No. of Printed Pages : 5

MSTE–004

POST GRADUATE DIPLOMA IN APPLIED STATISTICS (PGDAST)

Term-End Examination

December, 2025

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 : 5×2=10

   (a) For Chi-square test for homogeneity of proportions, the samples from two populations should be dependent and random.

   (b) A regression model Y = 2.5 + 0.15X is significant at $\alpha = 0.5$.

   (Given that : $n = 12$ and $SE(\hat{\beta}_1) = 0.05$).

   (c) For a logistic model if $\hat{\beta}_1 = 12$ and $SE(\hat{\beta}_1) = 2$, 95% confidence interval of β_1 is (8.1, 15.9).

   (d) If the hazard rate is $h(t) = t$, the survival function will be $\exp\left(-\frac{t^2}{2}\right)$.

   (e) For fitting a multiple regression model, all regressor variables should be correlated to each other.

2. The data classified according to different levels of alcohol consumption and cardiovascular disease (CVD) are given in the following table :

CVD	Alcohol Consumption				
	Never	Rarely	Sometimes	Often	Always
Yes	16	10	14	21	34
No	31	25	28	13	8

Test the linear association between cardiovascular disease and alcohol consumption at 5% level of significance. 10

3. The Systolic Blood Pressure (SBP), age and gender of 12 persons are given in the following table : 10

S. No.	SBP	Age	Gender
1	120	34	Male
2	115	34	Male
3	110	32	Male
4	145	38	Male
5	135	36	Male
6	160	42	Male
7	178	40	Female
8	165	37	Female

9	175	38	Female
10	185	41	Female
11	200	42	Female
12	190	42	Female

(i) Fit a multiple regression model of SBP on age and gender.

(ii) Determine the regression model for male and female separately.

4. For a logistic model :

$$Y = \frac{\exp(-3.5 + 0.07X)}{1 + \exp(-3.5 + 0.07X)},$$

the following information is given :

$$SE(\hat{\beta}_0) = 1.4 \text{ and } SE(\hat{\beta}_1) = 0.04 .$$

(i) Test the significance of the model parameters using Wald Chi-square statistic at 5% level of significance. 5

(ii) If the log-likelihood of the fitted, saturated and reduced models are – 64.7, – 64.6 and – 66.8, respectively, test the significance of the fitted logistic model using deviance and G statistic(s). (Given that : $n = 4$). 5

5. (a) If the survival time (in years) has the following probability density function : 7

$$f(t) = \begin{cases} \theta e^{-\theta t}; & \theta > 0, t \geq 0 \\ 0 \ ; & \text{otherwise} \end{cases}$$

then determine :

(i) Survival function

(ii) Cumulative distribution function

(iii) Hazard function

(iv) Median when $\theta = 0.2$

(b) Describe the Cox proportional hazard model. 3

6. (a) Describe the Baseline category logit models. 5

(b) Explain the receiver operating characteristic curve. 5

7. (a) Write a short note on the normal probability plot. 5

(b) Explain the hazard and cumulative hazard functions. 5

× × × × ×