

**STATISTICS**  
**Paper III**

Time Allowed : Three Hours

Maximum Marks : 200

**QUESTION PAPER SPECIFIC INSTRUCTIONS**

**Please read each of the following instructions carefully before attempting questions.**

There are **EIGHT** questions divided under **TWO** Sections.

Candidate has to attempt **FIVE** questions in all.

Both the questions in Section 'A' are compulsory.

Out of the **SIX** questions in Section 'B', any **THREE** questions are to be attempted.

The number of marks carried by a question/part is indicated against it.

Unless otherwise mentioned, symbols and notations have their usual standard meanings.

Assume suitable data, if necessary, and indicate the same clearly.

Attempts of questions shall be counted in sequential order. Unless struck off, attempt of a question shall be counted even if attempted partly.

Any page or portion of the page left blank in the Question-cum-Answer (QCA) Booklet must be clearly struck off.

Answers must be written in **ENGLISH** only.

## SECTION 'A'

- 1.(a) From a simple random sample of size  $n$  taken without replacement from  $N$  units, a simple random sample of size  $n_1$  is duplicated without replacement and added to the original sample. Show that the mean based on  $(n + n_1)$  units is an unbiased estimate of population mean. Obtain the variance of this estimate. 15

- 1.(b) Obtain Generalised Least Square Estimator (GLSE) of  $\underline{\beta}$  in,

$$\underset{n \times 1}{Y} = \underset{n \times k}{X} \underset{k \times 1}{\underline{\beta}} + \underset{n \times 1}{U}, \quad \text{with } E(\underline{U} \underline{U}') = \underset{n \times n}{\sigma^2 \Omega},$$

$\Omega$  is a positive definite matrix. Also derive variance covariance matrix of this GLSE and write down its unbiased estimator. 15

- 1.(c) Explain price elasticities of supply and demand. For the demand function

$$D = Cp^{-\alpha}, \quad C > 0, \quad \alpha > 0$$

Obtain price elasticity of demand and comment on its shape. What happens when  $\alpha = 1$ ? 10

- 2.(a) In systematic sampling when  $N = nK$ , prove that the variance of the  $r^{\text{th}}$  systematic sample is given by

$$V(\bar{y}_r) = \frac{(N-1)S^2}{N} - \frac{(n-1)}{n} S_{wsy}^2$$

where

$$S^2 = \frac{1}{(N-1)} \sum_{r=1}^K \sum_{i=1}^h (y_{ri} - \bar{Y})^2 \quad \text{and} \quad S_{wsy}^2 = \frac{1}{(n-1)} \sum_{i=1}^h (y_{ri} - \bar{y}_r)^2$$

Hence, show that systematic sampling is more efficient as compared to simple random sampling without replacement if within systematic sample variation is more as compared to the overall variation of population, that is,  $S^2$ . 10+5

2.(b) Consider the model

$$\underline{Y} = X\underline{\beta} + \underline{U}, \quad \text{with } E(\underline{U}\underline{U}') = \sigma^2 I.$$

$$\text{Given, } (X'X) = \begin{pmatrix} 270 & 240 \\ 240 & 630 \end{pmatrix} \text{ and } (X'Y) = \begin{pmatrix} 319 \\ 492 \end{pmatrix},$$

obtain the ordinary least square estimator (OLSE) of  $\beta$  and obtain the variances and covariances between its components. 10

2.(c) State the methods for estimating seasonal variation in time series. Compare “Ratio to trend” and “Ratio to moving average” methods.

Assuming that the trend is absent, determine seasonal indices for the following data and interpret your result.

<i>Year</i>	<i>1st Quarter</i>	<i>2nd Quarter</i>	<i>3rd Quarter</i>	<i>4th Quarter</i>
2007	30	40	36	34
2008	34	52	50	44
2009	40	58	54	48
2010	54	76	68	62
2011	80	92	86	82

15

### SECTION 'B'

3.(a) Differentiate between ordered and unordered estimators in pps sampling without replacement. Write down the Des Raj Estimator, its mean and variance for estimating population total when a sample of size 2 is drawn from population of size  $N$  by ppswor. 15

3.(b) Writing down the assumptions usually made in general linear regression model

$$\underline{Y} = X\underline{\beta} + \underline{U}, \quad \text{where } \underline{U} \sim N_p(0, \sigma^2 I).$$

Obtain the maximum likelihood estimator (mle) of  $\sigma^2$ . Also examine its unbiasedness. 15

- 3.(c) The Ministry of Labour retail price index was 104.6 given for a certain time period. The price indices in July 2004 were as follows :

Rent and Rates 65, Clothing 220, Fuel and Light 110, Miscellaneous 125.

Given that the weights of different items in the group were as follows :

Food 60, Rent and rates 16, Clothing 12, Fuel and Light 8 and Miscellaneous 4.

What was the price index of the food group ? 10

- 4.(a) From a list of 3042 names and addresses, a simple random sample of 200 names, taken without replacement, showed on investigation 18 wrong addresses. Estimate the total number of addresses needing corrections in the list and find the standard error of this estimate. 10

- 4.(b) State the rank and order conditions of identification for structural form of the system.

Discuss the identification of the following Keynesian model of income determination :

$$\begin{aligned} C_t &= \alpha_0 + \alpha_1 Y_t + \alpha_2 T_t + U \\ I_t &= \beta_0 + \beta_1 Y_{t-1} + V \\ T_t &= V_0 + V_1 Y_t + W \\ Y_t &= C_t + I_t + G_t \end{aligned} ,$$

where  $C_t$ ,  $I_t$ ,  $T_t$  and  $Y_t$  are endogenous variables and  $Y_{t-1}$  and  $G_t$  are predetermined variables. 15

- 4.(c) Explain the concept of unit root test. How unit root test is used while applying Dickey Fuller (DF) test for testing stationarity of a given time series. Interpret the following models of time series  $y_t$  for testing stationarity of the time series.

(i)  $\hat{y}_t^* = 0.00576 y_{t-1}$   
 $t = (5.7980)$

(ii)  $\hat{y}_t^* = 28.2054 - 0.00136 y_{t-1}$   
 $t = (1.1576) \quad (-0.2191)$

(iii)  $\hat{y}_t^* = 190.3857 + 1.4776t - 0.0603 y_{t-1}$   
 $t = (1.8389) \quad (1.6109) \quad (-1.6252)$

where  $y_t^* = (y_t - y_{t-1})$  and  $t = (\tau)$  value are the critical values for the models. (The tabulated values at 5% for

- (i) there is no constant is  $-1.9434$   
(ii) there is a constant is  $-2.8947$   
(iii) there is a constant and trend is  $-3.4614$

15

5.(a) If variates  $y_i, x_i$  are measured on each unit of a simple random sample of size  $n$

(assumed to be large) then show that variance of  $\hat{R} = \frac{\bar{y}}{\bar{x}}$  is approximately

$$V(\hat{R}) = \frac{(1-f)}{n \bar{X}^2} \frac{\sum_{i=1}^N (y_i - R x_i)^2}{(N-1)}$$

where  $R = \frac{\bar{Y}}{\bar{X}}$  is the ratio of population means and  $f = \frac{n}{N}$ ,  $N$  is the size of population. 15

5.(b) Considering the model

$$Y_i = \alpha + \beta X_i + U_i,$$

use the Durbin-Watson test for autocorrelation, with the following observations on  $Y$  and  $X$  and give your conclusions :

$X :$	1	2	3	4	5	6	7
$Y :$	2	2	2	1	3	5	6

(Given : the values of  $d_L$  and  $d_u$  on 5% level of significance with  $n = 7$  and one explanatory variable are  $d_L = 0.700$  and  $d_u = 1.356$ ) 10

5.(c) Describe autocorrelation function and correlogram. If a given time series of profit consists of 90 observations then while doing correlogram what is appropriate lag length to be taken and why ? Explain the role of correlogram in testing stationarity of given time series. 15

6.(a) If there are two strata and  $\phi$  be the ratio of actual  $\left(\frac{n_1}{n_2}\right)$  and  $\left(\frac{n_1}{n_2}\right)$  obtained through Neyman Allocation in stratified random sampling, then ignoring f.p.c, show that whatever be the values of  $N_1, N_2, S_1$  and  $S_2$  the ratio

$$\frac{V(\bar{y}_{st})_N}{V(\bar{y}_{st})}$$

is never less than  $\frac{4\phi}{(1+\phi)^2}$  where  $V(\bar{y}_{st})_N$  and  $V(\bar{y}_{st})$  represent the variances of sample mean under Neyman Allocation and actual allocation respectively. 15

- 6.(b) Obtain the Least Square Estimator of  $\underline{\beta}$  in General Linear Model (GLM)

$$\underline{Y}_{n \times 1} = \underline{X}_{n \times k} \underline{\beta}_{k \times 1} + \underline{U}_{n \times 1}$$

with usual assumptions, when linear restrictions are imposed in the form

$$\underline{r}_{g \times 1} = R_{g \times k} \underline{\beta}_{k \times 1}, \quad r \text{ and } R \text{ are known constants.}$$

Also obtain variance-covariance matrix of this estimator. 15

- 6.(c) Describe the scope and problems which generally occur in the construction of consumer price index numbers. 10

- 7.(a) What is the main drawback of Cumulative Total Method for drawing a pps sample ? Explain Lahiri's Method of drawing a pps sample.

Select a pps sample of size 4 by Lahiri's Method from the population given below if random numbers selected are (8, 16), (10, 13), (4, 26), (5, 35), (7, 26) 6+4

<i>Serial Number of unit</i>	<i>Size of the unit</i>
1	50
2	30
3	45
4	25
5	40
6	26
7	44
8	35
9	28
10	27

- 7.(b) Explain the concept of heteroscedastic disturbances in General Linear Regression Model (GLRM). Also explain their consequences. 15

- 7.(c) Differentiate between fixed base and chain base indices. In which type of situations chain base indices are derived? Write merits and demerits of chain base indices.

For the following series of index numbers of price of four commodities, calculate the chain index number and fixed base index number. Write your conclusion. 15

Commodity	Price in rupees				
	2007	2008	2009	2010	2011
A	2	3	4	3	7
B	3	6	9	1	3
C	4	12	20	8	19
D	5	7	18	11	22

- 8.(a) In case of cluster sampling with  $N$  clusters, each of size  $M$ ,  $n (< N)$  clusters are selected using simple random sampling without replacement (SRSWOR). Write down an unbiased estimator ( $\bar{y}_c$ ) for estimating population mean ( $\bar{Y}$ ) in this case. Obtain its variance and show that

$$V(\bar{y}_c) = \left( \frac{1}{n} - \frac{1}{N} \right) S_b^2$$

where  $S_b^2 = \frac{(NM - 1)S^2 - N(M - 1)S_w^2}{M(N - 1)}$  15

- 8.(b) What is the nature of autocorrelation disturbances in General Linear Regression Model (GLRM)? Also state the consequences of autocorrelated disturbances on GLRM. 10

- 8.(c) Explain the following terms using proper examples :

- (i) Random walk with Drift
- (ii) Trend stationary time series
- (iii) Difference stationary time series
- (iv) Deterministic trend
- (v) Time series integrated of order  $d$

15

