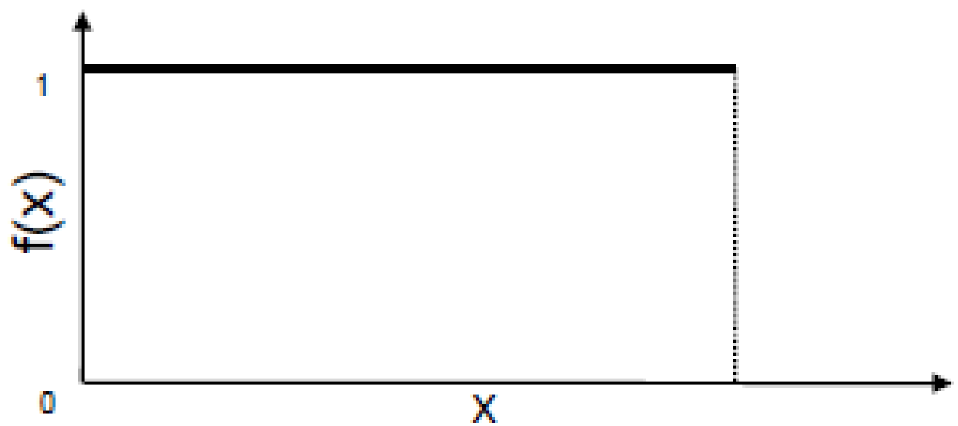
- A sequence of random numbers, R_1, R_2, R_3 … .. must have two important properties:
  - **Uniformity**, i.e. they are equally probable every where
  - **Independence**, i.e. the current value of a random variable has no relation with the previous values

Properties of Random Numbers

- Each random number R_t is an independent sample drawn from a continuous uniform distribution between 0 and 1

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

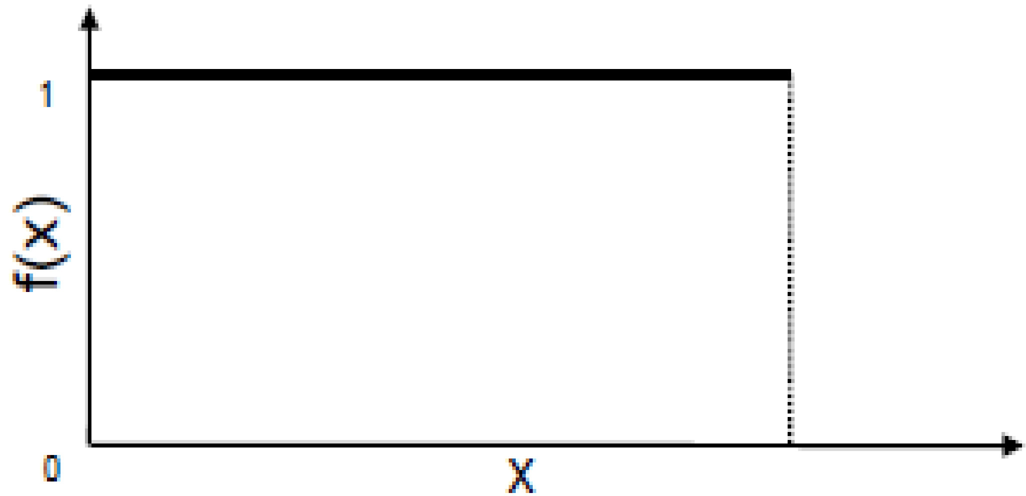
(Probability Distribution Function)



Properties of Random Numbers

- Expectation:

PDF:



$$E(R) = \int_0^1 x dx = [x^2 / 2]_0^1 = 1/2$$

Properties of Random Numbers

- Variance

$$V(R) = \int_0^1 x^2 dx - [E(R)]^2$$
$$= [x^3 / 3]_0^1 - (1/2)^2 = 1/3 - 1/4$$
$$= 1/12$$

So,

$$V(R) = \int_0^1 x^2 dx - [E(R)]^2 = \frac{1}{12}$$

Properties of Random Numbers

Consequences of Uniformity and Independence Properties:

- If the interval (0,1) is divided into n sub-intervals of equal length, the expected number of observations in each interval is N/n, where N is the total number of observations. *Note that N has to be sufficiently large to show this trend.*
- The probability of observing a value in a particular interval is independent of the previous values drawn.

Types of random numbers

- There are three types of random numbers, *quasi*-, *pseudo*- and *true*- random numbers.
- These different types of random numbers have different applications.

1.True Random Number:

The most often used example for "truly" random numbers is the decay of a radioactive material.

If a Geiger counter is put in front of such a radioactive source, the intervals between the decay events are truly random.

True random numbers are gained from physical processes like radioactive decay or also rolling a dice. But rolling a dice is difficult, perhaps someone could control the dice so well to determine the outcome.

Types of random numbers

2.Pseudo Random Number:

These numbers are generated by a computer or that is to say, by an algorithm and because of this not truly random.

Every new number is generated from the previous ones by an algorithm.

This means that the new value is fully determined by the previous ones. But, depending on the algorithm, they often have properties making them very suitable for simulations.

Types of random numbers

3. Quasi Random Number : Quasi (Virtual) random numbers are not designed to appear random, rather to be uniformly distributed.

One aim of such numbers is to reduce and control errors in Monte Carlo simulations.

Pseudo Random Numbers

- Pseudo means false, so false random numbers are being generated.
- The goal of any generation scheme is to produce a sequence of numbers between zero and 1 which simulates, or imitates, the ideal properties of uniform distribution and independence as closely as possible.

Pseudo Random Numbers

- It is not possible to generate a perfect random number that is the random numbers are generated by some known arithmetic operations or formulas which is pseudo-random number or false random number.
- Since the arithmetic operation is known and the sequence of random numbers can be repeated, the numbers cannot be called *truly random number*.
- However the pseudo-random numbers generated by many computer routines very closely fulfill the requirement of desired randomness.

Pseudo Random Numbers

When generating pseudo-random numbers, certain problems or errors can occur. Some examples include the following:

- The generated numbers may not be uniformly distributed.
- The generated numbers may be discrete-valued instead continuous valued. { Numbers are discrete valued and not continuous on [0,1] }
- The mean of the generated numbers may be too high or too low. $E(R) = \frac{1}{2}$
- The variance of the generated numbers may be too high or low. $Var(R) = \frac{1}{12}$
- There may be dependence. The following are the examples:
  a) Autocorrelation between numbers.
  b) Numbers successively higher or lower than adjacent numbers.
  c) Several numbers above the mean followed by several numbers below the mean.

Properties of Good random Number Generators

Usually, random numbers are generated by a digital computer as part of the simulation. Numerous methods can be used to generate the values. In selecting among these methods, or routines, there are a number of important considerations.

Properties of Good random Number Generators

1. **The routine should be fast**. *The total cost can be managed by selecting a computationally efficient method of random-number generation.*
2. **The routine should be portable to different computers, and ideally to different programming languages** *.This is desirable so that the simulation program produces the same results wherever it is executed.*
3. **The routine should have a sufficiently long cycle.** *The cycle length, or period, represents the length of the random-number sequence before previous numbers begin to repeat themselves in an earlier order. Thus, if 10,000 events are to be generated, the period should be many times that long; a special case cycling is degenerating. A routine degenerates when the same random numbers appear repeatedly. Such an occurrence is certainly unacceptable. This can happen rapidly with some methods.*

Properties of Good random Number Generators

4 . **The random numbers should be replicable.** *Given the starting point (or conditions), it should be possible to generate the same set of random numbers, completely independent of the system that is being simulated. This is helpful for debugging purpose and is a means of facilitating comparisons between systems.*

5. Most important, and as indicated previously, the generated random numbers should closely approximate the ideal statistical properties of uniformity and independences

Method / Techniques to Generate Random Numbers

Congruence or Residue Method (Uniform - Linear Congruential Method)

The linear congruential method, initially proposed by Lehrer [1951], produces a sequence of integers, X1, X2,... between zero and m — 1 according to the following recursive relationship:

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

The initial value X_0 is called the **seed**, **a** is called the **constant multiplier**, **c** is the **increment**, and **m** is the **modulus**.

Congruence or Residue Method (Uniform - Linear Congruential Method)

- Case 1: If $c \neq 0$ in Equation (i), the form is called the **mixed congruential** method.
- Case 2: When $c = 0$, the form is known as the **multiplicative congruential** method.

The selection of the values for a, c, m and Xo drastically affects the statistical properties and the cycle length. An example will illustrate how this technique operates.

Congruence or Residue Method (Uniform - Linear Congruential Method)

Mixed Congruential Method: $c \neq 0$

Here ,

$$X_{i+1} = (aX_i + c) \bmod m$$

Let a = 9, c = 3 , m = 31 & $X_0 = 2$

Then , $X_1 = (aX_0 + c) \bmod m$
$= (9 \times 2 + 3) \bmod 31$
$= 21 \bmod 31$
$= 21$

$X_2 = (aX_1 + c) \bmod m$
$= (9 \times 21 + 3) \bmod 31$
$= 192 \bmod 31$
$= 6$

And X_3=26 , X_4=20, X_5=28

Hence, random number are 2,21,6,26,20,28,7,4,8

Congruence or Residue Method (Uniform - Linear Congruential Method)

Multiplicative Congruential Method: c = 0

Here ,

$$X_{i+1} = (aX_i) \bmod m$$

Let a = 9 , m = 31 & X_0 = 2

Then , $X_1 = (aX_0) \bmod m$

$= (9 \times 2) \bmod 31$

$= 18 \bmod 31$

$= 18$

$X_2 = (aX_1) \bmod m$

$= (9 \times 18) \bmod 31$

$= 162 \bmod 31$

$= 7$

And X_3=1 , X_4=9, X_5=19

Hence, random number are 2,18,7,1,9,19,16,20…..

Congruence or Residue Method (Uniform - Linear Congruential Method)

Additive Congrutial Method : **a =1**

Here ,

$$X_{i+1} = (X_i + c) \bmod m$$

Let c = 17 , m = 29 & X_0 = 7

Then , $X_1 = (X_0 + c) \bmod m$

$= (7 + 17) \bmod 29$

$= 24 \bmod 29$

$= 24$

$X_2 = (X_1 + c) \bmod m$

$= (24 + 17) \bmod 29$

$= 41 \bmod 29$

$= 12$

And X_3=0 , X_4=17, X_5=5

Hence, random number are 7,24,12,0,17,5,22,10,27…..

Congruence or Residue Method (Uniform - Linear Congruential Method)

Arithmetic Congruential Method :

In this method random number are generated by the eq:

$$X_{i+1} = (X_{i-1} + X_i) \bmod m$$

Let , X_1=9, X_2=13, m=17

$X_3=(X_1 + X_2) \bmod m = (9 + 13) \bmod 17 = 5$

$X_4=(X_2 + X_3) \bmod m = (13 + 5) \bmod 17 = 1$

…………………

So

The random numbers are 9,13,5,1,6,7,13,3……….

EXAMPLE 3.1

Q.N. Use the linear congruential method to generate a sequence of random numbers with $X_0 = 27$, $a= 17$, $c = 43$, and $m = 100$. Here, the integer values generated will all be between zero and 99 because of the value of the modulus. These random integers should appear to be uniformly distributed the integers zero to 99. Random numbers between zero and 1 can be generated by

$$R_i = X_i/m, \quad i= 1,2,\ldots\ldots (3.1)$$

solution

The sequence of Xi and subsequent Ri values is computed as follows:

$X_0 = 27$

$X_1 = (17x27 + 43) \bmod 100 = 502 \bmod 100 = 2$

$R_1 = 2/100 = 0.02$

$X_2 = (17 \times 2 + 43) \bmod 100 = 77 \bmod 100 = 77$

$R_2 = 77/100 = 0.77$

$X_3 = (17 \times 77 + 43) \bmod 100 = 1352 \bmod 100 = 52$

$R_3 = 52/100 = 0.52$

First, notice that the numbers generated from Equation (3.2) can only assume values from the set i= {0,1 /m, 2/m,..., (m — l)/m), since each Xi is an integer in the set {0,1,2,..., m -1}. Thus, each Ri is discrete on i, instead of continuous on the interval [0, 1], This approximation appears to be of little consequence, provided that the modulus m is a very large integer (Values such as m = 231 -1 and m = 248 are in common use in generators appearing in many simulation languages). By maximum density is meant that the values assumed by Ri = 1,2,..., leave no large gaps on [0,1].

EXAMPLE 3.2

Q.N >Let m = 100, a = 19, c = 0, and X0 = 63, and generate a sequence random integers. Find first 7 random number generate using any suitable method??

Solution

X0 = 63

X1 = (19)(63) mod 100 = 1197 mod 100 = 97

X2 = (19) (97) mod 100 = 1843 mod 100 = 43

X3 = (19) (43) mod 100 = 817 mod 100 = 17

EXAMPLE 3.3

Q.N > Let a = 75, m = 231-1 and c= 0. These choices satisfy the conditions that insure a period of P = m-1. Further, specify a seed, X_0 = 123,457.

Solution:

The first few numbers generated are as follows:

X1= 75(123,457) mod (231 - 1) = 2,074,941,799 mod (231 - 1)

X1 = 165

X2 = 75(2,074,941,799) mod (231 - 1) = 559,872,160 =185

X3 = 75(559,872,160) mod (231 - 1) = 1,645,535,613 =

EXAMPLE 3.4

Q.N > Using the multiplicative congruential method, find the period of the generator for $a = 13$, $m = 64$, and $X_0 = 1, 2, 3,$ and 4. Prove that the solution is given , when the seed is 1 and 3, the sequence has period 16, a period of length eight is achieved when the seed is 2 and a period of length four occurs when the seed is 4.

solution

Period Determination Using Various seeds

i	X_i	X_i	X_i	X_i
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	

Q.N.1>Let m = 47, a = 19, and X_0= 46, and generate a sequence c random integers. Find first 4 random number generate using any suitable method??

Q.n.2>Let m = 100, a = 19, c = 6, and X_0 = 63, and generate a sequence c random integers. Find first 5 random number generate using any suitable method??

Q.N.3>Find random number, with first two random number is 10 and 15 respectively with modulus 50 .

Testing for Randomness

The desirable properties of random numbers — **uniformity** and **independence** to ensure that these desirable properties are achieved, a number of tests can be performed (fortunately, the appropriate tests have already been conducted for most commercial simulation software}.

The tests can be placed in two categories according to the properties of interest.

a) ***Testing for uniformity***
b) ***Testing for independence.***

Testing for Randomness

The desired properties of random numbers are **uniformity** and **independence**. So, the test of random numbers means uniformity and independence test. There are different types of test used for these purpose. They are as follows:

1. Frequency test:

Uses the **Kolmogorov-Smirnov (KS)** or **Chi-square test** to compare the distribution of the set of numbers generated to a uniform distribution

2. Runs test:

Tests the runs up and down or the runs above or below the mean by comparing the actual value to expected value.

The statistics for comparison in **Chi-square.**

3. Auto correlation test:

Tests the correlation between numbers and compares the sample correlation to the expected correlation of zero.

4. Gap test:

Counts the number of digits that appear between repetition of a particular digit and then uses KS test to compare with the expected size of gaps.

5. Poker test:

Treats numbers group together as a poker hand. Then the hands obtained are compared to what is expected using the **Chi-square test**.

Testing for uniformity

The testing for uniformity can be achieved through different frequency test.

These tests use the **Kolmogorov-Smirnov** or the **chi- square** test to compare the distribution of the set of numbers generated to a uniform distribution. Hence in this category we will discuss two types of test

Testing for uniformity

1.The Kolmogorov-Smirnov (KS) test.

This test compares the continuous cdf, F(x), of the uniform distribution to the empirical cdf, $S_N(x)$, of the sample of N observations.

By definition,

F(x) = x, 0 <= x <= 1

If the sample from the random-number generator is $R_1, R_2, \ldots R_N$, then the empirical cdf, $S_N(X)$, is defined by

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

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

1.The Kolmogorov-Smirnov (KS) test.

The **Kolmogorov-Smirnov test** is based on the largest absolute deviation or difference between F(x) and $S_N(X)$ over the range of the random variable.

i.e. it is based on the statistic

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

Algorithm for K-S test

Step 1. Rank the data from smallest to largest. Let $R_{(i)}$ denote the i th smallest observation, so that $R_{(1)} <= R_{(2)} <= \cdots <= R_{(N)}$

Step 2. Compute

$$D^{+} = \max\left\{\frac{i}{N} - R_i\right\}$$

$$D^{-} = \max\left\{R_i - \frac{(i-1)}{N}\right\}$$

Step 3: Compute D=max{D+, D⁻}

Step 4. Determine the critical value, Dα, from Table A.8(in your Text book) for the specified significance level α and the given sample size N.

Step 5.

If the sample statistic D is greater than the critical value, Dα, the null hypothesis that the data are a sample from a uniform distribution is rejected.

If D <= Da, conclude that no difference has been detected between the true distribution of { R1 R2, ,• • •, Rn } and the uniform distribution. Hence the null hypothesis is accepted.

Table A.8 Kolmogorov-Smirnov Critical Values

n \ α	0.01	0.05	0.1	0.15	0.2
1	0.995	0.975	0.950	0.925	0.900
2	0.929	0.842	0.776	0.726	0.684
3	0.828	0.708	0.642	0.597	0.565
4	0.733	0.624	0.564	0.525	0.494
5	0.669	0.565	0.510	0.474	0.446
6	0.618	0.521	0.470	0.436	0.410
7	0.577	0.486	0.438	0.405	0.381
8	0.543	0.457	0.411	0.381	0.358
9	0.514	0.432	0.388	0.360	0.339
10	0.490	0.410	0.368	0.342	0.322
11	0.468	0.391	0.352	0.326	0.307
12	0.450	0.375	0.338	0.313	0.295
13	0.433	0.361	0.325	0.302	0.284
14	0.418	0.349	0.314	0.292	0.274
15	0.404	0.338	0.304	0.283	0.266
16	0.392	0.328	0.295	0.274	0.258
17	0.381	0.318	0.286	0.266	0.250
18	0.371	0.309	0.278	0.259	0.244
19	0.363	0.301	0.272	0.252	0.237
20	0.356	0.294	0.264	0.246	0.231
25	0.320	0.270	0.240	0.220	0.210
30	0.290	0.240	0.220	0.200	0.190
35	0.270	0.230	0.210	0.190	0.180
40	0.250	0.210	0.190	0.180	0.170
45	0.240	0.200	0.180	0.170	0.160
50	0.230	0.190	0.170	0.160	0.150
OVER 50	$\frac{1.63}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.07}{\sqrt{n}}$

EXAMPLE 3.5

Q.N > Suppose that the five numbers 0.44 , 0.81, 0.14, 0.05, 0.93 were generated, and it is desired to perform a test for uniformity using the Kolmogorov-Smirnov test with a level of significance a of 0.05.

Solution

First, the numbers must be ranked from smallest to largest

i.e. 0.05 , 0.14 , 0.44 , 0.81 , 0.93

Then ,

The computations for **D+**, namely ***{ i /N -R(i) }*** and for **D-**, namely ***{ R(i) - (i - l) / N }*** ,

R(i)	0.05	0.14	0.44	0.81	0.93

The statistics are computed as **D+ = 0.26** and **D- = 0.21**.

Therefore,

D = max{0.26, 0.21} = **0.26.**

The critical value of D, obtained from Table A.8 for a = 0.05 and N= 5, is 0.565.

Since the computed value, 0.26, is less than the tabulated critical value, 0.565, the hypothesis of no difference between the distribution of the generated numbers and the uniform distribution is not rejected.

Example 3.6

- Suppose that the five numbers 0.24 , 0.80, 0.11, 0.05, 0.93 were generated, and it is desired to perform a test for uniformity using the Kolmogorov-Smirnov test with a level of significance a of 0.01.

Example 3.7

- Suppose that the four numbers 0.80, 0.14, 0.05, 0.5 were generated, and it is desired to perform a test for uniformity using the Kolmogorov-Smirnov test with a level of significance a of 0.10.

Example 3.8

- Suppose that the seven numbers 0.44 , 0.81, 0.14, 0.05, 0.93, 0.01, 0.02 were generated, and it is desired to perform a test for uniformity using the Kolmogorov-Smirnov test with a level of significance a of 0.05.

The Chi-square Test

- The chi-square test uses the sample statistic

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

Where,

O_i is the observed number in the i-th class,

E_i is the expected number in the i-th class, and

n is the number of classes.

For the uniform distribution, **Ei** the expected number in each class is given by:

Ei = N/n

for equally spaced classes, where **N** is the total number of observations. It can be shown that the sampling distribution of X_0^2 is approximately the chi-square distribution with **n - 1** degrees of freedom

The Chi-square Test

Algorithm

Step 1: Determine Order Statistics

R1<=R2<=…….Rn

Step 2: Divided Range Rn – R1 in n equidistant intervals [ai,bi], such that each interval has at least 5 observations.

Step 3: Calculate for i = 1,……,N

O_i=N . { $S_N(b_i)$ - S_N(ai) },

E_i=N . { F(bi) – F(ai) }

Step 4: Calculate

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

Step 5: Determine for significant level a , $X^2_{a,n-1}$

$$\begin{cases} \chi_0^2 \le \chi_{\alpha,n-1}^2 & \text{Accept: No Difference between } S_N(x) \text{ and } F(x) \\ \chi_0^2 > \chi_{\alpha,n-1}^2 & \text{Reject: Difference between } S_N(x) \text{ and } F(x) \end{cases}$$

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

Example 3.8

Q.N > Use the chi-square test with a = 0.05 to test whether the data shown below are uniformly distributed.

0.34	0.83	0.96	0.47	0.79	0.99	0.37	0.72	0.06	0.18	0.90
0.76	0.99	0.30	0.71	0.17	0.51	0.43	0.39	0.26	0.25	0.79
0.77	0.17	0.23	0.99	0.54	0.56	0.84	0.97	0.89	0.64	0.67
0.82	0.19	0.46	0.01	0.97	0.24	0.88	0.87	0.70	0.56	0.56
0.82	0.05	0.81	0.30	0.40	0.64	0.44	0.81	0.41	0.05	0.93
0.66	0.28	0.94	0.64	0.47	0.12	0.94	0.52	0.45	0.65	0.10
0.69	0.96	0.40	0.60	0.21	0.74	0.73	0.31	0.37	0.42	0.34
0.58	0.19	0.11	0.46	0.22	0.99	0.78	0.39	0.18	0.75	0.73 0.79
0.29	0.67	0.74	0.02	0.05	0.42	0.49,	0.49	0.05	0.62	0.78

Solution

Class interval(i)	Oi	Ei	(Oi – Ei)	$(Oi-Ei)^2$	$(Oi-Ei)^2/Ei$
1	8	10	-2	4	0.4
2	8	10	-2	4	0.4
3	10	10	0	0	0.0
4	9	10	-1	1	0.1
5	12	10	2	4	0.4
6	8	10	-2	4	0.4
7	10	10	0	0	0.0
8	14	10	14	16	0.16
9	10	10	0	0	0.0
10	11	10	1	1	0.1
Total	N=100	N=100			$\sum = 3.4$

Above Table contains the essential computations for chi square test. The test uses n = 10 intervals of equal length, namely [0.0, 0.1), [0.1, 0.2), . . . , [0.9, 1.0). The value of X^2 is 3.4.

Here degree of freedom is n-1=10-1=9 and α=0.05. The tabulated value of $X^2_{0.05, 9}$ =16.9.Since X_0^2 is much smaller than the tabulated value of chi square, the null hypothesis of a uniform distribution is not rejected.

Both the Kolmogorov-Smirnov and the chi-square test are acceptable for testing the uniformity of a sample of data, provided that the sample size is large. However, the Kolmogorov-Smirnov test is the more powerful of the two and is recommended. Furthermore, the Kolmogorov-Smirnov test can be applied to small sample sizes, whereas the chi-square is valid only for large samples, say N>=50.

Imagine a set of 100 numbers which are being tested for independence where the first 10 values are in the range 0.01-0.10, the second 10 values are in the range 0.11-0.20, and so on. This set of numbers would pass the frequency tests with ease, but the ordering of the numbers produced by the generator would not be random. The tests in the remainder of this chapter are concerned with the independence of random numbers which are generated. The presentation of the tests is similar to that by Schmidt and Taylor [1970].

Example 3.9

Q.N > Use the chi-square test with a = 0.05 to test whether the data shown below are uniformly distributed

In first ranges there are 15 random number , in second there are 5 random number, in 3rd there are 10 random number, in fourth there are 10 and in 5 th there are 20 random number.

Example 3.10

Q.N > Use the chi-square test with a = 0.99 to test whether the data shown below are uniformly distributed

In first ranges there are 10 random number , in second there are 10 random number, in 3rd there are 15 random number, in fourth there are 15 , in 5th there are 5 random number. And then 6th ,7th and 8th has 10 , 5 , 10 random numbers.

Example 3.11

Q.N > Use the chi-square test with a = 0.25 to test whether the data shown below are uniformly distributed

In first ranges there are 20 random number , in second there are 9 random number, in 3rd there are 15 random number, in fourth there are 15 and in 5 th there are 13 random number.

Example 3.12

Q.N > Use the chi-square test with a = 0.75 to test whether the data shown below are uniformly distributed

In first ranges there are 25 random number , in second there are 15 random number, in 3rd there are 10 random number, in fourth there are 17 and in 5 th there are 13 random number.

Chi-Square Vs. K-S Test

K-S test	Chi-Square Test
Small samples	Large sample
Continuous Distribution	Discrete Distribution
Difference between Observed and Expected cumulative probabilities (CDF)	Differences between observed and hypothesized probabilities (PDFs or PMFs).
Uses each observation in the sample without any grouping => makes a better use of the data Cell size is not a problem	Group observation into a small number of cells =>Cell sizes effect the conclusion but no firm guidelines
Exact	Approximate

Test for independence includes the three types of tests as given below:

1) **Autocorrelation Test** tests the correlation between numbers and compares the sample correlation to the expected correlation of zero.
2) **Gap test** Counts the number of digits that appear between repetitions of particular digit and then uses the Kolmogorov-Smirnov test to compare with the expected size of gaps,
3) **Poker test**: Treats numbers grouped together as a poker hand. Then the hands obtained are compared to what is expected using the chi-square test.

Tests for Autocorrelation

The tests for autocorrelation are concerned with the dependence between numbers in a sequence. As an example, consider the following 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

From a visual inspection, these numbers appear random, and they would probably pass all the tests presented to this point. However, an examination of the 5th, 10th, 15th (every five numbers beginning with the fifth), and so on indicates a very large number in that position.

Now, 30 numbers is a rather small sample size to reject a random-number generator, but the notion is that numbers in the sequence might be related. In this particular section, a method for determining whether such a relationship exists is described. The relationship would not have to be all high numbers. It is possible to have all low numbers in the locations being examined, or the numbers may alternately shift from very high to very low.

Tests for Autocorrelation

Autocorrelation Test is a statistical test that determines whether a random number generator is producing independent random number in a sequence. The test for the auto correlation is concerned with the **dependence between numbers in a sequence**. The test computes the auto correlation between every m numbers (m is also known as lag) starting with i^{th} index.

The variables involved in this test are:

- **m** is the lag, the space between the number being tested.
- **i** is the index or number from we start.
- **N** is the number of random numbers generated.
- **M** is the largest integer such that i+(M+1)m <= N

Tests for Autocorrelation

- Now the autocorrelation between
  $R_i, R_{i+m}, R_{i+2m}, \ldots\ldots R_{i+(M+1)m}$ is computed as

Now the autocorrelation ρ_{im} between R_i, Ri+m, Ri+2m, ……Ri+(M+1)m is computed as

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

Now the test Statics is

$$Z_0 = \frac{\rho_{im}}{\sigma_{im}}$$

where

$$\sigma_{im} = \frac{\sqrt{13M+7}}{12(M+1)}$$

After computing Z0, do not reject the null hypothesis of independence if $-z_{a/2} \le Z0 \le z_{a/2}$, where a is the level of significance.

Tests for Autocorrelation

Q.N.> Test whether the 3rd, 8th, 13th, and so on, numbers in the sequence at the beginning of this section are auto-correlated. (Use a = 0.05.) Here, $i = 3$ (beginning with the third number), $m = 5$ (every five numbers), $N = 30$ (30 numbers in the sequence).

Solution:

First we calculate the value of M using the condition

$$i + (M+1)m \le N$$

since $i=3$, $m=5$, and $N=30$

we have,

$$3 + (M+1)5 \le 30.$$

i.e. $3+5M+5 \le 30$

I.e. $5M \le 22$

i.e. $M \le 22/5$

Hence $M=4$

Then, ρ_{35} = 1/(4 + 1)[(0.23)(0.28) + (0.28)(0.33) + (0.33)(0.27) + (0.27)(0.05) + (0.05)(0.36)] – 0.25 = -0.1945

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

And

σ_{35}= √ (13(4) + 7) / 12(4 + 1) = 0.1280

Then, the test statistic assumes the value

Z_0 = -0.1945/0.1280 = -1.516

Now, the critical value is

$Z_{0.025}$ = 1.96 ($Z_{\alpha/2}$ is taken in this test)

Therefore, the hypothesis of independence cannot be rejected on the basis of this test.

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

Then, $\rho_{35} = 1/(4 + 1)[(0.23)(0.28) + (0.28)(0.33) + (0.33)(0.27) + (0.27)(0.05) + (0.05)(0.36)] – 0.25 = -0.1945$

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

Q.N.> Test whether the 2nd, 8th, 14th, and so on, numbers in the sequence at the beginning of this section are auto-correlated. (Use a = 0.05.) Here, i = 2 (beginning with the second number), m = 6 (every six numbers), N = 30 (30 numbers in the sequence).

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**

Q.N.> Test whether the 6th, 10th, 14th, and so on, numbers in the sequence at the beginning of this section are auto-correlated. (Use a = 0.05.) Here, i = 6 (beginning with the fifth number), m = 4 (every five numbers), N = 30 (30 numbers in the sequence).

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**

Q.N.> Test whether the 5th, 10th, 15th, and so on, numbers in the sequence at the beginning of this section are auto-correlated. (Use a = 0.05.) Here, i = 5 (beginning with the fifth number), m = 5 (every five numbers), N = 30 (30 numbers in the sequence).

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

Q.N.> Test whether the 2th, 12th, 22th, and so on, numbers in the sequence at the beginning of this section are auto-correlated. (Use a = 0.05.) Here, i = 2 (beginning with the fifth number), m = 10 (every five numbers), N = 30 (30 numbers in the sequence).

Gap test

The gap test is used to determine the significance of the interval between the recurrences of the same digit. A gap of length x occurs between the recurrences of some specified digit. The following example illustrates the length of gaps associated with the digit 3:

4, 1, **3**, 5, 1, 7, 2, 8, 2, 0, 7, 9, 1, **3**, 5, 2, 7, 9, 4, 1, 6, **3** , **3**, 9, 6, **3**, 4, 8, 2, **3**, 1, 9, 4, 4, 6, 8, 4, 1, **3**.

There are 7 three's are there. Thus only six gaps can occur. The first gap is of length 10 and second gap of length 7 and third gap of length zero. And so on. Similarly the gap associated with other digits can be calculated. The theoretical probability of first gap (of length 10 for digit 3) can be calculated as

Gap test

The probability of a particular gap length can be determined by a Bernoulli trail.

$$P(\text{gap of } n) = P(x \neq 3)P(x \neq 3)...P(x \neq 3)P(x = 3)$$

If we are only concerned with digits between 0 and 9, then

$$P(\text{gap of } n) = 0.9^n 0.1$$

The theoretical frequency distribution for randomly ordered digits is given by

$$P(gap \leq x) = F(x) = 0.1 \sum_{n=0}^{x} (0.9)^n = 1 - 0.9^{x+1}$$

Gap test

1. Specify the CDF (Cumulative Distributive frequency) from theoretical frequency distribution given by,

   $F(x) = 1 - 0.9^{x+1}$

   Based on the selected class interval
2. Arrange the observed sample of gaps in cumulative distribution with the same class
3. Find D, the maximum deviation between $F(x)$ and $SN(x)$ as equation,

   $D = |F(x) - SN(x)|$

   Where $S_N(x) = \frac{No.of\ gaps \leq x}{total\ no.of\ gaps}$

4. Determine the critical value of $D\alpha$ from the table for the specified value of α and sample size N. (KS table)
5. If $D_{cal} < D_{\alpha}$, Null hypothesis is not rejected.

Example 3.11

Q.N->Based on the frequency with which gaps occur, analyze the 110 digits below to test whether they are independent. Use α = 0.05.

4, 1, 3, 5, 1, 7, 2, 8, 2, **0**, 7, 9, 1, 3, 5, 2, 7, 9 4, 1, 6, 3, 3, 9, 6, 3, 4, 8, 2, 3, 1, 9, 4, 4, 6, 8, 4, 1, 3, 8, 9, 5, 5, 7, 3, 9, 5, 9, 8, 5, 3, 2, 2, 3, 7, 4, 7, **0**, 3, 6, 3, 5, 9, 9, 5, 5 5, **0**, 4, 6, 8, **0**, 4, 7, **0**, 3, 3, **0**, 9, 5, 7, 9, 5, 1, 6, 6, 3, 8, 8, 8, 9, 2, 9, 1, 8, 5, 4, 4, 5, **0**, 2, 3, 9, 7, 1, 2, **0**, 3, 6, 3

Solution

The number of gaps is given by the number of data values minus the number of distinct digits, or 110 —10 = 100 in the example. The numbers of gaps associated with the various digits are as follows:

Digit	0	1	2	3	4	5	6	7	8	9
# of Gaps	7	8	8	17	10	13	7	8	9	13

Gap Length	Frequency	Relative Frequency	Cum. Frequency S(X)	Theoretical Frequency F(X)	$\|F(x) - S_N(x)\|$
0-3	35	0.35	0.35	0.3439	0.0061
4-7	22	0.22	0.57	0.5695	0.0005
8-11	17	0.17	0.74	0.7176	**0.0224**
12-15	9	0.09	0.83	0.8147	0.0153
16-19	5	0.05	0.88	0.8784	0.0016
20-23	6	0.06	0.94	0.9202	0.0198
24-27	3	0.03	0.97	0.9497	0.0223
28-31	0	0.00	0.97	0.9657	0.0043
32-35	0	0.00	0.97	0.9775	0.0075
36-39	2	0.02	0.99	0.9852	0.0043
40-43	0	0.00	0.99	0.9903	0.0003
44-47	1	0.01	1.00	0.9936	0.0064

The critical value of D is given by $D_{0.05} = 1.36 / \sqrt{100} = 0.136$

Since

$D = \max |F(x) - SN(x)| = 0.0224$

is less than $D_{0.05}$,

we do not reject the hypothesis of independence on the basis of this test.

If $D_{cal} < D_{\alpha}$ Null hypothesis is not rejected.

Example 3.12 : Gap test

Q.N.> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use $\alpha = 0.05$.

Gap Length	Frequency
0-4	25
5-9	15
10-14	10
15-19	3
20-24	2
25-29	0
30-34	5
35-39	10
40-44	30

Example 3.13 : Gap test

Q.N.> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use $\alpha = 0.05$.

Gap Length	Frequency
0-2	25
3-5	20
6-8	15
9-11	3
12-14	12
15-17	0
18-20	5
21-23	15
24-26	30

Example 3.14 : Gap test

Q.N.> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use α = 0.2.

Gap Length	Frequency
0-5	15
6-10	20
11-15	5
16-20	3
21-25	12
26-30	15

Example 3.15 : Gap test

Q.N.> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use α = 0.05.

Gap Length	Frequency
0-10	5
11-20	2
21-30	15
31-40	3
41-50	10
51-60	0
61-70	15
71-80	15
81-90	10

Example 3.16 : Gap test

Q.N.> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use α = 0.05.

Gap Length	Frequency
0-9	15
10-19	20
20-29	15
30-39	3
40-49	10

Example 3.17 : Gap test

Q.N.> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use $\alpha = 0.05$.

Gap Length	Frequency
0-1	25
1-3	15
4-6	10
7-9	3
10-12	2
13-15	0

Poker Test

- The Poker Test is the test for independence based on the frequency with which certain digits are repeated with in a series of numbers.
- This test not only tests for the randomness of the sequence of numbers, but also the digits comprising of each of the numbers.

Poker Test

- The expected value of each of the combination of digits in a number is compared with the observed value by means of the chi-square test for independence.
- The acceptance is done if the observed value of ***chi-square*** sums for all the possible combinations of digits is less than the acceptable value for the given degree of freedom at the specified confidence interval.

Poker Test

- This test gets its name from a game of cards called poker
- This test not only tests the randomness of the sequence of numbers, but also the digits comprising of each number
- Every random number of five digits or every sequence of five digits is treated as poker hand.

Poker Test

- 71549 are five different digits
- 55137 would be pair
- 33669 would be two pairs
- 55513 would be three of a kind
- 44477 would be a full house
- 77774 would be four of a kind
- 88888 would be five of a kind
  - The occurrence of five of a kind is rare.

Poker Test

- In 10,000 random and independent numbers of five digits each, you may be expect the following distribution of various combinations.

Five different digits	3024 or 30.24%
pairs	5040 or 50.40 %
Two-pairs	1080 or 10.80 %
Three of a kinds	720 or 7.20 %
Full houses	90 or 0.90 %
Four of a kinds	45 or 0.45 %
Five of a kinds	1 or 0.01 %

Poker Test

- *Poker Test* - based on the frequency with which certain digits are repeated.

Example:

0.255 0.577 0.331 0.414 0.828 0.909 0.303 0.001...

Note: a pair of like digits appear in each number generated.

Poker Test

- Frequency with certain digits are repeated in a series of numbers
- Example
  - 0.255, 0.577, 0.414, 0.828, 0.909, 0.303, 0.001
- Pair of like digits generated
- For three digits: three possibilities
  - **All different**
  - **All equal**
  - **One pair of like digits**

$$P(\text{exactly one pair}) = \binom{3}{2}(0.1)(0.9) = 0.27$$

Given a fixed digit, this digit different

Given a fixed digit, this digit is the same

no. of possibilities

Poker Test

- P**(three different digits)**

$$= P(\text{second different from first})\, P(\text{third different from first and second})$$
$$= (0.9)(0.8) = 0.72$$

- P**(three like digits)**

$$= P(\text{second digit same as first})\, P(\text{third digit same as first and second})$$
$$= (0.1)(0.1) = 0.01$$

Poker test:

- Measure observed frequency for the three cases
  Compute expected frequency E_i
  (probabilities*1000) Perform chi-square test

Poker Test

In 3-digit numbers, there are only 3 possibilities.

P(3 different digits) =

= P(2nd diff. from 1st) * P(3rd diff. from 1st & 2nd)

= (0.9) (0.8) = 0.72

P(3 like digits) =

= P (2nd digit same as 1st) * P(3rd digit same as 1st)

= (0.1) (0.1) = 0.01

P(exactly one pair) = 1 - 0.72 - 0.01 = 0.27

Example 3.18

Q.N.> A sequence of 1000 three-digit numbers has been generated and an analysis indicates that 680 have three different digits, 289 contain exactly one pair of like digits, and 31 contain three like digits. Based on the poker test, are these numbers independent ? Let $\alpha = 0.05$. Test these numbers using poker test for three digits.

Solution

Combination, i	Observed Frequency, (Oi)	Expected Frequency (Ei)	(Oi-Ei)	$(Oi-Ei)^2 / Ei$
Three Different digit	680	0.72X1000=720	-40	2.22
Three Like digit	31	0.01X1000=10	21	44.10
Exactly one pair	289	0.27X1000=270	19	1.33
	1000	1000		47.65

The appropriate degrees of freedom are one less than the number of class intervals. Since
$47.65 > X^2_{0.05,2} = 5.99$ (tabulated value), the independence of the numbers is rejected on the basis of this test. Here 2 or n-1 is the degree of freedom since there are only 3 (n) classes.

Example 3.19

Q.N.> A sequence of three-digit numbers has been generated and an analysis indicates that 380 have three different digits, 389 contain exactly one pair of like digits, and 231 contain three like digits. Based on the poker test, are these numbers independent ? Let $\alpha = 0.05$. Test these numbers using poker test for three digits.

Example 3.20

Q.N.> A sequence of three-digit numbers has been generated and an analysis indicates that 320 have three different digits, 420 contain exactly one pair of like digits, and 160 contain three like digits. Based on the poker test, are these numbers independent ? Let $\alpha = 0.05$. Test these numbers using poker test for three digits.

Example 3.21

Q.N.> A sequence of three-digit numbers has been generated and an analysis indicates that 300 have three different digits, 500 contain exactly one pair of like digits, and 200 contain three like digits. Based on the poker test, are these numbers independent ? Let $\alpha = 0.05$. Test these numbers using poker test for three digits.

In four digit number, there are five different possibilities

P(four different digits)

= 4c4 x 10/10 x 9/10 x 8/10 x 7/10 = 0.504

P (one pair)

= 4c2 x10/10x1/10x9/10x8/10=0.432

P (two pair)

= 4c2/2 x 10/10 x 1/10 x 9/10 x 1/10 = 0.027

P (three digits of a kind)

= 4c3 x 10/10 x 1/10 x 1/10 x 9/10 = 0.036

P (four digits of a kind)

= 4c4 x 10/10 x 1/10 x 1/10 x 1/10 = 0.001

Example 3.22 (TU 2067/ 10mars)

Q.N.> Explain the independence test. A sequence of 1000 four digit numbers has been generated and an analysis indicates the following combinations and frequencies.

Combination (i)	**Observed frequency (Oi)**
Four different digits	560
One pair	394
Two pair	32
Three digits of a kind	13
Four digit of a kind	1
	1000

Based on poker test, test whether these numbers are independent. Use $\alpha=0.05$ and N=4 is 9.49.

solution

Now the calculation table for the Chi-square statistics is:

Combination (i)	**Observed frequency (Oi)**	**Expected frequency(Ei)**	**(oi-Ei)**	**$(oi\text{-}Ei)^2/Ei$**
Four different digits	560	0.504x1000=504	56	6.22
One pair	394	0.432x1000=432	-38	3.343
Two pair	32	0.027x1000=27	5	0.926
Three digits of a kind	13	0.036	-23	14.694
Four digit of a kind	1	0.0001x1000=1	0	0.000
	1000	1000		25.185

Here the calculated value of chi-square is 25.185 which is greater than the given value of chi-square so we reject the null hypothesis of independence between given numbers.

Example 3.23 (TU 2072/ 10 marks)

- Define frequency test for random numbers. Develop the Poker test for tour digit numbers, and use it to test whether a sequence of following 1000-four digit numbers are independent. (Use $\alpha = 0.05\ and\ N = 4\ is\ 9.49$)

Combination i	Observed Frequency 0_i
Four different digit	565
One pair	392
Two pair	17
Three like digits	24
Four like digits	2
	1000

Example 3.24 (TU 2073/ 10 marks)

- Define and develop a Poker test for four-digit random numbers. A sequence of 10,000 random numbers, each of four digits has been generated. The analysis of the numbers reveals that in 5120 numbers all four digits are different, 4230 contain exactly one pair of like digits, 560 contain two pairs, 75 have three digits of a kind and 15 contain all like digits. Use Poker test to determine whether these numbers are independent> (Critical value of chi-square for $\alpha\ 0.05\ and\ N = 4\ is\ 9.49$)

- Define and develop a Poker test for four-digit random

Example 3.24 (TU 2071/ 10 marks)

- What do you mean by uniformity test? Explain the Poker test with example. [ ]

Calculation of Expected Value for Poker Test of 5-Digit Random Numbers

- **5 different Digits**

  P(5diff)= 1 * 0.9 * 0.8 * 0.7 * 0.6 = 0.3024

- **1 Pair and 3 different digits**

  P(1pair)= $^5C_2$ * 1 * 0.1 * 0.9 * 0.8 * 0.7 = 0.5040

- **2 Pairs**

  P(2pairs)= ($^5C_2$)/2 * $^3C_2$ * 1 * 0.1 * 0.9 * 0.1 * 0.8 = 0.108

- **3 of a kind**

  P(3same)= $^5C_3$ * 1 * 0.1 * 0.1 * 0.9 * 0.8= 0.072

Calculation of Expected Value for Poker Test of 5-Digit Random Numbers

- **Full House**

  P(full)= $^5C_2$ * 1 * 0.1 * 0.1 *0.9 * 0.1= 0.009

- **Four of a Kind**

  P(four)= $^5C_1$ * 1 * 0.1 * 0.1 * 0.1 * 0.9 = 0.0045

- **Five of a Kind**

  P(five)= 1 * 0.1 * 0.1 * 0.1 * 0.1 = 0.0001

Q.N.1 Use the linear congruential method to generate a sequence of random numbers with X_0 = 17, a= 10, c = 43, and m = 50.Random numbers between zero and 1 can be generated by (Ri =Xi/m)

Q.N.2 .Let m = 100, a = 29, and X0 = 63, and generate a sequence c random integers. Find first 5 random number generate using any suitable method??

Q.N.3 Suppose that the five numbers 0.4 , 0.8, 0.34, 0.06, 0.3 were generated, and it is desired to perform a test for uniformity using the Kolmogorov-Smirnov test with a level of significance a of 0.05. (*from Table A.8 for a = 0.05 and N= 5, is 0.565.)*

Q.N.4 Use the chi-square test with a = 0.99 to test whether the data shown below are uniformly distributed

In first ranges there are 30 random number , in second there are 9 random number, in 3rd there are 25 random number, in fourth there are 10 and in 5th there are 15 random number. (From table $X^2_{0.99,4}$= 0.297)

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**

Q.N.5> Test whether the 5th, 10th, 15th, and so on, numbers in the sequence at the beginning of this section are auto-correlated. (Use a = 0.05.) Here, i = 5 (beginning with the fifth number), m = 5 (every five numbers), N = 30 (30 numbers in the sequence). {$Z_{0.025} = 1.96$ ($Z_{\alpha/2}$ is taken in this test)}

Q.N.6. Based on the poker test, are these numbers independent ? Let α = 0.05. Test these numbers using poker test for four digits.

Combination (i)	**Observed frequency (Oi)**
Four different digits	300
One pair	400
Two pair	150
Three digits of a kind	35
Four digit of a kind	15

Based on poker test, test whether these numbers are independent. Use α=0.05 and N=4 is 9.49.

Q.N.7> Based on the frequency with which gaps occur, analyze the digits below to test whether they are independent. Use α = 0.05.

Gap Length	Frequency
0-4	35
5-9	5
10-14	10
15-19	4
20-24	1
25-29	0
30-34	5
35-39	15
40-44	25

The critical value of D is given by $D_{0.05} = 1.36 / \sqrt{100} = 0.136$

Q.N.8> A sequence of 1000 three-digit numbers has been generated and an analysis indicates that 780 have three different digits, 189 contain exactly one pair of like digits, and 31 contain three like digits. Based on the poker test, are these numbers independent ? Let α = 0.05. Test these numbers using poker test for three digits. *{$X^2_{0.05, 2}$ = 5.99 (tabulated value)}*

Non Uniform Random Number Generation / Random Variate generation

A random variable is a measurable mapping having some distribution, and a **random Variate is just a member of the co-domain of a random variable. A random Variate is a particular outcome of a random variable.**

Random Variates are the samples generated from a known distribution i.e. Random Variable and Random Variates have an inverse relationship.

Suppose *X* is a random variable which stands for the outcome of tossing a fair dice. So *X* can take value from 1 through 6 with equal probability of 1/6. Now you actually toss a dice and get a number 4. This number is a particular outcome of *X*, and thus a random Variate. If you toss again, you may get another different value.

1. Non Uniform Transformation Method / Inverse Transform Method

The inverse transform technique can be used to sample from the *exponential, uniform, triangular distribution* etc. by inversing the CDF of those probability distributions. The inverse transform technique can be utilized for any distribution when the cdf, F(x), is of a form that its inverse, F^{-1} can be computed easily.

1. Non Uniform Transformation Method / Inverse Transform Method

a) Exponential Distribution

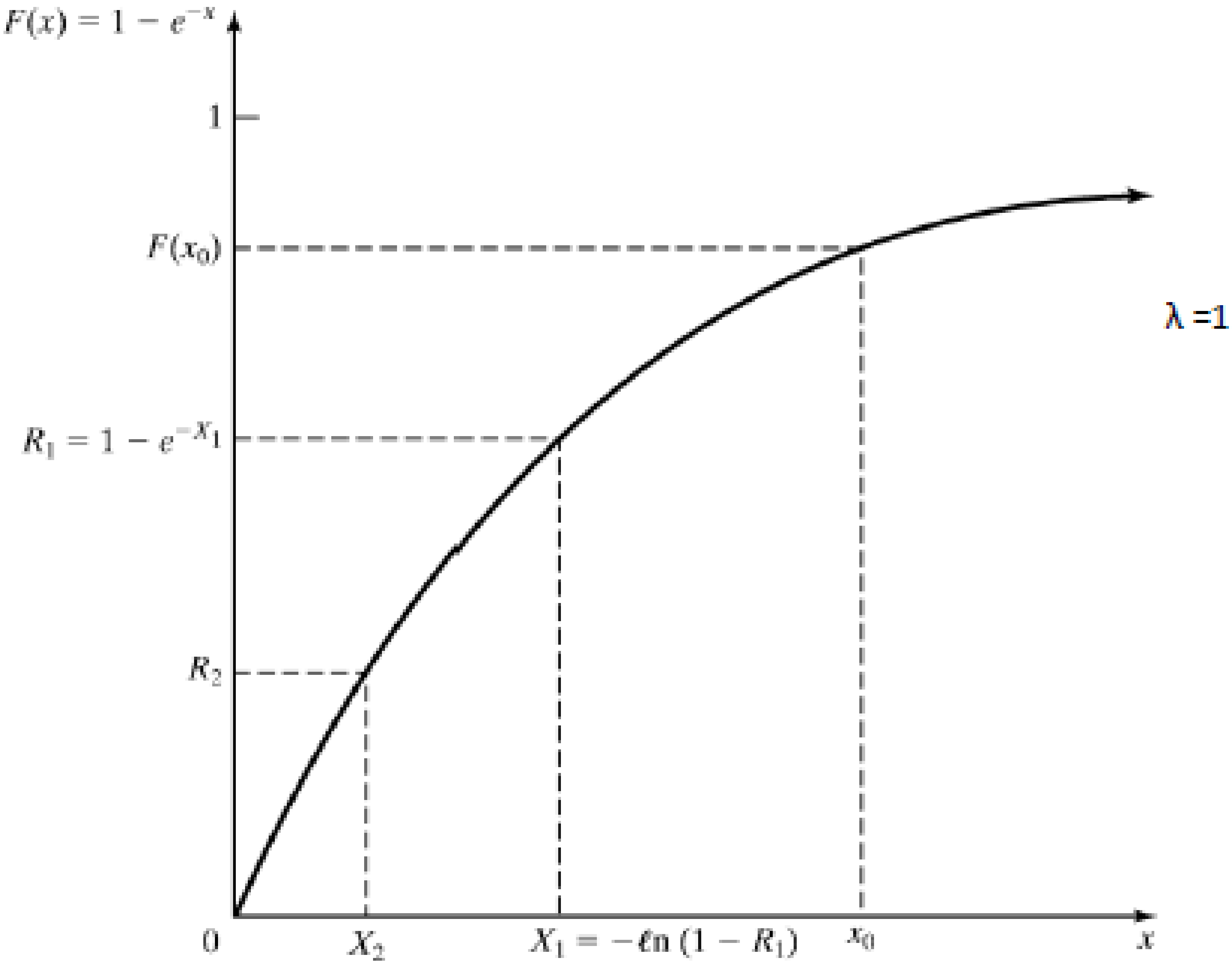
The exponential distribution has the probability function (pdf)

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

and the cumulative distribution function (cdf)

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

λ is the number of occurrences per unit time.



$F(x) = 1 - e^{-x}$
1
$F(x_0)$
$R_1 = 1 - e^{-X_1}$
R_2
0
X_2
$X_1 = -\ell n\,(1 - R_1)$
x_0
x
λ =1

a) Exponential Distribution

The random Variate generation process is summarized in following steps:

Step 1: Compute the cdf of the random variable X for exponential distribution.

Step 2: Set F(X) = R on the range of X i.e. $1 - e^{-\lambda X} = R$

Step 3: Solve the equation $1 - e^{-\lambda X} = R$ in terms of R.

$$1 - e^{-\lambda X} = R$$

$$e^{-\lambda X} = 1 - R$$

$$-\lambda X = \ln(1 - R)$$

$$X = -\frac{1}{\lambda}\ln(1 - R) \quad \text{———————} \quad (1)$$

Equation (1) is called random Variate generator for the exponential distribution. In general equation (1) is written as $X = F^{-1}(R)$

Example 3.16 : Generation of Exponential Variates Xi with mean 1 (λ=1), given random numbers Ri

i	1	2	3	4	5
R_i	0.1306	0.0422	0.6597	0.7965	0.7696

Solution:

$$R_1 = 1 - e^{-\lambda X}$$

$$X_1 = -\frac{1}{\lambda}\ln(1 - R_1)$$

$$X_1 = -\ln(1 - R_1) \qquad (\text{since } \lambda = 1)$$

$X_1 = -\ln(0.1306) = 0.1400$ and so on.

Then, required random Variates are:

X_i	0.1400	0.0431	1.078	1.592	1.468

Example 3.18 : Generation of Exponential Variates Xi with number of occurrence per unit time is 3 , given random numbers Ri

i	1	2	3	4	5	6	7	8	9
Ri	0.32	0.2	0.01	0.121	0.55	0.11	0.4	0.7	0.22

Example 3.19 : Generation of Exponential Variates Xi with number of occurrence per unit time is 2 , given random numbers Ri

i	1	2	3	4	5	6	7	8	9
Ri	0.231	0.12	0.11	0.021	0.45	0.112	0.54	0.731	0.12

Example 3.20: Generation of Exponential Variates Xi with number of occurrence per unit time is 5 , given random numbers Ri

i	1	2	3	4	5	6	7	8	9
Ri	0.1	0.2	0.01	0.31	0.435	0.512	0.514	0.173	0.122

Example 3.17 : Generation of Exponential Variates Xi with number of occurrence per unit time is 8 , given random numbers Ri

i	1	2	3	4	5	6	7	8	9
Ri	0.2	0.1	0.411	0.121	0.415	0.312	0.154	0.531	0.132

Example 3.17 : Generation of Exponential Variates Xi with number of occurrence per unit time is 4 , given random numbers Ri

i	1	2	3	4	5	6	7	8	9
Ri	0.261	0.172	0.11	0.721	0.745	0.172	0.574	0.771	0.172

b) Uniform Distribution:

The pdf for X in uniform distribution is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & otherwise \end{cases}$$

and the cdf is given by

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

Now, set F(X) = $\frac{X-a}{b-a}$ = R

X= a + (b – a) R is the random Variate generator for the uniform distribution.

Generation of **Uniform Distribution** Xi with number of given random numbers Ri , where a=0.2, b=0.9

i	1	2	3	4	5	6	7	8	9
Ri	0.261	0.172	0.11	0.721	0.745	0.172	0.574	0.771	0.172

Generation of **Uniform Distribution** Xi with number of given random numbers Ri , where a=0.15, b=0.8

i	1	2	3	4	5	6	7	8	9
Ri	0.2	0.1	0.411	0.121	0.415	0.312	0.154	0.531	0.132

Generation of **Uniform Distribution** Xi with number of given random numbers Ri , where a=0.2, b=0.9

i	1	2	3	4	5	6	7	8	9
Ri	0.1	0.2	0.01	0.31	0.435	0.512	0.514	0.173	0.122

c) Triangular Distribution

Consider a random variable X that has pdf

$$f(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2 - x, & 1 \le x \le 2 \\ 0, & otherwise \end{cases}$$

This distribution is called triangular distribution with endpoints (0, 2) and mode at 1. Its cdf is given by

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{x^2}{2}, & 0 \le x \le 1 \\ 1 - \frac{(2-x)^2}{2}, & 1 < x \le 2 \\ 1, & x > 2 \end{cases}$$

For $0 < X \le 1$,

$R = X^2/2$

$\therefore X = \sqrt{2R}$

For $1 \le X \le 2$,

$R = 1 - (2 - X)^2/2$

$\therefore X = 2 - \sqrt{2(1-R)}$

Generation of **Triangular Distribution** Ri with number of given random numbers Xi

i	1	2	3	4	5	6	7	8	9
xi	0.1	0.2	0.01	1.31	0.435	0.512	0.514	1.173	2.122

Generation of **Triangular Distribution** Ri with number of given random numbers Xi

i	1	2	3	4	5	6	7	8	9
xi	0.231	1.2	0.01	1.31	2.435	0.512	0.514	1.173	2.122

2. Acceptance /Rejection method

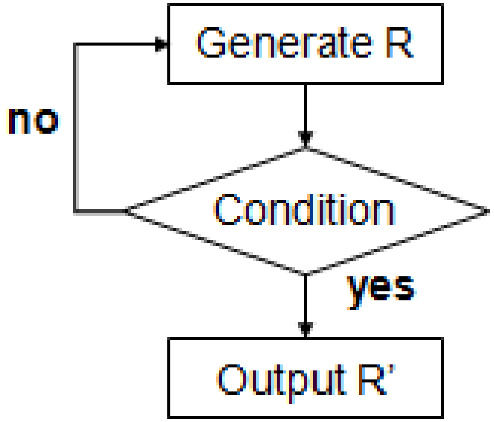
- Useful particularly when inverse cdf does not exist in closed form
- Illustration: To generate random Variates, $X \sim U(1/4, 1)$

Procedures:

Step 1. Generate R ~ U[0,1]

Step 2a. If R >= ¼, accept X=R.

Step 2b. If R < ¼, reject R, return to Step 1



- R does not have the desired distribution, but R conditioned (R') on the event $\{R \geq ¼\}$ does.
- Efficiency: Depends heavily on the ability to minimize the number of rejections.

Finished Unit 3 !!!!