



QP CODE: 26019601



Reg No :

Name :

M.Sc DEGREE (CSS) EXAMINATION, MARCH 2026

Fourth Semester

MSc Statistics with Data Science

ELECTIVE - ST880403 - ANALYSIS OF CATEGORICAL DATA

2020 Admission Onwards

9D38C078

Time: 3 Hours

Weightage: 30

Part A (Short Answer Questions)

*Answer any **eight** questions.*

Weight 1 each.

1. Distinguish between inferential and descriptive statistics giving examples.
2. What are the two important measures of disease in epidemiology? Briefly explain any one measure.
3. Define sensitivity and specificity.
4. What is a latent variable?
5. Derive the relation between odds ratio and the parameter of a Logistic regression model.
6. What are the Simple Residuals in logistic regression?
7. Define a Poisson regression model for count data and rate data.
8. Discuss the effect of application of the negative binomial model to data that are not over dispersed relative to the Poisson.
9. What are random number generators? What are the different types of random number generators?
10. What are the different methods to generate random numbers from a probability distribution?
(8×1=8 weightage)

Part B (Short Essay/Problems)

*Answer any **six** questions.*

Weight 2 each.

11. In an outbreak of varicella (chickenpox) in Oregon in 2002, varicella was diagnosed in 18 of 152 vaccinated children compared with 3 of 7 unvaccinated children. Calculate the risk ratio and interpret it.





12. Give the different chi-square test used in analysing categorical data. Explain any two of them.
13. Give the advantages of GLMs over the traditional (OLS) regression.
14. Discuss the Hosmer–Lemeshow test to evaluate a Logistic regression model fit.
15. If Poisson means follows a gamma distribution with shape parameter $(1-p)/p$ then what is the resulting distribution? Give its pdf. Does negative binomial distribution allow overdispersion and under dispersion? Explain.
16. What is Cox proportional hazard model? What are its assumptions? State the advantages of a Cox proportional hazard model.
17. What you mean by Monte Carlo simulation? What are the different mechanisms to generate random numbers from a distribution of our choice?
18. State and prove Bayes' Theorem and explain how do you translate this theorem as a tool for inference.

(6×2=12 weightage)

Part C (Essay Type Questions)

*Answer any **two** questions.*

*Weight **5** each.*

19. Explain the general linear probability model and logit model and compare the two giving examples.
20. Explain the generalized linear model for binary data and discuss the use of logit in logistic regression model.
21. (a) Derive the deviance formula for Poisson regression model, (b) How will you test the lack of fit of a Poisson regression model?
22. Explain Gibbs sampling procedure to generate a random sample from a multivariate distribution, explain the use of this method with an example.

(2×5=10 weightage)

