Anekant Education Society's

Tuljaram Chaturchand College of Arts, Science and Commerce, Baramati

Autonomous

QUESTION BANK

FOR

F.Y.B.Sc (Computer Science) SEM-II

STATISTICS

PAPER-II: CSST-1202

Statistical Testing of Hypothesis and Use of R Software

(With effect from June 2019)

PAPER-II: CSST-1202

Statistical Testing of Hypothesis and Use of R Software

Unit-1 Introduction to R-software

A) Questions for 1 mark:-

I. Choose the correct alternative:-

1. If X→B(n=5,p=0.4), to find P(X=3) we use command

   a) dbinom(5,3,0.4) b) dbinom(3,5,0.4)

   c) dbinom(0.4, 5,3) d) pbinom(3,5,0.4)

2. If X→P(m=4.5), to find P(X≤4) we use command

   a) ppois(4,4.5) b) rpois(4,4.5)

   c) dpois(4,4.5) d) ppois(4.5,4)

3. To obtain random sample of size 5 from P(m=2) distribution we use command

   a) ppois(5,2) b) rpois(5,2)

   c) dpois(2,4) d) ppois(5,2)

4. If X→B(n=10,p=0.25), to find P(X=4) we use command

   a) dbinom(10,4,0.25) b) dbinom(4,10,0.25)

   c) pbinom(10,4,0.25) d) pbinom(4,10,0.25)

5. Transform command is used to

   a) add new records b) add new columns

   c) extracting data d) deletion data

6. If y=c(1,5,1,2,3,7,1,2) then the result of unique(y) command is

   a) 1 5 1 2 3 7 1 2 b) 1 2 3 5 7

   c) 7 5 3 2 1 d) 1 5 2 3 7

7. If y=c(1,5,1,2,3,7,1,2) then the result of unique(y) command is

   a) 1 5 2 3 7 b) 5 3 7

   c) 1 2 d) 1 7

8. If y=c(1,2,3,4,5) then the result of prod(y) command is

   a) 60 b) 120

   c) 15 d) 110

II. State True or False of the following:-

1. Transform command is used to extract elements of vector conditionally.
2. Subset command is used to augment two vectors X and Y, each containing same number of elements.
3. seq( ) function is used to generate a vector elements in sequence.
4. rep( ) function is used to generate a vector with repeated elements.
5. c( ) function is used to generate a vector with repeated elements.
6. length( ) function is used to count the number of elements in vector.
7. R –software is a case sensitive language.

B) Questions for 2 marks:-

1. Define: Data frame.
2. Define: Summary function from R- Software.
3. Define: fivenum function from R-Software.
4. Define: Resident data sets in R-Software.
5. If $X \rightarrow B(n=5, p=0.4)$, write R command for computing $P(X=3)$ and $P(X<3)$.
6. Create a vector y of numbers between 1 to 200 which are divisible by 5.
7. If $X \rightarrow B(10, 0.5)$ write R-command for calculating $P(X<=3)$.
8. Write output for the following R-commands :

```
>x=c(1,5,2,3)
>y=c(6,7)
>z=x + y
>z
```

9. Write a R-command for creating a vector z having even numbers between 1 and 100.

C) Questions for 4 Marks:-

1. Create a vector ‘WEIGHT’ containing following weight in Kg of 6 students 56, 59, 76, 87, 54, 77.
   i) Access weight of 2^{nd} and 4^{th} student.
   ii) Create a vector WEIGHT60 whose weight is greater than 60.

1. Create a vector x of 2 elements 3 and 7. Create a new vector y from x containing elements 3, 7, 3, 7, 3, 7.

D) Questions for 6 Marks:-

1. Create a data frame containing seat number and marks in two subjects.

2. Create a data frame containing roll number and marks in two subjects.

Unit-2 Concepts and definitions related to Testing of Hypothesis

A) Questions for 1 mark:-

I] Choose the correct alternative:-

1. A statistical hypothesis is
a) statement about the test
b) an imaginary abstract concept
c) an ideal value of parameter
d) a statement about the parameter of distribution

2. A null hypothesis is a
a) hypothesis of interest
b) hypothesis which assigns value 0 to parameter
c) hypothesis of no difference
d) hypothesis which is simple

3. A critical region is a region
a) of rejection for H_0
b) of acceptance for H_0
c) of rejection for either H_0 or H_1
d) both a and b

4. Type I error is
a) accepting H_0 when it is false
b) rejecting H_0 when it is false
c) accepting H_0 when it is true
d) rejecting H_0 when it is true

5. Type II error is
a) accepting H_0 when it is false
b) rejecting H_0 when it is false
c) accepting H_0 when it is true
d) rejecting H_0 when it is true

6. Rejecting H_0 when it is true leads to
a) type I error
b) type II error
c) both type I and type II errors
d) nor type I error and nor type II error

7. Level of significance is
a) proportion of wrong decisions
b) proportion of wrong decision regarding H_0 when it is true
c) proportion of wrong decision regarding H_1
d) proportion of correct decisions

8. Testing $H_0 : \mu = 50$ against $H_1 : \mu \neq 50$ is a
a) one sided left tailed test
b) one sided right tailed test
c) two sided test
d) both a and b

9. Suppose X_1, X_2,------,X_n is a random sample from $N(\mu,\sigma^2)$, σ^2 known. To test $H_0 : \mu=0$, the test statistic is

a) $\dfrac{\overline{X}-\mu}{s/\sqrt{n}}$

b) $\dfrac{\overline{X}}{\sigma/\sqrt{n}}$

c) $\dfrac{\overline{X}}{s/\sqrt{n}}$

d) $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n-1}}$

10. A function of population values is known as :

a) Statistic b) Level of significance
c) Sample d) Parameter

11. Probability of Rejecting H_0 when H_0 is true is called as

a) Statistic b) Parameter
c) Sample d) Level of significance

II] State whether each of the following statement is True or False.

1. Hypothesis is a statement of possible errors about H_0.
2. Critical region is a region of rejection for H_0.
3. Level of significance lies between 0 and 1.
4. In test of significance two types of errors are possible.
5. Test statistic is independent of unknown parameter on parent distribution.
6. The level of significance is probability of rejecting H_1 hypothesis when it is true.
7. Type II error is accepting H_0 when H_1 is true.
8. The standard error of sample mean is $s/\sqrt{n}$.
9. Hypothesis a statement regarding the population of probability distribution.

B) Questions for 2 marks:-

Define the following terms:-

1. Population
2. Random Sample
3. Parameter
4. Statistic
5. Sampling distribution
6. Standard error

7. Types of Errors
8. Null (H_0) and Alternative (H_1) Hypothesis
9. Level of significance
10. Critical region
11. one –tailed test
12. two-tailed test
13. p-value

C) Questions for 4 marks:-

1. Describe the test procedure for testing of Hypothesis.

Unit-3 Large Sample Tests

D) Questions for 4 Marks:-

1. It is claimed that the following is random sample of size 12 from normal distribution with mean 45.
   41, 46, 45, 38, 32, 44, 50, 53, 39, 47, 51, 42
   Examine whether the claim is true at 5% level of significance.
2. In a certain township A, 450 were income tax payers out of a sample of 1000 persons. In another township B, 400 were income tax payers out of a sample of 800 persons. Do these data indicate a significant difference between two townships as far as proportion of income tax payers is concerned? Use 5% l.o.s.
3. A random sample of 1000 school children from rural area shows mean height of 150 cm with standard deviation 45.2 cm. A random sample of 800 school children from urban area shows mean height of 146 cm with standard deviation 37.3 cm. Can we conclude that the children from rural and urban area significantly differ in their mean heights at 5% l.o.s.?
4. A low-noise transistor for use in computing products is being developed. It is claimed that the mean noise level will be below 2.5 dB level of products currently in use. A sample of 16 transistors yields mean noise level 1.8 dB level with standard deviation 0.8 dB level. Test the claim at 5% level of significance.
5. A new computer network is being designed. The maker's claim that it is compatible with more than 99% of the equipment already in use. A sample of 300 programs is run and 298 of these run with no changes necessary that is, they are compatible with the new network. Test the maker's claim at 5% l.o.s.
6. In a sample of 500 parts manufactured by a company, the number of defective parts was found to be 42. The company, however, claimed that 6% of their product is defective. Test the claim at 5% level of significance.
7. Two groups A and B each consisting of 100 people who have a disease. A serum is given to group A but not to group B (which is called the control group),otherwise , the two groups are treated identically. It is found that in groups A and B, 75 and 65 people, respectively, recover from the disease. Test hypothesis that the serum helps to cure disease using a level of significance.
8. A stenographer claims that she can take dictation at the rate of 120 words per minute. Can we reject her claim on the basis of 100 trials in which she demonstrate a mean of 116 words with S.D. of 15 words? Use 5% l.o.s.
9. A wholesaler in apples claims that only 4% of the apples supplied by him are defective. A random sample of 600 apples contained 36 defective apples. Test the claim of the wholesaler.
10. The mean height of 50 male students who showed above average participation in college athletics was 68.2 inches with standard deviation of 2.5 inches, while 50 male students who showed no interest in such participation had a mean height of 67.5 inches with a standard deviation of 2.8 inches. Test the hypothesis that male students who participate in college athletics are smaller than other male students.

11. Daily sales figures of 40 shopkeepers showed that their average sales and standard deviation were 528 rupees and 600 rupees respectively. Can we conclude daily sale on an average is greater than 400. (Use 5% l.o.s.)

12. In a big city 325 men out of 600 were found to smokers. Does this information support the conclusion that the majority of men in this city are smokers?

13. A certain factory runs in two shifts. A sample of 1000 items selected from production of day shift gave 62 defective articles. However a sample of 700 items selected from production of night shift reveals 50 items defectives. Can we conclude that proportion of defective items in the day shift is less than that of night shift. (Use 5% l.o.s.)

E) Questions for 6 Marks:-

1. Describe test procedure for testing equality of single population mean $H_0 : \mu = \mu_0$ against

i. $H_1: \mu < \mu_0$ (left tailed test)
ii. $H_1: \mu > \mu_0$ (right tailed test)
iii. $H_1: \mu \neq \mu_0$ (two tailed test)

2. Describe test procedure for testing equality of two population means i.e. $H_0 : \mu_1 = \mu_2$ against

i. $H_1: \mu_1 < \mu_2$ (left tailed test)
ii. $H_1: \mu_1 > \mu_2$(right tailed test)
iii. H1: $\mu_1 \neq \mu_2$(two tailed test)

3. Describe test procedure for testing equality of single population proportion. i.e. $H_0 : P = P_0$ against

i. $H_1: P < P_0$ (left tailed test)
ii. $H_1: P > P_0$(right tailed test)
iii. H1: $P \neq P_0$(two tailed test)

4. Describe test procedure for testing equality of two population proportions i.e. $H_0 : P_1 = P_2$ against

i. $H_1: P_1 < P_2$ (left tailed test)
ii. $H_1: P_1 > P_2$(right tailed test)
iii. H1: $P_1 \neq P_2$(two tailed test)

Unit-4 Test based on t distribution

A) Questions for 1 mark:-

I] Choose the correct alternative:-

1. Let X_1,X_2,------,X_n be a random sample from a normal population with mean μ and unknown variance σ^2 . Then under H_0 : μ =15 , the statistic $\frac{(\overline{X}-15)\sqrt{12}}{s}$, where s^2 is sample mean square follows:

a) t distribution with 12 d.f. b) t distribution with 15 d.f.

c) t distribution with 14 d.f. d) t distribution with 11 d.f.

2. Let X_1,X_2,------,X_{17} is a random sample from a normal population with mean μ and variance σ^2 (unknown). Suppose sample mean square is denoted by s^2. Then which of the following is the appropriate test statistic used to test H_0 : μ =15 ,

a) $\frac{(\overline{X}-15)\sqrt{16}}{s}$

b) $\frac{(\overline{X}-15)7}{s}$

c) $\frac{(\overline{X}-15)5}{s}$

d) $\frac{(\overline{X}-16)\sqrt{17}}{s}$

3. Let X_1,X_2,------,X_9 be a random sample from a normal population with mean μ and unknown variance σ^2. We want to test H_0 : μ =16 against H_0 : μ ≠16 at level of significance α .In this case, the critical region is given by calculated value of the appropriate test specific is

a) $< t_{8,\alpha/}2$ b) $\geq t_{8,\alpha/2}$

c) $\geq t_{9,\alpha/2}$ d) $< t_{9,\alpha/2}$

4. Paired t- test was applied using 13 pairs of observations (X_i,Y_i);i=1,2,---,13. In this case the distribution of test statistic under the null hypothesis H_0: $\mu_d = 0$ is

a) t distribution with 13 d.f. b) t distribution with 12 d.f.

c) t distribution with 26 d.f. d) t distribution with 24 d.f.

II] State whether each of the following statement is true or false.

1. A t-test cannot be applied only if observations in the population from which random sample is drawn follow normal distribution.
2. In a paired t test observations in two samples are independent of each other.
3. A random sample $X_1, X_2, \text{------}, X_{25}$ drawn from normal population with unknown parameters μ and σ^2 has mean square 625. Then under H_0: $\mu = 50$ the statistic $\frac{\overline{X} - 50}{5}$ follows t distribution with 24 df.
4. In a test based on t distribution, the value of the test statistic cannot be negative.

B) Questions for 2 marks:-

1. Define t test
2. Define paired t- test

C) Questions for 4 marks:-

1. Describe test procedure for paired t- test.
2. Describe test procedure for testing significance of correlation coefficient.
3. Describe test procedure for testing significance of regression coefficient of Y on X.
4. It is claimed that the sales of a product (in lakh Rs.)(Y) is a linear function of advertisement cost (in thousand Rs.) (X). Consider the following data :

X	41	67	92	38	80	84
Y	46	52	85	50	83	60

   Test the above claim using the test of significance of regression coefficient of Y on X at 5% l.o.s.

5. Let X denotes the number of lines of executable SAS code, and let Y denote the execution time in seconds. Use the following summary information:

   n=10, $\sum_{i=1}^{10} x_i = 16.75$, $\sum_{j=1}^{10} y_j = 170$, $\sum_{i=1}^{10} x_i^2 = 28.64$, $\sum_{j=1}^{10} y_j^2 = 2898$, $\sum_{i=1}^{10}\sum_{j=1}^{10} x_i y_j = 285.625$

   i. Compute the value of regression coefficient of Y and X and
   ii. Test the significance of regression coefficient of Y on X at 1% l.o.s.

6. Let X denotes marks of students in Statistics, and let Y denote marks of students in Mathematics. Use the following summary information:

$$n=10\ ,\ \sum_{i=1}^{10} x_i = 55,\quad \sum_{j=1}^{10} y_j = 40,\ \sum_{i=1}^{10} x_i^2 = 385,\ \sum_{j=1}^{10} y_j^2 = 192,\ \sum_{i=1}^{10}\sum_{j=1}^{10} x_i y_j = 185$$

i. Compute the value of regression coefficient of Y and X and
ii. Test the significance of regression coefficient of Y on X at 1% l.o.s.

7. Memory capacity of 10 students was tested before and after training are as follows:

Roll No.	1	2	3	4	5	6	7	8	9	10
Before training	12	14	11	8	7	10	3	0	5	6
After training	15	16	10	7	5	12	10	2	3	8

Test whether the training was effective or not. Use 5% l.o.s.

8. The table below gives the number of customers visiting a certain Post office on various days of the week :

Days	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Number of Customers	130	120	110	115	110	135

Test whether the customers visiting the post office are uniformly distributed. Use 5% l.o.s.

9. A machine designed to produce insulating washers for electrical devices of average thickness of 0.025 cm. A random sample of 10 washers was found to have an average thickness of 0.074 cm with S.D. of 0.002 cm. Test the significance of deviation at 5% l.o.s.

10. A random sample of 10 objects has the following weight in kg. 70, 120, 110, 101, 88, 83, 95, 88, 107, 100.Does this data support that the entire population from which above sample is drawn , has mean weight of 100 kg.

11. A random sample of 8 envelopes is taken from letter box of a post office and their weights in grams are found to be 12.1 ,11.9 , 12.4 , 12.3 , 11.9 , 12.5 , 12.8 , 12.1

Does this sample indicate at 1% level that the average weight of envelopes received at that post office is greater than 12.35 grams?

12. Two horses A and B were tested according to time (in seconds) to run a particular track with the following results.

Horse A	28	30	32	33	33	29	34
Horse B	29	30	30	24	27	29	

Test whether the two horses have the same running capacity.

13. A random sample of size n_1 =15 taken from a normal population has standard deviation 5.2 and mean $\bar{x}_1$ =81. A second random sample of size n_2 = 16 taken from normal population has mean $\bar{x}_2$ = 76 and standard deviation 3.4. Test the hypothesis that $\mu_1 = \mu_2$ against the alternative $\mu_1 > \mu_2$.

14. Eleven school boys were given a test in statistics. They were given a month's tution and second test was held at the end of it. Do the marks give evidence that the students have benefited by the extra coaching?

Boys	1	2	3	4	5	6	7	8	9	10	11
Marks of 1st test	23	20	19	21	20	18	17	23	16	18	19
Marks of 2nd test	24	19	22	18	22	20	20	23	20	20	18

D) Questions for 6 marks:-

1. Describe t- test for testing H_0: $\mu = \mu_0$ against

i. H_1: $\mu \neq \mu_0$

ii. H_1: $\mu < \mu_0$

iii. H_1: $\mu > \mu_0$

2. Describe t-test for testing H_0: $\mu_1 = \mu_2$ against

i. H_1: $\mu_1 \neq \mu_2$

ii. H_1: $\mu_1 \neq \mu_2$

iii. H_1: $\mu_1 \neq \mu_2$

Unit-5 Test based on Chi-Square distribution

A) Questions for 1 mark:-

I] Choose the correct alternative:-

1. Suppose $o_1, o_2, \text{------}, o_i, \text{----}, o_k$ is a set of observed frequencies and $e_1, e_2, \text{------}, e_i, \text{----}, e_k$ are corresponding expected frequencies and $\sum_{i=1}^{k} o_i = N$. Then in order to test

   H_0: There is no significant difference between observed and expected frequencies

   the test statistic is:

   a) $\sum_{i=1}^{k} \left(\frac{o_i^{\,2}}{e_i} \right) - N$

   b) $\sum_{i=1}^{k} \left(\frac{e_i}{o_i^{\,2}} \right) - N$

   c) $\sum_{i=1}^{k} \left(\frac{e_i}{o_i} \right) - N$

   d) $\sum_{i=1}^{k} \left(\frac{e_i^{\,2}}{o_i} \right) - N$

2. Suppose $e_1, e_2, \text{------}, e_i, \text{----}, e_{10}$ ($e_i \geq 5$ for all i) is a set of expected frequencies obtained fitting a probability distribution in which 2 parameters were estimated. Then under H_0 : Fitting of the probability distribution is good, the test statistic used follows:

   a) χ^2 distribution with 5 d.f. b) χ^2 distribution with 9 d.f.

   c) χ^2 distribution with 7 d.f. d) χ^2 distribution with 10 d.f.

3. A 4*3 contingency table was obtained to test H_0 : two attributes A and B are independent, then under H_0 , the distribution of statistic used in this case is

   a) χ^2 with 6 d.f. b) χ^2 with 12 d.f.

   c) χ^2 with 7 d.f. d) χ^2 with 11 d.f.

4. We want to test H_0 : Two attributes A and B are independent and both the attributes are at two levels. Then under H_0 , the statistic used is

   a) $\chi_2^{\,2}$ b) $\chi_4^{\,2}$

   c) $\chi_3^{\,2}$ d) $\chi_1^{\,2}$

II] State whether each of the following statement is True or False.

1. In a test based on χ^2 distribution, we need not pool frequencies of neighboring classes while computing the value of test statistic if any expected frequency is less than 5.
2. The critical region in χ^2 test of goodness of fit is always one sided.
3. Suppose A and B are two attributes each at 2 levels. We want to test H_0: A and B are independent. In this case we use test statistic which follow χ^2 distribution with 1 d.f. under H_0.

B) Questions for 2 marks:-

1. Define a chi-square variate.
2. State the conditions for validity of chi-square test for goodness of fit.
3. For what purpose goodness of fit is applied.

C) Questions for 4 marks:-

1. Describe the procedure for chi-square test for goodness of fit.
2. Describe the procedure for chi-square test for independence of Attributes.
3. In a radio listener's survey, 120 persons were interviewed and their opinions about preference to Hindi or English music and preference to classical or light music were asked. The results are as follows :

	English Music	Hindi Music
Classical Music	13	45
Light Music	39	23

   Examine at 5% l.o.s., whether the preference to music language is independent of type of music.

4. In a radio listener's survey 120 persons were interviewed and their opinions about preference to Hindi or Marathi music were asked. The results are as follows :

Type of Music	Opinion about Preference	
	Hindi	Marathi
I	13	45
II	39	23

Test whether the preference for music type is independent of language at 5% level of significance.

5. A random sample of 90 adults is classified according to gender and the number of hours they watch television during a week :

Time spent watching television during a week	Gender	
	Male	Female
Over 25 hours	15	29
Under 25 hours	27	19

Test whether the time spent in watching television is independent of gender at 5% level of significance.

6. The theory predicts that the proportion of beans in the four groups A, B, C and D should be 9:3:3:1. In an experiment among 1600 beans, the numbers in the four groups were 882,313,287 and 118. Does the experimental result support the theory?

7. A die is thrown 60 times with the following results:

Face	1	2	3	4	5	6
frequency	8	7	12	8	14	11

Test at 5% level of significance, if the die is honest.

8. Children having one parent of blood type M and the other type N will always be one of the three types M, MN, N and average proportions of these will be 1:2:1. Out of 300 children having one M parent and one N parent, 30% were found to be of type M, 45% of type MN and the remaining of type N. Use χ^2 test to test the hypothesis.

9. 1072 college students were classified according to their intelligence and economic conditions. Test whether there is any association between intelligence and economic conditions.

		Excellent	Good	Mediocre	Dull
Economic conditions	Good	48	199	181	82
	Not good	81	185	190	106

10. The results of a survey regarding radio listener's preference for different types of music are given in the following table, with listeners classified by age group. Is preference of types of music influence by age?

Age Group

Type of music preferred	19-25	26-35	Above 36
National Music	80	60	9
Foreign Music	210	325	44
Indifferent	16	45	132

11. In an experiment of immunization of cattle from tuberculosis, the following results were obtained

	Affected	Unaffected
Inoculated	12	28
Non-Inoculated	13	7

Examine whether the two underlying attributes are independent or not. Use 1% l.o.s.

12. On a particular proposal of national importance, the two parties, A and B, cast votes as shown below. At 1 % l.o.s. test hypothesis that there is no difference between the two parties regarding the proposal.

	In favour	Opposed	Undecided
Party A	85	78	37
Party B	118	61	25

D) Questions for 6 marks:-

1. Describe the procedure for chi-square test for goodness of fit.

Suppose that a die is rolled 150 times and the number of times each face comes up is recorded and results are obtained as:

Face	1	2	3	4	5	6
Observed Frequency	29	19	19	27	26	30

Are these results consistent with the hypothesis that the die is fair at 1 % level of significance?

2. Describe the procedure for chi-square test for independence of Attributes.

Two groups of certain types of patients, A and B, each containing of 200 people are used to test the effectiveness of a new serum. Both groups are treated identically except that Group A is given serum while group B is not. It was found that 140 and 120 of groups A and B respectively recover from the disease. Is the observed result sufficient evidence for the conclusion that the new serum helps to cure the disease, if we are willing to take a risk of 0.01?

Unit-6 Simulation

A) Questions for 1 mark:-

I] Choose the correct alternative:-

1. If U_1 and U_2 are two independent U(0,1) random variable then $X \rightarrow N(0,1)$ distribution if X=
   a) $[-2\log_e(U_2)^{1/2}]\cos(2\Pi U_1)$
   b) $[2\log(U_1)^{1/2}]\cos(2\Pi U_2)$
   c) $[-2\log_e(U_1)^{1/2}]\cos(\Pi U_2)$
   d) $[-2\log(U_1)^{1/2}]\cos(2\Pi U_2)$
2. In order to draw a random sample from U(a,b) where a and b are constants first obtain random sample from U(0,1) and then use formula
   a) a + (b – a) * rand( ) b) b - a * rand( )
   c) = a + (b – a) * rand( ) d) = b - a * rand( )
3. In a sign test suppose s^+ = number of the +ve signs , s^- = number of –ve signs and s=min (s^+,s^-) then we accept the hypothesis of randomness H_0 at l.o.s. α if
   a) $s \leq s_\alpha$ b) $s < s_\alpha$
   c) $s \geq s_\alpha$ d) $s > s_\alpha$
4. In run test used to test randomness of sequence of random numbers suppose R = total number of runs, a and b are critical values. Then hypothesis H_0 of randomness is accepted at l.o.s. α if
   a) $a < U < b$ b) $a < U \leq b$
   c) $a \leq U < b$ d) $a \leq U \leq b$

II] State whether each of the following statement is True or False.

1. Simulation is a method of imitating the real system with artificial data using computer.
2. In simulation random sample is drawn just by specifying the average without using a statistical model.
3. Pseudo-Random numbers do not satisfy all the tests of randomness.
4. Box-Muller transformation is used to draw a random sample from uniform distribution.
5. In run test we take alternative hypothesis H_1: The given sequence of numbers is not random.
6. In sign test the null hypothesis H_0: The population mean = 0.5.

B) Questions for 2 marks:-

1. What is simulation, Describe in brief?
2. State any two merits of simulation.
3. State any two demerits of simulation.
4. Define linear congruential generator.
5. Define Box-Muller transformation.
6. What do you mean by pseudo-random number?

C) Questions for 4 marks:-

1. What are requisites of Good Random Number Generator?
2. Explain the use of computers in simulation.
3. Describe the procedure of obtaining a random sample from uniform (a ,b) distribution.
4. Describe the procedure of obtaining a random sample from exponential distribution with mean θ, using simulation technique.
5. Explain how Box-Muller transformation is used to obtain a model sample from $N(\mu, \sigma^2)$ distribution.
6. Describe test procedure of goodness of fit test to check randomness of a given sample.
7. A driver buys petrol either at a petrol pump P or at petrol pump S and the following arrangement shows the order of the petrol pump from which he bought petrol over a certain period of time :

   P P P S P S P S S P P S P S P S P S S P S S

   Test the randomness of the above sequence at 5% level of significance (l.o.s.).

8. In a laboratory experiment, 18 determinations of the coefficient of friction between leather and metal yielded the following results :

   0.49, 0.56, 0.49, 0.55, 0.45, 0.55, 0.51, 0.4, 0.56, 0.47, 0.58, 0.41, 0.54, 0.48, 0.51, 0.57, 0.43, 0.56.

   Test using sign whether population median is 0.5 at 5% level of significance (l.o.s.).

D) Questions for 6 marks:-

1. Describe the procedure of sign test for testing symmetry of the sample.

   Following is the sequence of random numbers generated using MS-EXCEL RAND ( ) function. Test the randomness of the sequence.

   0.54, 0.77, 0.15, 0.51, 0.37, 0.60, 0.95, 0.81, 0.30, 0.66, 0.00, 0.86, 0.14, 0.08, 0.86

2. Describe the procedure of run test for testing randomness.

   A sequence of 15 random numbers generated using MS-EXCEL RAND ( ) function is given below. Examine whether the sequence has population median $\mu = 0.5$.

   0.96, 0.51, 0.41, 0.09, 0.50, 0.12, 0.12, 0.55, 0.52, 0.84, 0.77, 0.78, 0.28, 0.69, 0.38

3. Explain how Box-Muller transformation is used to obtain a model sample from $N(\mu, \sigma^2)$ distribution.

   Suppose the life length of a water heater produced by a certain company has an exponential distribution with mean 12 years. Simulate life lengths of 20 such water heaters and estimate Standard deviation of the distribution.

4. Describe the procedure of obtaining a random sample from uniform (a ,b) distribution.

   Buses ply at half-hourly intervals, starting at 8.00 a.m. On a given day, a commuter arrives at a time X in the morning which is uniformly distributed between 8.15 am and 8.45 am. Simulate the arrivals of the commuter for 15 days.

****