POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Examination

June, 2021

MST-004 : STATISTICAL INFERENCE

Time : 3 hours *Maximum Marks : 50*

Note :

(i) *Question no. **1** is **compulsory**.*

(ii) *Attempt any **four** questions from the remaining questions.*

(iii) *Use of scientific (non-programmable) calculator is allowed.*

(iv) *Use of Formulae and Statistical Tables Booklet for PGDAST is allowed.*

(v) *Symbols have their usual meanings.*

---

1. State whether the following statements are *True* or *False*. Give reasons in support of your answers. *5×2=10*

(a) If degrees of freedom of a t-distribution is 8, then its mean is 8.

(b) The number of runs in the sequence ABBAAABABBABABBB is 10.

(c) If estimators t_1 and t_2 of a parameter T have variances $\frac{1}{2}$ and $\frac{1}{3}$ respectively, then t_1 is more efficient than t_2.

(d) Rejecting the null hypothesis when it is not true is called type-I error.

(e) F test cannot be applied for comparing population means.

2. (a) The following data represents the number of days absent per year in a population of 5 employees of a small company :

25, 20, 30, 15, 10

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

(ii) Compute the mean of all samples of size 2 and set up the sampling distribution of the sample mean.

(iii) Compute mean and standard error of the sampling distribution of the sample mean. *7*

(b) The temperature of the coffee sold at 16 coffee shops was measured. If the sample mean and standard deviation of the temperature are respectively 162·0°F and 10·0°F, construct a 95% confidence interval for the population mean temperature of coffee sold. Assume that temperatures are approximately normally distributed. *3*

3. (a) A new computer software package was developed to reduce the project completion time. To evaluate the benefit of the new software package, the following data were collected :

Time for Current Software (in hours)	Time for New Software (in hours)
300	274
280	220
344	308
385	336
372	198
360	300
288	315
321	258
376	318
290	310
301	332
283	263

It is assumed that the project completion times of both softwares are normally distributed with equal variances. Test whether the average project completion times of both softwares differ significantly at 5% level of significance. *8*

(b) Distinguish between Rejection and Non-rejection regions. *2*

4. (a) Discuss the properties of a good estimator with examples. *4*

(b) Determine the maximum likelihood estimator for parameter p of the binomial distribution with probability mass function given by

$$P[X = x] = {}^{n}C_{x}\, p^{x}\, q^{n-x};\ 0 < p < 1,$$
$$x = 0, 1, 2, ..., n,\ q = 1 - p.$$ *6*

5. The tensile strength for parachutes woven with synthetic fibres from four different suppliers are provided in the following table :

Supplier			
I	II	III	IV
18·5	26·3	20·6	25·4
24·0	25·3	25·2	19·9
17·2	24·0	20·8	22·6
19·9	21·2	24·7	17·5
18·0	24·5	22·9	20·4

If the tensile strength of parachutes is not normally distributed, then test the significance between the tensile strength of the parachutes supplied by all suppliers at 5% level of significance. *10*

6. (a) Differentiate between parametric and non-parametric tests. Give two examples of each kind of test. *4*

(b) The management of a prime bank would like to know whether online banking is more popular amongst younger (under 40) as compared to older banking clients. A random sample of 240 younger clients found that 125 use online banking, while from the random sample of 310 older clients, 140 use online banking. Test the appropriate hypothesis at 10% level of significance. *6*

7. (a) The mean distances travelled by drivers each day are given as follows :

Age Group	Average Distance (miles)
18 – 19	20·7
20 – 29	31·0
30 – 49	37·0
50 – 64	30·4
65 – 74	30·4

A random sample of 50 drivers of ages 18 to 19 years is selected. What is the probability that the mean distance travelled each day is between 19·4 and 22·5 miles ? *7*

Assume $\sigma = 6{\cdot}5$ miles.

(b) Determine the sufficient statistic for the parameter β of the population having the following probability density function : *3*

$$f(x, \beta) = \frac{5^{\beta}}{\sqrt{\beta}} e^{-5x} x^{\beta-1}; x \geq 0$$