3 Hours **TOTAL MARKS: 100**

Instructions: Please check that you have received the correct question paper.

1) All the questions are compulsory. Choice is internal.
2) Figures to the right indicate full marks.
3) All questions carry equal marks.
4) Use of simple calculators is permissible.

Q1 A) Fill in the blanks (<u>any three</u>):

i) _____ is not an ideal measure of central tendency if data has extreme values. **3**

ii) _____ is a device of representing statistical data as pictures.

iii) If a person A is shorter than person B, then it is said to be _____ scale of measurement.

iv) The _____ quartile has 25% of the items of the distribution above it.

v) _____ is based on two extreme observations.

Q1 B) Answer in brief <u>any one</u>: **3**

i) Write a comparative account of histogram and frequency polygon.

ii) A firm of readymade garments make both men's and women's shirts. Its profit average is 6% of sales. Its profits in men's shirts average is 8% of sales and women's shirts comprise 60% of output. What is the average profit per sales rupee in women's shirts?

Q1 C) Answer <u>any one</u>: **6**

i) Find the value of mean and mode of the following data:

Weight (kg)	30-34	35-39	40-44	45-49	50-54	55-59	60-64
No. of students	3	5	12	18	14	6	2

ii) Write a note on standard deviation.

Q1 D) Answer <u>any one</u>: **8**

i) Calculate the median, P_{70} and mean deviation of the following data of marks obtained in biostatistics paper.

Marks	0-7	7-14	14-21	21-28	28-35	35-42	42-49
No. of students	3	4	7	11	2	14	9

ii) Calculate the mean and standard deviation from the following data of presence of AST in the blood samples of 520 patients at GIS hospital.

Range (mg/dl)	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of patients	20	44	60	101	109	84	66	10

Q2 A) Fill in the blanks (<u>any three</u>): **3**

i) The cases which ensure the occurrence of an event are said to be _____ to the event.

ii) The probability of an event which is certain to occur is _____

iii) The normal distribution is a _____ distribution.

iv) _____ is an example of protein database.
v) _____is a multiple sequence analysis tool.

Q2 B) Answer in brief <u>any one</u>: **3**

i) A problem in statistics is given to five students., A, B, C, D, E. Their chances of solving it are 1/2,1/3,1/4,1/5 and 1/6. What is the probability that the problem will be solved?

ii) Write a brief note on GenBank.

Q2 C) Answer <u>any one</u>: **6**

i) Give detailed account of microarray analysis supporting your answer with a schematic representation. Also add a note on its applications

ii) Assuming the mean height of the children to be 68.22 cm with a variance of 10.8 cm. how many children in a school of 1000 are expected to be over 72 cm tall? (For Z=1, area=0.3413, for Z= 1.1,area= 0.3643, for Z=1.15, area= 0.3749)

Q2 D) Answer <u>any one</u>: **8**

i) Give detailed account of the applications of bioinformatics in different fields.

ii) Two cards are selected at random one after the other from a well shuffled deck of 52 cards. 1) If the first card is replaced before the second is drawn, find probabilities that: (a) both are spades (b) one is ace and the other is jack (c) an honour card in the first draw and a 9 in the second draw.
2) If the first card is not replaced before the second, find the above probabilities.

Q3 A) Fill in the blanks <u>(any three)</u>: **3**

i) Any statistical measure computed from sample data is known as_____.
ii) The_____ separates the rejection region from the acceptance region.
iii) Z statistic for a two tailed test at 5% level of significance is _____.
iv) _____ test is used when sample size is more than 300.
v) The probability of making a type II error is denoted by _____.

Q3 B) Differentiate between <u>any one</u>: **3**

i) Type I and type II error
ii) Statistic and parameter

Q3 C) Answer <u>any one</u>: **6**

i) Random samples drawn from two places gave the following data relating to the heights of children:

	Place A	**Place B**
Mean height (cm)	68.50	68.58
Standard deviation (cm)	2.50	3.00
Number of Children	1200	1500
Test at 5% level that the mean height is the same for children at two places.		

ii) Write a note on different types of tailed tests used for testing hypothesis.

Q3 D) Answer <u>any one</u>: **8**

i) The standard deviation of the height of honours students is 5.0cm. Two samples are taken. The mean height of 1000 B. Com students is 158 cm with standard deviation of 5.5 cm and mean height of 580 B.A. students is 155 cm with standard deviation of 6.5 cm. Test the significance of standard deviation of the sample.

ii) Justify: Hypothesis testing is a multistep process.

Q4 A) Fill in the blanks <u>(any three)</u>: **3**

i) The theory of small sample was developed by _____.

ii) If the alternate hypothesis is $\mu > 10$, for a null hypothesis, $\mu = 10$, then it is a_____ tailed test.

iii) Chi square test was first used in testing statistical hypothesis by _____.

iv) Increase in weight of two groups of 10 and 14 students after providing nourishing food is to be tested. The degrees of freedom in this case would be _____.

v) The chi square value lies between 0 and _____.

Q4 B) Answer in brief <u>any one</u>: **3**

i) Write characteristics of chi square test.

ii) A soap manufacturing company was distributing a particular brand soap through a large number of retail shops. Before a heavy advertisement campaign, the mean sales per shop was 140 dozen. After the campaign, a sample of 26 shops was taken and the mean sales figure was found to be 147 dozen with standard deviation 16. Can you consider the campaign effective? (Given $t_{0.05, 25} = 2.06$, $t_{0.05, 26} = 2.056$)

Q4 C) Answer <u>any one</u>: **6**

i) Write an elaborative note on t-test.

ii) Ten rats were fed with rice in the first month and body weights of the rats were recorded. In the next month they were fed with grams and their body weights were measured again. The respective weights of ten rates in two months are as follows:

Weight in 1st month (g)	50	60	58	52	51	62	58	55	50	65
Weight in 2nd month (g)	56	58	68	61	56	59	64	60	50	62

Test the given data to find the impact of grams in rat's nutrition. (Given $t_{0.05, 9} = 2.26$, $t_{0.05, 10} = 2.228$ and $t_{0.05, 20} = 2.086$)

Q4 D) Answer <u>any one</u>: **8**

i) Mentioning the criteria for applying unpaired t- test, least the properties of the t-distribution. Apply a suitable test for the following data. In a nutritional study, 13 children were given a usual diet plus vitamins A and D tablets. While the second group of 12 children was taking the usual diet. After 12 months, the gain in weight in pounds was noted as given in the table. Can you say that vitamin A and D were responsible for this difference? ($t_{0.05, 23} = 2.07$, $t_{0.05, 24} = 2.064$, $t_{0.05, 25} = 2.060$, $t_{0.05,11} = 2.201$)

A=	5	3	4	3	2	6	3	2	3	6	7	5
B=	1	3	2	4	2	1	3	4	3	2	2	3

ii) In garden pea plant heterozygote for mutant for shortness was self-crossed. Results obtained from two different localities give the following results: Locality -1, tall-218, short 70, Locality -2, tall 346, short 116. Locality 3-, tall 567, short 175Using appropriate statistical method determine whether any environmental factors responsible for such deviation. ($x^2_{0.05,1}$= 3.84, $x^2_{0.05,2}$= 5.991)

Q5 A) Answer <u>any four</u>: **16**

i) Differentiate between frequency polygon and pie chart.

OR

i) Define arithmetic mean and give its merits and demerits.

ii) Discuss the characteristics of a normal curve.

OR

ii) Two cards are selected at random one after the other from a well shuffled deck of 52 cards. If the first card is replaced before the second is drawn, find the probabilities that – (a) both are spades (b) one is an ace and the other is jack.

iii) Give the characteristics of Z test.

OR

iii) A simple sample of 1000 members is found to have a mean 3.42 cm., Could it be reasonably regarded as a simple sample from a large population arose men is 3.30 cm and standard deviation is 2.6 cm?

iv) In a cross of red and white flowers, the F2 individual segregated into 787 red and 277 white flowers. Do these results agree with Mendel's laws?

OR

iv) A fertilizer mixing machine is set to give 12 kg of nitrate for every quintal bag of fertilizer. Ten 100 kg bags are examined. The percentage of nitrate is as follows: 10, 14, 13, 12, 13 ,12, 13, 14, 11, 12. Is there reason to believe that the machine is defective? ($t_{0.05,\ 9}$= 2.262, $t_{0.05,\ 10}$= 2.228)

Q5 B) State true or false (<u>any four</u>): **4**

i) Median and mean are both affected by extreme values.
ii) Dendrogram is a 3-D graph.
iii) For negatively skewed curve, mean<median<mode.
iv) Standard deviation is a measure of central tendency.
v) Probability can never be a negative value.
vi) Mode for ATTITUDE is T.
