AA–3354

Seat No. ______

M. Phil. Examination

April / May – 2003

Statistics (Compulsory)

(Research Methodology)

Time : **3** Hours] [Total Marks : **100**

Instructions : (1) Each question carries **20** marks.
(2) Use of statistical tables and scientific calculator is permitted.

1 (a) Derive OLS estimators and their dispersion matrix for general linear model with data matrix having full rank and when the parameters are subjected to given linear restrictions.

(b) What is the problem of heteroscadasticity ? Enumerate different tests of heteroscadasticity. How would you apply 2 SLS method to tackle the heteroscadastic situation ?

(c) Explain how dummy variables are useful in econometric analysis.

OR

1 (a) Discuss in detail Silvey's approach of tackling the problem of multicolinearity in linear models.

(b) Obtain Generalised Difference Equation for linear models and discuss how it will be useful to tackle the problem of auto correlation when serial correlation coefficient is known or unknown.

(c) Show how Koyck's transformation can convert dL Models into AR Models.

2 (a) State and prove Gauss Marfoff Theorem for linear models having less than full rank.

(b) Discuss important applications of linear models in different fields.

(c) Show that the general solution of the non-homogenous consistent linear equations $A\underline{x} = \underline{b}$ is given by $\underline{x} = \bar{A}\underline{b} + (\underline{T} - H)\underline{Z}$.

OR

2 (a) For the system of non homogenous linear equations $\underline{Y} = X\underline{\beta} + \underline{\in}$, having general solution $\underline{\hat{\beta}} = \bar{S}\ X'\underline{Y} + (I - H)\underline{Z}$ where $\bar{S}$ is any g-inverse of $X'X$, $H = \bar{S} - S$, show that (i) if $\bar{S}$ is a g-inverse of $X'X = S$ then $(\bar{S})'$ is also a g-inverse. (ii) $X = XH$ (iii) if $\bar{S}_a$ and $\bar{S}_b$ are two g-inverses of X, then $X\ \bar{S}_a X' = X\bar{S}_b X'$. (iv) a solution of the normal equation $X'\underline{Y} = (X'X)\underline{\beta}$ is unique if and only if $\rho(X) = \rho(X'X) = K$.

(b) Write explanatory note on :

'Piecewise Linear Regression Model'.

3 (a) Discuss briefly about (i) Monte Carlo method and (ii) Mid square method for simulation.

(b) How would you select a random sample of size n from the following populations using simulation method ?

(1) $f(x,\ \theta) = \theta\, e^{-\theta x}, \quad x \in R^+, \quad \theta > 0$

(2) $f(x,\ \theta) = \dfrac{e^{-\theta}\theta^x}{x!}, \quad x = 0,\ 1,\ 2........, \quad \theta > 0.$

(c) A certain company wants to market a new product. The fixed cost for the project is Rs. 4,000. There are three uncertain factors : Selling price, variable cost and sales volume for the new product to be launched. The product has a life of one year. Following data are for consideration by the management :

Factor - 1		***Factor - 2***		***Factor - 3***	
Selling Price (in Rs.)	***Probability***	***Variable cost (in Rs.)***	***Probability***	***Sales volume (in units)***	***Probability***
3	0.20	1	0.40	2000	0.40
4	0.50	2	0.50	3000	0.40
5	0.30	3	0.10	5000	0.20

Run simulation for 10 days and estimate average profit.

OR

3 (a) Define CAN estimator. Show that sample distribution function is a CAN estimator for population distribution function.

(b) If $T \sim AN\left(\theta, \dfrac{V(\theta)}{n}\right)$ where $V(\theta) > 0$ and if Ψ is such that $\dfrac{d\Psi}{d\theta} \neq 0$ and is continuous then prove that

$$\Psi(T) \sim AN\left[\Psi(\theta), \frac{V(\theta)}{n}\left(\frac{d\Psi}{d\theta}\right)^2\right].$$

(c) Let $(x_1, x_2, \ldots, x_n)$ is a random sample taken from a Poisson distribution with mean $\theta, \theta > 0$. Show that $\bar{x}\, e^{-\bar{x}}$ is CAN estimator for $\theta\, e^{-\theta}$ and derive its asymptotic variance.

4 (a) Consider multinomial distribution in four cells with $P_1(\theta) = \dfrac{2+\theta}{4}$, $P_2(\theta) = P_3(\theta) = \dfrac{1-\theta}{4}$, $P_4(\theta) = \dfrac{\theta}{4}$. Here $0 < \theta < 1$ is the linkage factor. Consider a sample of size n with data given by (n_1, n_2, n_3, n_4) as cell frequencies. Then show that $\hat{\theta} \sim AN\left(\theta, \dfrac{2\theta(1-\theta)(2+\theta)}{n(1+2\theta)}\right)$; where $\hat{\theta}$ is the mle of θ.

(b) Explain the method of constructing asymptotic confidence interval (ACI). For Cauchy distribution with parameter θ show that ACI for θ is $\left(\hat{\theta} \pm \sqrt{\dfrac{2}{n}}\; \zeta_{1-\frac{\alpha}{2}}\right)$, where n = sample size and $\hat{\theta}$ is the mle of θ.

(c) Explain the method to construct studentized version of the asymptotic confidence interval with suitable example.

OR

4 (a) Consider the Poreto distribution with pdf

$f(x;\ \theta) = \dfrac{\theta}{x^{\theta+1}},\ x \geq 1,\ \theta > 0,$ and zero otherwise. Show that it belongs to one parameter exponential family and obtain the MLE $\hat{\theta}$ and its asymptotic distribution.

(b) If $\{X_1,\ X_2, \ldots\ldots, X_n\}$ are iid random variables with $E(X_i) = \mu(\theta), V(X_i) = \sigma^2(\theta)$ such that $\dfrac{d\mu}{d\theta} \neq 0$ and is continuous then show that

$$\mu^{-1}(\bar{X}) \sim AN\left(\theta,\ \frac{\sigma^2(\theta)}{n}\Big/\left(\frac{d\mu}{d\theta}\right)^2\right)$$

where $\bar{x} =$ sample mean.

(c) Let $\{X_1,\ X_2, \ldots\ldots, X_n\}$ are n iid random variables from $N(\theta_1,\ \theta_2)$ then obtain ACI for $\theta_1 + \zeta_p\sqrt{\theta_2}$, where ζ_p is the 100 p % percentile of standard normal distribution.

5 (a) Discuss in detail the main steps involved in a sample survey.

(b) Discuss in detail family living surveys for working class.

(c) Explain the terms :

(1) Sampling and non–sampling errors.

(2) Sample check.

(3) Post census.

OR

5 (a) Discuss in detail National Sample Survey for rural sector.

(b) Define non-response errors. Suggest an estimator for population total when n_1 units for n units respond and if the units are selected with SRSWOR. Also find its variance.