

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 in **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 Booklet must be clearly struck off.

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

SECTION A

- Q1.** (a) The following table shows the initial probability of selection (p_i) of the seven units of a population $U : \{U_i ; i = 1, 2, 3, 4, 5, 6, 7\}$. Let a random sample of size 3 be selected from the population using probability proportional to size without replacement scheme. Let the units selected in the three successive draws be U_7 , U_3 and U_5 respectively. Find the selection probabilities with which these units were selected : 15

Population units (U_i)	U_1	U_2	U_3	U_4	U_5	U_6	U_7
Label (i)	1	2	3	4	5	6	7
Initial probability of selection (p_i)	0.15	0.20	0.12	0.29	0.14	0.07	0.03

- (b) What are 'Linear Systematic' and 'Circular Systematic' sampling schemes? Under what condition is Linear Systematic sampling scheme applied for selecting a sample of size n from the population of N units? Show that Linear Systematic sampling is also a probability sampling scheme which ensures equal probability of inclusion of each unit of the population in the sample. 10
- (c) Explain the processes for analysing an economic problem, being an econometrician following the classical methodology. Illustrate the proceeding steps with Keynesian theory of consumption. 15

- Q2.** (a) The joint pdf of (X, Y) is given by

$$f(X, Y) = \begin{cases} 1, & \text{if } |Y| < X, \quad 0 < X < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find the regression of Y on X and X on Y . Interpret the obtained results. 15

- (b) Distinguish clearly between fixed base and chain base index numbers, and point out their relative merits and demerits. Calculate fixed base index numbers from the chain base index numbers given below : 15

Year	2005	2006	2007	2008	2009	2010	2011
Chain Base Index Number	98	102	115	135	125	103	95

- (c) Discuss the components of national income according to the expenditure approach. How are these components calculated? 10

SECTION B

- Q3.** (a) In a class of 200 students, 15 failed in a paper of Mathematics. If a random sample of 5 students was selected from the population using simple random sampling without replacement scheme in order to estimate the proportion of failed students in the population, find the sampling variance of the suggested estimator. 10
- (b) What do you mean by interpenetrating sub-sampling ? With what purposes are they used often ? Explain its use in any sampling scheme. 15
- (c) Define Das and Des Raj estimators of population mean on the basis of a sample of size n . Taking $n = 2$, show that Des Raj's estimator is an unbiased estimator. Obtain the sampling variance of Des Raj's estimator for $n = 2$. 15

- Q4.** (a) A population is divided into two strata such that the ratio of their sizes, N_1 and N_2 is 11 and ratio of their standard deviations, σ_1 and σ_2 is 9. The ratio of the sample sizes for an arbitrary allocation is n_1/n_2 and that for Neyman allocation is n'_1/n'_2 and the ratio of n_1/n_2 and n'_1/n'_2 is 2. If the objective is to estimate the population mean by taking with replacement samples from within strata, compare the precision of the two allocations. 15
- (b) Describe the operational procedure of Lahiri's method for selecting a random sample with unequal probability of selection. Show theoretically that with this procedure the units are selected with probability proportional to their size. Mention the advantages and disadvantages of this method. 10
- (c) Give the operational procedure of two-stage sampling with the help of an example. Suggest an unbiased estimator of population mean. A population of 6 clusters, each having 6 values of the study variable Y is given below :

Cluster	Y-values	Cluster	Y-values	Cluster	Y-values
I	2, 4, 6, 1, 3, 5	III	4, 3, 6, 2, 1, 5	V	2, 4, 6, 8, 3, 5
II	2, 5, 3, 4, 7, 4	IV	3, 2, 5, 1, 6, 4	VI	4, 1, 2, 7, 5, 3

Let the clusters II, IV and VI be selected randomly in the first-stage sample. Further, let the Y -values (2, 7, 3) of cluster II, (6, 3, 1) of cluster IV and (7, 4, 3) of cluster VI be selected randomly in the second-stage sample. Estimate the population mean on the basis of the suggested estimator. 15

- Q5.** (a) Explain identification problem. What are the conditions for identification ?

Examine the identifiability of the following supply-demand model clearly mentioning the assumptions used, if any.

$$Q_d = \beta_0 + \beta_1 P + U$$

$$Q_s = \alpha_0 + \alpha_1 P + V$$

Given, $Q_d = Q_s$. 15

- (b) What is heteroscedasticity ? Explain the consequences of heteroscedasticity on OLS estimators. Also explain Spearman's Rank Correlation test for heteroscedasticity. 15
- (c) Discuss the Koyck method of estimating distributed-lag model and also explain the features of Koyck transformation. 10

- Q6.** (a) Describe the methods used for estimating simultaneous equation models, including the Two Stage Least Squares (2 SLS) and Three Stage Least Squares (3 SLS) methods with their advantages and limitations. 15

- (b) Explain the von Neumann test for auto-correlation and provide the conclusions. Draw the relationship between Durbin-Watson (DW) d-statistic and von Neumann test. 15

- (c) What is multicollinearity ? Why does multicollinearity arise ? How is multicollinearity detected using Farrar-Glauber Test ? 10

- Q7.** (a) (i) What do you understand by seasonal variations in a time series data ? Give example. Explain the method of link relatives of computing the seasonal indices.
- (ii) Write down the utility of time series for economic and business data. 9+6

- (b) What do you mean by income elasticity of demand ? Given family-budget data, how would you estimate this elasticity ? What adjustments would you make for variation in the size of the family ? 15

- (c) Find the first three autocorrelation coefficients of the following moving average process :

$$X_t = 1.6 e_t - 0.8 e_{t-1} + 0.4 e_{t-2} - 0.2 e_{t-3},$$

where $\{e_t\}$ are independent and identically distributed, each with normal distribution having mean 0 and the same variance. 10

- Q8.** (a) (i) What are the assumptions underlying the ARIMA model, and how do violations of these assumptions affect model performance ?
- (ii) How do you interpret ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots in the context of ARIMA model ? 8+7
- (b) What are the advantages and limitations of using a periodogram for spectral analysis ? 10
- (c) (i) Describe various types of errors that can occur in index numbers.
- (ii) During a certain period the cost of living index goes up from 110 to 200 and the salary of a worker is also raised from ₹ 500 to ₹ 900. Does the worker really gain and if so, by how much, in real terms ? 9+6

