[This question paper contains 2 printed pages.]

Your Roll No...............

Sr. No. of Question Paper	**:**	**8138**	**L**
Unique Paper Code	:	2373010020	
Name of the Paper	:	Generalized Linear Models	
Name of the Course	:	**B.Sc. (Honours) Statistics, NEP-UGCF**	
Semester	:	VIII	

Duration : 3 Hours — Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **FIVE** questions in all.
3. Use of a non-programmable scientific calculator is allowed.
4. Use of Statistical Tables is allowed.

1. (a) What is the key difference between a linear regression model and a nonlinear regression model?
   (b) Give one example of a nonlinear model that can be transformed to a linear model.
   (c) In logistic regression, what does the odds ratio represent? Write the logit function.
   (d) State the null hypothesis for a test of independence in a two-way contingency table using a log-linear model.
   (e) What is overdispersion in Poisson regression? Name one common cause.
   (f) What are the main challenges in the estimation of parameters in a nonlinear regression model compared to linear regression?

(3x6)

2. (a) Discuss the origins of nonlinear regression models in detail. Explain the difference between intrinsically linear and intrinsically nonlinear models with suitable examples.
   (b) Consider the nonlinear growth model $Y = \frac{\theta_1}{1+\theta_2 e^{-\theta_3 X}} + \varepsilon$ (logistic growth curve). Can this model be transformed to a linear model? If yes, show the transformation and state the limitations. If no, explain why not.

(9,9)

3. (a) What are the common methods for estimating parameters in nonlinear regression models?
   (b) The following data represent the length (Y) of a fish as a function of its age (X in years):

X	1	2	3	4	5	6	7	8
Y	3.5	5.8	7.2	8.1	8.6	8.9	9.1	9.2

It is hypothesized that $Y = \alpha(1 - e^{-\beta X}) + \varepsilon$. Using (i) a suitable transformation (if possible), or (ii) without transformation, estimate α and β. Also, predict the length at age 10 years.

(6,12)

P.T.O.

4. (a) Define the logistic regression model for multiple predictors. Explain the concept of link function in the context of GLM and identify the link function used in logistic regression.

(b) A study examines the effect of a new drug on recovery (1=recovered, 0=not recovered). Data are as follows:

Dose (mg)	Recovered	Not Recovered	Total
0	10	90	100
10	30	70	100
20	50	50	100
30	70	30	100

Fit a logistic regression model with dose as predictor. Estimate the coefficients, interpret the odds ratio, and find the ED50 (dose at which recovery probability = 0.5).

(6,12)

5. (a) Describe the Poisson regression model for count data. Explain the role of an offset variable with a real-life example.

(b) Define deviance in Poisson regression. How is the analysis of deviance used to compare nested Poisson models?

(9, 9)

6. (a) Write a note on goodness-of-fit tests for Poisson regression models. Describe the deviance-based tests.

(b) The number of insurance claims (Y) for 10 policyholders is recorded along with age (X_1 in years) and car type ($X_2 = 1$ for sports car, 0 otherwise). Fitted Poisson model gives:

$$\ln(\hat{\mu}) = -1.2 + 0.03 \times \text{Age} + 0.8 \times \text{CarType}$$

$SE(\hat{\beta}_{Age}) = 0.01$, $SE(\hat{\beta}_{CarType}) = 0.25$.

(i) Test whether Age is significant at 5% level.

(ii) Interpret the coefficient of CarType.

(iii) Predict the expected number of claims for a 30-year-old sports car owner.

(6,12)

7. (a) Explain log-linear models for three-way contingency tables. Write the models for:

(i) mutual independence,

(ii) joint independence,

(iii) conditional independence.

(b) A 2×2×2 table for variables X (Gender: M/F), Y (Treatment: Yes/No), Z (Recovery: Yes/No) is given:

Gender	Treatment	Recovered	Not Recovered
M	Yes	40	10
M	No	15	35
F	Yes	35	15
F	No	20	30

Test the hypothesis of conditional independence of Treatment and Recovery given Gender using a log-linear model. Compute expected frequencies and likelihood ratio test statistic.

(9,9)

(500)