M.Sc. Examination, 2018
Semester-III
Statistics
Course : MSC-32

(Categorical Data Analysis and Advanced Data Analysis)

Time : 3 Hours **Full Marks : 40**

Questions are of value as indicated in the margin

Answer **any four** questions

1. Distinguish between case-control study and cohort study. Discuss about their relative merits and demerits. 5+5=10
2. a) Define a measure of association between two ordinal categorical variables in a 2×2 contingency table. Find the large sample standard error of your measure.
   b) Develop a test procedure for testing the independence of the variables. (2+5)+3=10
3. a) Discuss about three major components of the GLM.
   b) Show that the likelihood equations for a generalized linear model (with usual notation) can be written as

   $$\Sigma w(y-\mu)\frac{\partial \eta}{\partial \mu}xj = 0 \ i \text{ for each j.}$$

   c) Define Pseudo – R^2 measure for the GLM. 3+5+2=10
4. a) Stating the structural defect of linear probability model, discuss the genesis of logistic regression model.
   b) Interpret the parameters of the logistic regression model with one covariate.
   c) Discuss the fitting procedure of logistic regression model. 3+2+5=10
5. Discuss about the use of EM-algorithm to cluster data which can be modelled as a mixture of univariate normal populations. 10
6. How can you apply Gibbs sampling technique in one way ANOVA model to estimate the parameters and hyper parameters. 10
7. a) Write the bootsrap algorithm to find the standard error of a statistic. Also distinguish between bootstrap and jackknife techniques.
   b) Write a short note on Metropolis – Hastings algorithm. 5+5=10

_____