

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

Term-End Examination

June, 2026

MSTE-004 : BIOSTATISTICS—II

Time : 3 Hours

Maximum Marks : 50

Note : (i) *Question No. 1 is compulsory.*

(ii) *Attempt any **four** questions from the remaining question nos. 2 to 7.*

(iii) *Use of scientific calculator (non-programmable) 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 :

$$5 \times 2 = 10$$

- (a) If we obtained paired observations in the contingency table, then we use Fischer's exact test.
- (b) If there are 27 true negative cases out of 300 non-diseased cases, then the specificity of the test will be 0.9.
- (c) To find the regression coefficient, in linear regression model, only the method of maximum likelihood is applied.
- (d) In a study, it was found that the probability of killing a bug using 0.4 mg of a pesticide is 0.75, then log-odds will be approximate 1.386.
- (e) If the risk of mortality is high at the initial stage and it decreases as time increases, then it is called wear-out phase.

2. A researcher is interested to evaluate the association between coronary heart disease (CHD) and obesity. For this purpose, a random sample of 100 patients was taken and obtained the following data :

	No Obesity	Obesity
No CHD	18	16
CHD	18	48

- (i) Test at 5% level of significance whether CHD and obesity status are associated.
- (ii) If the researcher obtained the following data :

	No Obesity	Obesity
No CHD	52	2
CHD	14	32

then why the test applied on the part (i) is not appropriate to the aim of the study in part (ii) ? Perform an appropriate analysis and interpret the results. 10

3. A researcher was interested to study the impact of the birthweight and age of the infant on the systolic blood pressure (SBP).

For this purpose, the researcher measured SBP (mm Hg), birth weight (Oz) and age (days) for 16 infants and the obtained data are given as follows :

SBP	Birth weight	Age	SBP	Birth weight	Age
89	135	3	96	120	5
90	120	4	95	90	4
83	100	3	80	120	2
77	105	2	79	95	3
92	130	4	86	120	3
98	125	5	97	150	4
82	125	2	92	160	3
85	105	3	88	125	3

If the researcher obtained the regression coefficients as $\hat{\beta}_0 = 53$, $\hat{\beta}_1 = 0.1$, $\hat{\beta}_2 = 6$, then :

- (i) Obtain the predicted values of SBP using multiple regression model.

(ii) Calculate the residuals. Check whether $\sum r_i = 0$ or not.

(iii) Obtain the estimate of σ^2 .

(iv) Determine the standardized residuals.

10

4. A researcher conducts a case control study in which 500 males are in case group having lung cancer and 500 males are in control group who do not have lung cancer. They were asked whether they were exposed to smoking or not. The obtained data are given as follows :

	Lung Cancer (Case)	No Lung Cancer (Control)
Smoker	226	105
Non-smoker	274	395
Total	500	500

(i) Fit a logistic model.

(ii) Determine the probability of having lung cancer in both groups using logistic model.

- (iii) Compute log-odds for smokers and non-smokers.
- (iv) Obtain the odds ratio for lung cancer. 10
5. A small prospective study is conducted and followed ten participants for the development of myocardial infarction (MI) over a period of 10 years. The observed data are given as follows :

Participant ID	MI Time (in year)
1	5
2	8
3	2
4	6 ⁺
5	9
6	6
7	10 ⁺
8	4
9	8
10	4 ⁺

Estimate the survival function.

10

6. (a) If a researcher obtained the following survival probabilities in a study : 5

Survival Time	Survival Probability
0	1
2	0.9
4	0.8
4 ⁺	0.8
5	.687
6	0.571
6 ⁺	0.571
8	0.286
9	0.143
10 ⁺	0.143

then compute standard error.

- (b) Mention the assumptions of multiple linear regression. 5
7. (a) State the assumptions for applying the chi-square test for association. 4
- (b) Describe the polytomous model. 6

× × × × ×