

**POST GRADUATE DIPLOMA IN
APPLIED STATISTICS (PGDAST)**

Term-End Examination

December, 2016

01014

MST-004 : STATISTICAL INFERENCE

Time : 3 hours

Maximum Marks : 50

Note :

- (i) *Attempt all questions. Questions no. 2 to 5 have internal choices.*
- (ii) *Use of scientific calculator is allowed.*
- (iii) *Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.*
- (iv) *Symbols have their usual meaning.*

-
1. State whether the following statements are *True* or *False*. Give reasons in support of your answers. $5 \times 2 = 10$
- (a) A function of sample values to estimate a parameter is called sampling unit.
 - (b) An estimator (T) is said to be unbiased for parameter θ , if $\text{Var}(T) = \theta$.
 - (c) Rejecting null hypothesis (H_0) when it is true is called type-I error.

- (d) The number of runs in the sequence HTHHTTTHHHTTT is 6.
- (e) The estimators obtained by the method of moments are more efficient than the maximum likelihood estimators.

2. (a) The salaries (in thousands) of 5 workers in a factory are 20, 15, 14, 10 and 25.

(i) How many samples of size 2 are possible without replacement ? Write them. 3

(ii) How many samples of size 2 are possible with replacement ? 1

(b) A machine produces a large number of items of which 15% are found to be defective. If a random sample of 200 items is taken from the population and the sample proportion is calculated, then find

(i) the mean and standard error of the sampling distribution of the proportion defective. 2

(ii) the probability that less than or equal to 16% defectives are found in the sample. 4

OR

- (a) The probability density function (pdf) of a t-distribution is given by

$$f(t) = \frac{1}{\sqrt{5} \beta \left(\frac{1}{2}, \frac{5}{2}\right) \left(1 + \frac{t^2}{5}\right)^3}; \quad -\infty < t < \infty.$$

- (i) What is the degree of freedom of the t-distribution? 2
- (ii) What are the mean and variance of the t-distribution? 3
- (iii) Write any two applications of the t-distribution. 1

- (b) Differentiate between the following with examples :

- (i) Population and Sample 2
- (ii) Estimator and Estimate 2

3. To estimate the average nicotine content (μ) in a newly launched cigarette, an analyst randomly selects 44 of these cigarettes and determines their nicotine content. It is known from past experience that the standard deviation of the nicotine content of a cigarette is equal to 0.7 milligrams.

Answer the following questions :

- (i) What is a 95 percent confidence interval estimate of the average nicotine content in the cigarettes, if the average nicotine content finding in the sample is 1.75 milligrams? 5

- (ii) How large a sample is necessary for the length of the 95 percent confidence interval so that the estimate of the average nicotine content lies between ± 0.3 milligrams ?

5

OR

- (a) Show that the sample mean (\bar{X}) is an unbiased estimator of the population mean (μ). A survey was undertaken to learn about the weekly average number of working hours of a university professor. A sample of 10 professors yields the following data :

48, 25, 20, 65, 72, 35, 50, 62, 42, 30

Use this data to estimate the population mean of the number of working hours per week.

5

- (b) Find the maximum likelihood estimator for parameter λ of the exponential distribution whose probability density function is given by

$$f(x) = \frac{1}{\lambda} e^{-x/\lambda}; \quad x > 0, \lambda > 0.$$

5

4. (a) An agency claimed that 70% of its employees felt that work stress had a negative impact on their personal lives. To check this statement, an analyst took a random sample of 100 employees and found that out of the 100 employees, 75 felt that work stress had a negative impact on their personal lives. Formulate the null and alternative hypotheses and test the claim at 5% level of significance.

5

(b) A company uses two sources of raw materials. Both the sources seem to have similar characteristics but the company is not sure about their respective uniformity. To check it, a sample of 10 lots from source A and a sample of 11 lots from source B are taken. The following data are obtained :

	Source A	Source B
Sample mean	84	86
Sum of squares from their mean	117	110

Is it likely that the variance of source A is significantly greater than the variance of source B at 1% level of significance ?

5

OR

A researcher wants to compare the effect of two different types of drugs, A and B, for increasing weight. She gives drug A to 5 persons and drug B to 6 persons and records the increase in weight (in pounds) which are given below :

Drug A	8	11	13	15	8	
Drug B	10	9	9	10	12	10

It is known that the effects of both drugs are normally distributed with equal variances. To examine that the two drugs differ significantly with regard to their effect in increasing weight, answer the following :

- (a) Are both samples paired or independent ? 1
- (b) State the null and alternative hypotheses. 1
- (c) Which test is used for testing the null hypothesis in this case, and why ? 2
- (d) Conduct the suitable test at 5% level of significance and interpret the result. 6
5. The table given below shows the relation between the performances of students in statistics and mathematics. Test the hypothesis that the performance of the students in statistics is independent of their performance in mathematics using 0.01 significance level. 10

Mathematics Statistics	High Grades	Medium Grades	Low Grades
High Grades	36	72	42
Medium Grades	34	122	44
Low Grades	50	56	44

OR

A company is trying to improve the work efficiency of its employees. It has organised a special training programme for all employees. In order to assess the effectiveness of the training programme, the company has selected 10 employees randomly and administered a well-structured questionnaire. The scores (out of 50) obtained by the employees are given in the following table :

Employee Number	Before Training	After Training
1	30	35
2	32	34
3	37	31
4	34	33
5	36	33
6	33	37
7	39	37
8	33	42
9	30	40
10	32	43

It is known that the distribution of the differences of the scores before and after the training programme is symmetrical about its median. To examine whether the training programme has improved the efficiency of the employees, answer the following questions :

- (a) Are both samples paired or independent ? 1
- (b) State the null and alternative hypotheses. 1
- (c) Which test is used for testing the null hypothesis in this case, and why ? 2
- (d) Conduct the suitable test at 1% level of significance and interpret the result. 6