

**Q.P. Code : 05377**

**( 2½ Hours)**

**[ Total Marks : 75**

- N.B. :** (1) All the question are **compulsory**. Choice is **internal**.  
 (2) **Figures** to the **right** indicate **Full Marks**.  
 (3) **All** questions carry **equal marks**.  
 (4) Use of **statistical tables & simple calculators is permitted**.

1. (A) Choose the **most appropriate** answer (**any three**) : **3**

- (i) The arithmetic mean is 10 for 5 items and is 10 for 18 items. The combined mean is \_\_\_\_\_.  
 (a) 10 (b) 8 (c) 18
- (ii) \_\_\_\_\_ is affected by extreme values.  
 (a) Mode (b) Mean (c) Median
- (iii) The value of the variable which divides an array of series when arranged in ascending order into \_\_\_\_\_ equal parts are called percentiles.  
 (a) 100 (b) 99 (c) 4
- (iv) \_\_\_\_\_ is the simplest measure of dispersion.  
 (a) Mean deviation (b) Standard deviation (c) Range
- (v) The Sample mean  $\bar{X}$  is a \_\_\_\_\_.  
 (a) Parameter (b) Statistic (c) Constant
- (vi) The mode in 'ALABAMAMONTANA' is \_\_\_\_\_.  
 (a) M (b) A (c) N

(B) Define and explain **any one** of the following : **2**

- (i) Histogram (ii) Mean

(C) Answer **any one** of the following : **4**

- (i) A bookshop in Delhi had sold solve books to students of different ages by a particular publisher on a certain day with the following distribution :

Age Group	1 - 3	4 - 6	7 - 9	10 - 12	13 - 15	16 - 18	19 - 21
No. of books sold	10	15	24	35	16	3	1

Find the mean of the distribution.

- (ii) Three dices were thrown together. The following table shows the frequency of the sum of their numbers obtained. Calculate the standard deviation :

Sum of the numbers	0-4	4 - 8	8 - 12	12 - 16	16 - 20
Frequency	3	10	16	12	9

**TURN OVER**

(D) Answer **any one** of the following :

(i) Measures of dispersion.

(ii) Calculate the mean and median for the following data :

Variable	Frequency
1 - 10	80
11 - 20	77
21 - 30	72
31 - 40	65
41 - 50	55
51 - 60	43
61 - 70	28
71 - 80	26
81 - 90	10
91 - 100	8

2. (A) Choose the **most appropriate** answer (**any three**) :

(i) In positive skewness of a distribution, \_\_\_\_\_ is observed.

(a) Mean > Median > Mode

(b) Mean < Median > Mode

(c) Mean < Median < Mode

(ii) EMBL is an example of \_\_\_\_\_ database.

(a) Primary

(b) Secondary

(c) Tertiary

(iii) \_\_\_\_\_ event is the total number of all the possible outcomes of an experiment.

(a) Dependent

(b) Independent

(c) Exhaustive

(iv) A micro array is a glass slide onto which \_\_\_\_\_ molecules are fixed in an orderly manner.

(a) RNA

(b) DNA

(c) Protein

(v) PDB contains various protein structure NOT obtained by \_\_\_\_\_.

(a) NMR

(b) X-ray crystallography

(c) autoradiography

(vi) The probability that all 3 coins when tossed together, will show heads is \_\_\_\_\_.

(a) 1/3

(b) 1/8

(c) 3/6

TURN OVER

(B) Define and Explain **any one** of the following :

- (i) Probability (ii) GenBank

2

(C) Attempt **any one** :

- (i) Application of bioinformatics.  
(ii) Two students X and Y work independently on a problem. The probability that X will solve it is  $\frac{3}{4}$  and the probability that Y will solve it is  $\frac{2}{3}$ . What is the probability that the problem will be solved?

4

(D) Elaborate on **any one** :

- (i) Microarray analysis.  
(ii) (a) Out of 4 boys and 5 girls, a team of 4 students is to be selected for a quiz programme. Find the probability that 2 are girls and 2 are boys.  
(b) The probability that a boy will get a scholarship is 0.9 and that a girl will get is 0.8. What is the probability that at least one of them will get the scholarship?

6

3. (A) Choose the **most appropriate** answer (**any three**) :

3

- (i) The \_\_\_\_\_ hypothesis is regarded to be true in the zone of acceptance of the normal curve.  
(a) Alternative (b) Null (c) Statistical
- (ii) The probability of rejecting the null hypothesis when it is true is called \_\_\_\_\_.  
(a) level of confidence  
(b) level of significance  
(c) power of the test
- (iii) The alternative hypothesis for a left-tailed test is \_\_\_\_\_.  
(a)  $\mu \neq 10$  (b)  $\mu > 10$  (c)  $\mu < 10$
- (iv) The probability associated with committing type I error is \_\_\_\_\_.  
(a)  $\alpha$  (b)  $\beta$  (c)  $1 - \beta$
- (v) For a two tailed test, the area of rejection is at the \_\_\_\_\_ of a normal distribution curve tail.  
(a) centre (b) one end (c) both ends
- (vi) For the application of Z-test, the sample size must be \_\_\_\_\_.  
(a)  $> 30$  (b)  $< 40$  (c)  $> 50$

TURN OVER

4

(B) Define and explain **any one** of the following :

- (i) Alternative hypothesis (ii) Zone of rejection

2

(C) Attempt **any one** of the following :

- (i) Write a note on Type I and II errors.  
(ii) In an intelligence test administered to 52 students, the average score was 89 and standard deviation 18. Test the hypothesis that average score of the population is less than 78.

4

(D) Attempt **any one** of the following :

- (i) In a group of 196 adults in the age group 45-53 years belonging to social class I, the mean serum cholesterol was 180 mg% with a standard deviation of 42 mg%. In a comparable group of 144 adults belonging to social class V, the mean serum cholesterol was 150 mg% with a standard deviation of 48 mg%. Is the difference in cholesterol level of the two classes statistically significant?  
(ii) Determine the significance of standard deviation with gender.

6

Gender	Number	Mean height (cm)	SD
Boys	69	168	14
Girls	54	153	8

4. (A) Choose the **most appropriate** answer (**any three**) :

3

(i) The degree of freedom for unpaired t - test can be calculated as \_\_\_\_\_.

- (a)  $n_1 + n_2$  (b)  $n_1 + n_2 - 1$  (c)  $n_1 + n_2 - 2$

(ii) The values of  $\chi^2$  distribution can be \_\_\_\_\_.

- (a) -10 (b) +50 (c) -100

(iii) The mean difference between 16 paired observations is 25 and the standard deviation of differences is 10. The value of t-stastic is \_\_\_\_\_.

- (a) 10 (b) 16 (c) 25

(iv) If  $\chi^2$  value is \_\_\_\_\_ than the  $\chi^2$  table value, the hypothesis is accepted.

- (a) Less (b) Equal or greater (c) Greater

(v) \_\_\_\_\_ was given by W. S. Goresett.

- (a) z-test (b)  $\chi^2$  test (c) t-test

(vi) \_\_\_\_\_ is applied for a sample with correlated observations.

- (a) Paired t-test (b) Unpaired t-test (c)  $\chi^2$  test

TURN OVER

(B) Answer **any one** of the following :

- (i) Paired t test (ii) Test of association

2

(C) Attempt **any one** of the following :

- (i) Explain  $\chi^2$  test as Test of population variance.  
 (ii) Genetic theory states that children having one parent of Blood group A and other of B will always be one of the three types : A, AB and B & that proportion of the three types will be on an average 2: 3: 1. A report states that out of 300 children having one 'A' parent and one 'B' parent, 30% were found to be type A, 40% type AB, remainder type B. Test the hypothesis.

$(t_{0.05, 2}) = 1.81; (\chi^2_{0.05, 1}) = 3.84; (t_{0.05, 1}) = 1.51; (\chi^2_{0.05, 2}) = 5.44.$

4

(D) Discuss in detail **any two** of the following :

- (i) Serum digoxin levels were determined for nine healthy males aged 20 - 45 years following rapid intravenous injection of the drug. The measurements were made 4 hours after the injection and again at the end of an 8 hour period.

6

**Serum digoxin concentration  $\mu\text{g} / \text{ml}$**

4 hours	1.0	1.3	0.9	1.0	1.0	0.9	1.3	1.1	1.0
8 hours	1.0	1.3	0.7	1.0	0.9	0.8	1.2	1.0	1.0

Is the difference in the serum digoxin concentration at the end of 4 hours and at the end of 8 hours statistically significant?

$(t_{0.05, 9}) = 2.26; (t_{0.05, 8}) = 2.31; (t_{0.05, 16}) = 2.12; (t^2_{0.05, 18}) = 2.10.$

- (ii) An oil company has explored three different areas for possible oil reserves. The results of the test were as given below :

	Area		
	A	B	C
<b>Strikes</b>	7	10	8
<b>Dry holes</b>	10	18	9
<b>Total</b>	<b>17</b>	<b>28</b>	<b>17</b>

Do the three areas have the same potential, at the 5% level of significance?

$(\chi^2_{0.05, 2}) = 5.991; (\chi^2_{0.05, 3}) = 7.82.$

TURN OVER

5. (A) Answer **any one** of the following :

3

(i) Calculate the mean number of computers per family from the following table :

No. of Computers	1 - 2	2 - 3	3 - 4	4 - 5	5 - 6	6 - 7
No. of Families	42	49	34	10	2	2

(ii) Write a note on applications of biostatistics in medicine.

(B) Attempt **any one** of the following :

3

(i) The face cards are removed from a full pack of 52 playing cards. Out of the remaining 40 cards, 4 are drawn at random. What is the probability that they belong to different suits?

(ii) Elaborate on Normal distribution.

(C) Attempt **any one** of the following :

3

(i) Level of significance.

(ii) The mean weekly sale of the biscuit was 140 per store. After an advertisement campaign, the mean weekly sales in 45 stores for a typical week increased to 153 and showed a standard deviation of 17. Was the advertising campaign successful.

(D) Attempt **any one** of the following :

3

(i) From the table given below, whether the colour of son's eyes is associated with that of father's eyes.

		Eyes colour in sons	
		Not light	Light
Eye Colour in father	Not light	23	27
	Light	17	33

$$(\chi^2_{0.05, 1}) = 3.84; (\chi^2_{0.05, 2}) = 5.99.$$

(ii) Write a note on characteristics of t-test.

(E) State True or False (**any three**) :

3

(i) In  $\chi^2$  test, the smaller the number of degrees of freedom, the more skewed is the distribution.

(ii) Mode is also known as norm.

(iii) The probability of Type I error is usually determined in advance.

(iv) Normal distribution is symmetrical.

(v) Databases are fundamental tools for biological research.

(vi) FASTA is a more sensitive tool than BLAST for nucleic acid sequence analysis .