

UNIT-5Random Numbers

The numbers generated by a process, whose outcome is unpredictable, and which cannot be subsequently reliably reproduced are called random numbers. Random numbers are basic ingredient in the simulation of almost all discrete systems. Computer languages have a subroutine, object, or function that will generate a random number.

### Properties of random numbers: [Imp]

Random numbers have two important properties:

i) Uniformity: (i.e, they are equally probable every where).

ii) Independence: (i.e, the current value of random variable has no relation with the previous values).

⇒ Each random number  $R_i$  is an independent sample drawn from a continuous uniform distribution between zero and one.

↳ pdf is given by;  $f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$

↳ expectation is given by:  $E(R) = \int x dx = \frac{1}{2}$

↳ variance is given by:  $V(R) = \int x^2 dx - [E(R)]^2 = \frac{1}{12}$

⇒ Some consequences of the uniformity and independence properties are:

↳ 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.

[Imp] asked in micro-syllabus/model set

Types of random numbers

## ⊗. Differences between true and pseudo random numbers:

| Pseudo-random   | True-random  |
|---|--|
| i) It is an algorithm of mathematical formula, later translated into relatively bits of programming code. | i) It is an extract randomness from physical phenomena and introduce it into a computer. |
| ii) It has fast responses in generating numbers.  | ii) It has slow responses in generating numbers.   |
| iii) In this, sequence of numbers can be reproduced.  | iii) In this, sequence of numbers cannot be reproduced.                                  |
| iv) In this, sequence of numbers is repeated.   | iv) In this, sequence of numbers will or will not get repeated.                          |
| v) It is deterministic  | v) It is non-deterministic.  |

## ⊗. Methods of generation of Random Numbers:

### 1) Linear Congruential Method: [Imp]

The linear congruential method produces a sequence of integers,  $x_1, x_2, \dots$  between zero and  $m-1$  according to the following recursive relationship:

$$x_{q+1} = (ax_q + c) \bmod m, q = 0, 1, 2, \dots \quad (1)$$

The initial value  $x_0$  is called the seed,  $a$  is called the constant multiplier,  $c$  is the increment, and  $m$  is the modulus.

Case 1: If  $c \neq 0$  in eqn (1), the form is called 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  $x_0$  drastically affects the statistical properties and the cycle length.

Example 1: Use the linear congruential method to generate a sequence of random numbers with  $X_0=27$ ,  $a=17$ ,  $c=43$ , and  $m=100$ .

Solution:

theory for understanding only

Here, the integer values generated will all be between 0 and 99 because of the value of the modulus. These random integers should appear to be uniformly distributed between the integers 0 to 99. Random numbers between 0 and 1 can be generated by  $R_g = X_g/m$ ,  $g=1, 2, \dots$

The sequence of  $X_g$  and subsequent  $R_g$  values is computed as follows:

$$X_0 = 27$$

$$\begin{aligned} X_1 &= (17 \times 27 + 43) \bmod 100 \\ &= 502 \bmod 100 \\ &= 2 \end{aligned}$$

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

$$R_2 = \frac{77}{100} = 0.77$$

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

$$R_3 = \frac{52}{100} = 0.52$$

We use left side value of  $a$  if given like this

We keep right side value for actual solving

Example 2: Let  $a=75=16,807$ ,  $m=231-1=2,147,483,647$  (a prime number), and  $c=0$ . These choices satisfy the conditions that ensure a period of  $P=m-1$ . Further, specify a seed  $X_0=123,457$ . Calculate/Generate sequence of random numbers.

Solution:

Given,  $a=75$ ,  $m=(231-1)$ ,  $c=0$  &  $X_0=123,457$ .

The first few numbers generated are as follows:

$$X_1 = (75)(123,457) \bmod (231-1) = 2,074,941,799 \bmod (231-1)$$

$$\text{or, } X_1 = 2,074,941,799$$

$$R_1 = \frac{2,074,941,799}{231-1} = 0.966$$

actual value is  
2,147,483,647  
given

$$X_2 = (75)(2,074,941,799) \bmod (231-1) = 559,872,160.$$

$$R_2 = \frac{X_2}{231-1} = 0.2607$$

$$X_3 = (75) (559, 872, 160) \bmod (231-1) = 1,645, 535, 613$$

$$R_3 = \frac{X_3}{231-1} = 0.7662$$

## 2) Mid-Square Method:

Mid square method is also used to generate pseudo random numbers. For this we follow following steps:

→ Select a seed number ( $X$ ) with  $n$  digits ( $x_1 x_2 \dots x_n$ ).

$X$  is the initial random number. Note:  $n$  is even.

→ Square  $X$  to obtain number ( $Y$ ) with  $m$ -digits.

$$Y_1 Y_2 \dots Y_m = \{x_1 x_2 \dots x_n\}^2$$

→ Add zeros to the left of  $Y$  to form  $Z$  number with  $(2n)$  digits,

$$Z_1 Z_2 \dots Z_{2n} = 00 \dots 0 Y_1 Y_2 \dots Y_m$$

→ Extract the middle  $n$  digits of  $Z$ , which is represented by number  $R$ ,

$$R_1 R_2 \dots R_n = Z_{(n/2)} Z_{(n/2)+1} \dots Z_{(n/2)}$$

Example: Calculate the first five random numbers using mid-square method for seed = 1920.

Solution:

$$1920, \quad (1920)^2 = 3686400, \quad 03686400, \quad \text{then } R_1 = 6864$$

$$6864, \quad (6864)^2 = 47114496, \quad 47114496, \quad \text{then } R_2 = 1144$$

$$1144, \quad (1144)^2 = 1308736, \quad 01308736, \quad \text{then } R_3 = 3087$$

$$3087, \quad (3087)^2 = 9529569, \quad 09529569, \quad \text{then } R_4 = 5295$$

$$5295, \quad (5295)^2 = 28037025, \quad 28037025, \quad \text{then } R_5 = 0370$$

## \* Tests for Randomness:

### a) Uniformity testing:

#### 1) Kolmogorov-Smirnov (K-S) test: [Imp]

→ 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:

$$F(x) = x, \quad 0 \leq x \leq 1.$$

→ If the sample from the random-number generator is  $R_1, R_2, \dots, R_N$ , then the empirical cdf,  $S_N(x)$ , is defined by:

$$S_N(x) = (\text{number of } R_1, R_2, \dots, R_N \text{ which are } \leq x) / N.$$

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

→  $D = \max |F(x) - S_N(x)|$ . When the sampling distribution  $D$  is known, it is tabulated as a function of  $N$  on Table.

#### 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  $R_{(1)} \leq R_{(2)} \leq \dots \leq R_{(N)}$ .

#### 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, $D_{\alpha}$ , from Table.

Step 5: If the sample statistic  $D$  is greater than the critical value  $D_{\alpha}$ , the null hypothesis that the data are a sample from a uniform distribution is rejected.

Step 6: If  $D < D_{\alpha}$ , conclude that no difference has been detected between the true distribution of  $\{R_1, R_2, \dots, R_N\}$  and the uniform distribution. Hence the null hypothesis is accepted.

theory can be escaped  
algorithm and numerical,  
mostly solving numerical  
qs only grp. xv

Example: 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 of 0.05.

Solution:

Ranking numbers from smallest to largest we get,

$$0.05, 0.14, 0.44, 0.81, 0.93.$$

The calculations can be facilitated by use of Table below:

| $R_g$           | 0.05 | 0.14 | 0.44 | 0.81 | 0.93 |
|-----------------|------|------|------|------|------|
| $i/N$           | 0.2  | 0.4  | 0.6  | 0.8  | 1.0  |
| $i/N - R_g$     | 0.15 | 0.26 | 0.16 | -    | 0.07 |
| $R_g - (i-1)/N$ | 0.05 | -    | 0.04 | 0.21 | 0.13 |

Now,

$$\begin{aligned} D^+ &= \max \{ i/N - R_g \} \\ &= \max \{ 0.15, 0.26, 0.16, 0.07 \} \\ &= 0.26 \end{aligned}$$

$$\begin{aligned} D^- &= \max \{ R_g - (i-1)/N \} \\ &= \max \{ 0.05, 0.04, 0.21, 0.13 \} \\ &= 0.21 \end{aligned}$$

$$\therefore D = \max \{ D^+, D^- \} = \max \{ 0.26, 0.21 \} = 0.26.$$

The critical value of  $D$ , obtained from table for  $\alpha = 0.05$  and

$$N=5 \text{ is } 0.565.$$

Since the computed value, 0.26, is less than the tabulated value, 0.565, the hypothesis that the distribution of the generated numbers is the uniform distribution is not rejected.

We can also solve same method as we did in statistics by setting up hypothesis it is better

if values go beyond given values  
We write,  
since here we got 0.95 and  
-0.06 values (two)  
which are outside 0.05 to 0.93 so rejected

table A.8 given in  
book in appendix

If  $D \leq D_{\alpha}$  accepted otherwise rejected.

## 2) 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,  $E_i$  the expected number in each class is given by;

$E_i = N/n$ . for equally spaced classes where  $N$  is the total number of observations.

It can be shown that the sampling distribution of  $\chi^2$  is approximately the chi-square distribution with  $n-1$  degrees of freedom.

Example: Use the chi-square test with  $\alpha=0.05$  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:

The table for chi square statistics is:

| Class Interval ( $i$ ) | $O_i$ | $E_i$ | $(O_i - E_i)$ | $(O_i - E_i)^2$ | $(O_i - E_i)^2 / E_i$ |
|------------------------|-------|-------|---------------|-----------------|-----------------------|
| 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                   |

$$N = 100$$

$$\sum (O_i - E_i)^2 / E_i = 3.4$$

Now,

$$\chi^2_0 = \sum_{i=0}^n \frac{(O_i - E_i)^2}{E_i} = 3.4$$

i.e., critical value  $\chi^2_{0.05, 9}$  tabulated value so accepted  
i.e., not rejected

The value of  $\chi^2_0$  is 3.4.  $\chi^2_{0.05, 9} = 16.9$ . Hence, the data given are uniformly distributed.

### b) Independent Testing:

1) Gap test: Gap test counts the number of digits that appear between repetitions of particular digit and then uses the Kolmogorov-Smirnov test to compare with expected size of gaps. The following example illustrates the length of gaps associated with the digit 3.

4, 1, 3, 5, 1, 7, 2, 8, 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,

There are 7 three's, so only 6 gaps can occur. The first gap is of length 9, second gap of length 7, third gap of length zero and so on.

Example: Based on the frequency with which gap occurs, analyze the 110 digits below to test whether they are independent. Use  $\alpha = 0.05$ .

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,  
i.e.,  $N = 110 - 10 = 100$

The numbers of gaps associated with the various digits are as follows:

no. of repetition - 1

| Digit      | 0 | 1 | 2 | 3  | 4  | 5  | 6 | 7 | 8 | 9  |
|------------|---|---|---|----|----|----|---|---|---|----|
| No. of Gap | 7 | 8 | 8 | 17 | 10 | 13 | 7 | 8 | 9 | 13 |

Upper limit  
of gap length

The calculation for gap test is shown in following table:

$\frac{f}{N}$

(adding each freq)

$1 - 0.9^{\frac{x+1}{x+1}}$   
i.e.  $1 - 0.9^{x+1}$

| Gap length | Frequency $f(f)$ | Relative frequency $\frac{f}{N}$ | Cumulative frequency $S(x)$ | Theoretical Frequency $F(x)$ | $D =  F(x) - S(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                                | 0.97                        | 0.9657                       | 0.0043              |
| 32-35      | 0                | 0                                | 0.97                        | 0.9775                       | 0.0075              |
| 36-39      | 2                | 0.02                             | 0.99                        | 0.9852                       | 0.0043              |
| 40-43      | 0                | 0                                | 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  $D = \max |F(x) - S(x)| = 0.0224$  is less than  $D_{0.05}$ , so

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

## 2) Auto Correlation test:

Auto correlation test is a statistical test that determines whether a random number generator is producing independent random numbers in a sequence. The test for the auto correlation is concerned with the dependence between numbers in sequence.

The variables involved in this test are:

$m \rightarrow$  It is the lag, the space between the number being tested.

$i \rightarrow$  It is the index or number from which we start.

$N \rightarrow$  It is the number of random numbers generated.

$M \rightarrow$  It is the largest integer such that  $i + (M+1)m \leq N$ .

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

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

Now the test statistics is;

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

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

After computing  $Z_0$ , accept the null hypothesis if independence of  $-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$ . where,  $\alpha$  is level of significance.

Example:

0.19, 0.16, 0.82, 0.63, 0.04, 0.16, 0.3, 0.22, 0.88, 0.48, 0.23, 0.56, 0.44, 0.05, 0.81, 0.38, 0.59, 0.37, 0.75, 0.43, 0.92, 0.45, 0.57, 0.99, 0.2, 0.14, 0.64, 0.5, 0.73, 0.15, 0.02, 0.49, 0.86, 0.24, 0.9, 0.74, 0.41, 0.09, 0.86, 0.42, 0.16, 0.23, 0.77, 0.08, 0.69, 0.46, 0.39, 0.18, 0.21, 0.98.

Take  $\alpha = 0.05$ , Test numbers at position 2nd, 7th, 12th are auto-correlated.

Solution:

Here,  $i=2$  and  $m=5$

Now, we have,

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

$$\Rightarrow 2 + (M+1)5 \leq 50$$

$$\Rightarrow 2 + 5M + 5 \leq 50$$

$$\Rightarrow 5M \leq 43$$

$$\Rightarrow M \leq 8.6 \text{ or, } M=8$$

Now,

$$S_{2m} = \frac{1}{M+1} \left[ \sum_{k=0}^M R_{2+km} \cdot R_{2+(k+1)m} \right] - 0.25$$

$$\text{or, } S_{25} = \frac{1}{8+1} \left[ \sum_{k=0}^8 R_{2+k \times 5} \cdot R_{2+(k+1)5} \right] - 0.25$$

$$= \frac{1}{9} \left[ R_2 R_7 + R_7 \cdot R_{12} + R_{12} R_{17} + R_{17} R_{22} + R_{22} R_{27} + R_{27} R_{32} + R_{32} R_{37} \right. \\ \left. + R_{37} \cdot R_{42} + R_{42} \cdot R_{47} \right] - 0.25$$

$$= \frac{1}{9} \left[ 0.16 \times 0.3 + 0.3 \times 0.56 + 0.56 \times 0.59 + 0.59 \times 0.45 + 0.45 \times 0.64 \right. \\ \left. + 0.64 \times 0.49 + 0.49 \times 0.41 + 0.41 \times 0.23 + 0.23 \times 0.39 \right] - 0.25$$

$$= 0.172$$

Again,

$$\sigma_{2m} = \sigma_{25} = \sqrt{\frac{13M+7}{12(M+1)}} \\ = \sqrt{\frac{13 \times 8 + 7}{12(8+1)}}$$

$$= 0.0975$$

$$\text{and } Z_0 = \frac{S_{25}}{\sigma_{25}} = \frac{0.172}{0.0975} = 1.764$$

The tabulated value of  $Z_{0.05/2} = 1.96$

Since  $-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$ . Therefore numbers are independent.

### 3) Poker Test: [Imp]

mainly focus on numerical. Numericals are more imp

The poker test for independence is based on the frequency with which certain digits are repeated in a series of numbers. The poker test uses the chi-square statistics to accept or reject the hypothesis. The following example shows an unusual amount of repetition:

- 0.255, 0.577, 0.331, 0.414, 0.828, 0.909, 0.303, 0.001, ...
- In three digit numbers there are only three possibilities as follows:
  - 1) The individual numbers can all be different.
  - 2) The individual numbers can all be same.
  - 3) There can be one pair of like digits.

The probability associated with each of these possibilities is given by the following:

$$\begin{aligned} \text{i)} P(\text{three different digits}) &= P(\text{second different from the first}) \times \\ &\quad P(\text{third different from the first and second}) \\ &= (0.9)(0.8) = 0.72 \end{aligned}$$

$$\begin{aligned} \text{ii)} P(\text{three like digits}) &= P(\text{second digit same as the first}) \times \\ &\quad P(\text{third digit same as the first}) \\ &= (0.1)(0.1) = 0.01. \end{aligned}$$

$$\begin{aligned} \text{iii)} P(\text{exactly one pair}) &= 1 - 0.72 - 0.01 \\ &= 0.27 \end{aligned}$$

Again For Four digit: there are five possibilities;

$$\begin{aligned} \text{i)} P(\text{four different digits}) &= C(4,4) \times P(\text{second different from first}) \times P(\text{third} \\ &\quad \text{different from first and second}) \times P(\text{fourth different from 1st, 2nd and 3rd}) \\ &= 1 \times 0.9 \times 0.8 \times 0.7 \\ &= 0.504 \end{aligned}$$

$$\begin{aligned} \text{ii)} P(\text{one pair}) &= C(4,2) \times P(\text{second similar as first}) \times P(\text{third different from first}) \\ &\quad \times P(\text{fourth different from first and third}) \\ &= 6 \times 0.1 \times 0.9 \times 0.8 \\ &= 0.432 \end{aligned}$$

$$\begin{aligned} \text{iii)} P(\text{three digits of a kind}) &= C(4,3) \times P(\text{second same as first}) \times P(\text{third same as} \\ &\quad \text{first}) \times P(\text{fourth not same as first}) \\ &= 4 \times 0.1 \times 0.1 \times 0.9 \\ &= 0.036 \end{aligned}$$

$$\begin{aligned} \text{iv)} P(\text{four digits of a kind}) &= C(4,4) \times P(\text{second same as first}) \times P(\text{third same as} \\ &\quad \text{first}) \times P(\text{fourth same as first}) \\ &= 1 \times 0.1 \times 0.1 \times 0.1 \\ &= 0.001 \end{aligned}$$

$$\begin{aligned} \text{v)} P(\text{two pair}) &= 1 - P(\text{four different digits}) - P(\text{three digits of a kind}) - \\ &\quad P(\text{four digits of a kind}) - P(\text{one pair}) \\ &= 1 - 0.504 - 0.036 - 0.001 - 0.432 \\ &= 0.027 \end{aligned}$$

### Example: (Three digit):

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 poker test, are the numbers independent?

Let  $\alpha = 0.05$ .

Solution:

#### Defining Hypothesis:

$$H_0: R_i \sim \text{independent}$$

$$H_1: R_i \not\sim \text{independent.}$$

Calculation:

| Combination           | Observed frequency ( $O_i$ ) | Expected Frequency ( $E_i$ ) | $\frac{(O_i - E_i)^2}{E_i}$ |
|-----------------------|------------------------------|------------------------------|-----------------------------|
| Three different digit | 680                          | $0.72 \times 1000 = 720$     | 2.22                        |
| One pair              | 289                          | $0.27 \times 1000 = 270$     | 1.33                        |
| Three like digit      | 31                           | $0.01 \times 1000 = 10$      | 44.1                        |
|                       |                              |                              | $\sum = 47.65$              |

$$\text{Here, } \chi^2 = 47.65$$

$$\text{And, Tabulated } \chi^2_{\alpha, n-1} = \chi^2_{0.05, 3} = 5.99$$

Since, calculated  $\chi^2 >$  tabulated  $\chi^2$ . So,  $H_0$  is rejected and we conclude numbers are dependent.

### Example: (Four digits):

Q. A sequence of 1000 four digit numbers have been generated and an analysis indicates the following combinations and frequencies.

| Combination           | Observed frequency $O_i$ |
|-----------------------|--------------------------|
| Four different digits | 560                      |
| One pair              | 394                      |
| Two pair              | 32                       |
| Three digits of kind  | 13                       |
| Four digits of kind   | 1                        |

Test the independence.

Solution:

Define Hypothesis:

$H_0: R_j \sim \text{Independent}$

$H_1: R_j \propto \text{dependent.}$

Calculation:

| Combination.          | Observed Frequency ( $O_i$ ) | Expected Frequency ( $E_i$ ) | $\frac{(O_i - E_i)^2}{E_i}$ |
|-----------------------|------------------------------|------------------------------|-----------------------------|
| Four different digits | 560                          | $0.504 \times 1000 = 504$    | 6.22                        |
| One pair              | 394                          | $0.432 \times 1000 = 432$    | 3.434                       |
| Two pairs             | 32                           | $0.027 \times 1000 = 27$     | 0.926                       |
| Three digits of kind  | 13                           | $0.036 \times 1000 = 36$     | 14.694                      |
| Four digits of kind   | 1                            | $0.001 \times 1000 = 1$      | 0                           |
|                       |                              |                              | $\sum = 25.165$             |

Here,  $\chi^2 = 25.165$

And Tabulated  $\chi^2_{\alpha, n-1} = \chi^2_{0.05, 4} = 9.49$

Since, calculated value is greater than tabulated value. So,  $H_0$  is rejected. Hence numbers are dependent.

## ④ Random Variate Generation:

### 1) Inverse Transform Method:

The inverse transform method can be used to sample from the exponential, uniform, triangular distribution etc. by inverting 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.

## Uniform Distribution:

The pdf for random variable  $X$  in uniform distribution is given by;  $f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise.} \end{cases}$

and the cdf is given by;

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

Now, Set  $F(x) = \frac{x-a}{b-a} = R$ .

$X = a + (b-a)R$  is the random variate generator for uniform distribution.

## 2) Acceptance / Rejection Method:

It is useful particularly when inverse cdf does not exist in closed form. Suppose that an analyst needed to think up a method for generating random variates,  $X$  uniformly distributed between  $1/4$  and  $1$ . One way to proceed would be to follow these steps:

Step 1: Generate a random number  $R$ .

Step 2: (a) If  $R \geq 1/4$ , accept  $X=R$ , then go to step 3.

(b) If  $R < 1/4$ , reject  $R$ , and return to step 1.

Step 3: If another uniform random variate on  $[1/4, 1]$  is needed, repeat the procedure beginning at step 1. If not, stop.

The efficiency of an acceptance-rejection technique depend heavily on being able to minimize the number of rejections.

## Additional: [Impl]

Q. Write a computer program in C that will generate four digit random numbers using the multiplicative congruential method. Allow the user to input values of  $x_0, a, c$  and  $m$ .

### Solution:

```
#include <stdio.h>
#include <conio.h>
```

```
/* Function to generate random numbers */
void linearCongruentialMethod(int X0, int m, int a, int c)
```

```
int main() {
    int X0, X1; /* X0 = seed, X1 = next random no. that
                   we will generate */
    int a, c, m; /* a = constant multiplier, c = increment,
                   m = modulus */
    int i, n; /* i for loop control, n for no. of random numbers */
    n = 4; /* Given in question */
    int array[20]; /* to store random numbers generated */

    printf("Enter the seed value X0: ");
    scanf("%d", &X0);
    printf("\n");

    printf("Enter the constant multiplier a: ");
    scanf("%d", &a);
    printf("\n");

    printf("Enter the increment c: ");
    scanf("%d", &c);
    printf("\n");

    printf("Enter the modulus m: ");
    scanf("%d", &m);
    printf("\n");
```

```
for (q=0; q<n; q++) /*loop to generate random numbers */.  
{  
    x1 = (a*x0+c)%m;  
    array[q] = x1;  
    x0 = x1;  
}  
  
printf ("The generated random numbers are: ");  
for (q=0; q<n; q++)  
{  
    printf ("%d", array[q]);  
    printf ("\t");  
}  
getch();  
return (0);  
}
```