No. of Printed Pages : 8 **MSTL–003(Set-II)**

POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Practical Examination

June, 2025

MSTL–003(Set-II) : BIOSTATISTICS LAB

Time : 3 Hours *Maximum Marks : 50*

***Note** : (i) Attempt any **two** questions.*

(ii) Solve the questions in Microsoft Excel.

(iii) Use of Formulae and Statistical Table Booklet for PGDAST program is allowed.

(iv) Mention necessary steps, hypotheses, interpretations, etc.

---

1. (a) A researcher wants to check the effect of smoking on Respiratory Tract

Infections (RTIs). There are 60 smokers as well as 60 non-smokers. The data are as follows :

Respiratory Tract Infections	Smoker	Non-Smoker
Yes	35	22
No	25	38
Total	60	60

Test the hypothesis that the proportions of the person having respiratory tract infections are different in smokers and non-smokers groups at 5% level of significance. 10

(b) A clinical trial was run to compare three Treatments A, B and C on patients who are suffering from cancer at advance stage. Thirty six participants who agreed to participate in the trial were randomly assigned to receive Treatments A, B and C. The primary outcome was death ad participants were

followed for upto 48 months. The survival data are as follows :

Survival Time	Outcome	Treatment
2	Died	B
4	Died	B
5	Unknown	B
6	Unknown	A
9	Unknown	B
10	Died	C
12	Died	B
12	Died	A
15	Unknown	B
18	Unknown	A
20	Died	B
21	Died	C
21	Died	A
24	Unknown	C
25	Died	C
25	Died	C
27	Died	A
27	Survived	C
29	Died	A
30	Died	C
30	Unknown	B
30	Died	A

32	Unknown	B
35	Died	C
42	Died	A
40	Died	B
40	Unknown	B
42	Unknown	C
42	Died	B
45	Died	B
45	Unknown	A
48	Survived	C
48	Survived	B
48	Survived	B
48	Survived	A
48	Survived	C

Test whether all treatments are equally effective in term of distributions of the survival times at 5% level of significance. 15

2. A sample of 20 houses was selected to develop a linear model for the electricity consumption of a household and to predict the electricity consumption during summer. We have recorded the electricity consumption (in kWh), size of house (in sq. ft) and AC (0 for not having AC and 1 for having AC) in the following table :

Table : Electricity consumption data

S. No.	Unit (in kWh)	Area (in sq. ft)	AC
1	510	725	1
2	923	1000	1
3	703	900	1
4	1043	1350	0
5	1193	1400	1
6	1048	1200	1
7	710	825	0
8	513	750	1
9	368	675	0
10	1058	1350	0
11	733	825	1
12	588	850	0
13	863	1000	1
14	778	925	1
15	918	1050	0
16	868	1100	0
17	803	1075	0
18	878	1000	1
19	663	875	1
20	818	1025	0

(a) Prepare a scatter plot to get an idea about the relationship among the variables.

(b) Develop a linear regression model and perform related analysis at 5% level of significance.

(c) Check the linearity and normality assumptions for the regression analysis.

(d) Draw both fitted regression lines on the scatter plot. 5+8+6+6

3. A researcher wants to study the relationship of number of deaths with waiting time of the patients to be attended by a doctor and number of doctors on duty in a hospital. The compiled data recorded of 30 consecutive days are as follows :

S. No.	Waiting Time (minutes)	Number of Doctors	Number of Deaths
1	5	3	4
2	3	3	5
3	4	3	1
4	3	1	0

5	4	3	5
6	4	4	6
7	4	4	2
8	4	5	1
9	3	5	6
10	3	6	7
11	3	6	3
12	3	7	2
13	2	7	7
14	2	8	8
15	4	3	0
16	4	2	5
17	4	3	1
18	3	1	0
19	4	2	6
20	4	2	6
21	3	1	0
22	3	2	5
23	4	3	2
24	4	3	2

25	2	1	0
26	2	1	0
27	4	3	0
28	3	2	3
29	2	2	0
30	3	3	1

(a) Fit a simple Poisson regression model for number of deaths considering waiting time as a regressor variable.

(b) Fit a multiple Poisson regression model for the number of deaths considering both waiting time and number of doctors as regressor variables.

(c) Determine the value of $-2 (\log L)_F$ for both fitted Poisson regression models.

10+10+5

× × × × ×