I.E.S./I.S.S. Examination – 2011

E-JTT-L-TUC

STATISTICS – III

Time Allowed : Three Hours | *Maximum Marks : 200*

INSTRUCTIONS

Candidates should attempt FIVE questions in ALL including Questions No. 1 and 5 which are compulsory. The remaining THREE questions should be answered by choosing at least ONE question each from Section A and Section B.

The number of marks carried by each question is indicated against each.
Answers must be written only in ENGLISH.

(Symbols and abbreviations are as usual.)

If any data/value is to be assumed for answering a question, the same must be mentioned clearly.

SECTION A

1. Answer any *five* parts : *8×5=40*

(a) Discuss Lahiri's method of selecting samples for PPS scheme. Give a suitable example to select 4 units when N = 10, M = 60, by taking your choice of sample number and its corresponding size.

(b) Explain the concept of simple random sampling with and without replacement. In a simple random sampling, show that s^2 is an unbiased estimator of S^2.

(c) For the ratio estimator of the population mean $\overline{Y}$, obtain the bias of $\overline{y}_r$ in terms of covariance. Further show that if the coefficient of variation of $\overline{x}$ is sufficiently small, the bias compared to standard error of $\overline{y}_r$ may be considered to be negligible.

(d) For symmetric BIBD, show that $|N| = \sqrt{r(r-\lambda)^{v-1}}$, where N is the incidence matrix of SBIBD.

(e) Discuss 3^2 factorial experiments. Explain a method to estimate all the main effects and interaction effects at single degree of freedom for 3^2 factorial experiment. Write its ANOVA table considering replication size 2.

(f) Construct a key block of a 3^4 confounded factorial experiment into a block of size 9 by confounding the interactions ABC and AB^2D. Write all its generalized confounded interactions. Construct its one more block.

2. (a) Let for a stratified random sampling

$$n_i = \frac{W_i S_i}{\sqrt{\mu_o c_i}}$$

where μ_0 is constant, c_i is the cost per unit in i^{th} stratum, $W_i = \frac{N_i}{N}$, S_i^2 is the mean square based on N_i units. Estimate the sample size n under optimum allocation for fixed cost c_0.

(b) In two stage sampling with equal first stage units, obtain the variance of the sample mean $\bar{\bar{y}}_2$.

(c) Discuss cluster sampling. Let $\bar{\bar{y}}$ be the sample mean based on a sample of nM elements. $\bar{y}$ is the sample mean based on a sample of nM elements drawn for SRS without replacement from NM elements in the population. Obtain Relative efficiency of $\bar{\bar{y}}$ with respect to $\bar{y}$ (without getting the derivation of $V(\bar{y})$ and $V(\bar{\bar{y}})$). Write its ANOVA table.

(d) Explain aligned sample and unaligned systematic sampling method. Give an example of each with $m = 3$, $l = 3$, $n = 3$ and $k = 4$, where nm units denotes systematic samples, provided units of population are arranged in the form of ml rows each containing nk units. *10×4=40*

3. (a) Discuss a method of construction of BIBD with parameters $v = b = 11$, $r = k = 6$ and $\lambda = 3$.

(b) Explain Randomized block design. Obtain non-zero eigen value of C matrix of RBD. Estimate t_i ($i = 1, 2, ..., v$) using non-zero eigen value of C matrix.

(c) Discuss missing plot techniques in a RBD. Suppose one observation, say, t_i is missing in one block of a RBD, derive a method to estimate that missing value.

(d) Discuss symmetrical BIBD. For a SBIBD, show that any two blocks have exactly λ treatments in common. *10×4=40*

4. (a) Explain difference estimator of population mean. Obtain its mean and variance. Hence obtain the linear regression estimator of $\overline{Y}$.

(b) In a PPSWOR scheme, let $z_i = \frac{y_i}{NP_i}$. Obtain $E(\bar{z})$ and $V(\bar{z})$.

(c) Discuss group divisible PBIB design of two associate classes. Construct a PBIB design with parameters $v = 8 = b$, $r = 3 = k$, $\lambda_1 = 0$ and $\lambda_2 = 1$. Further obtain the value of m and n.

(d) Explain the layout of split plan designs. Write its model and assumptions. Give ANOVA table of sub-plot observations (only df and sum of squares). *10×4=40*

SECTION B

5. Attempt any ***five*** parts : *8×5=40*

(a) Explain moving average [m, p]. Let $U_t = a_0 + a_1 t$, with p = 1 and m = 2k + 1, obtain the coefficient c_j of U_t and [m, p].

(b) Let $U_t = a\xi + e_t$, $-\infty < t < \infty$, where e_t's are i.i.d. with $E(e_t) = 0$ and $V(e_t) = 1$. Show that the process is stationary with correlation.

(c) For the following table :

Commodity

	A	B
p_0	1	1
q_0	10	5
p_1	2	X
q_1	5	2

where p and q stand for price and quantity for 0 and 1 time periods respectively. Find the value of X if the ratio between Laspeyres' (L) and Paasche's (P) index number is

L : P : : 28 : 27

(d) For the general linear model $Y = X\beta + e$, following observations

$\Sigma x_1^2 = 50$, $\Sigma x_2^2 = 960$, $\Sigma x_1 x_2 = -60$,

$\Sigma x_1 y = 30$, $\Sigma x_2 y = 40$, $\overline{X}_1 = 5$, $\overline{X}_2 = 6$ and

$\overline{Y} = 5$

are given, where lower case letters $y_j\, x_j$ (j = 1, 2) are given as $y = Y - \overline{Y}$ and $x_j = X_j - \overline{X}_j$.

Obtain OLS estimator.

(e) Discuss the effect of imperfect multicollinearity on tests and errors. Consider a model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (e - \bar{e}) \text{ with}$$
$$\sum x_1 = \sum x_2 = 0 \text{ and } x_2 = \theta x_1 + v \text{ with}$$
$$\sum x_1^2 = \sum x_2^2 = 1, \quad \sum x_1 x_2 = \theta, \quad \sum v = 0 \text{ and}$$
$$\sum x_1 v = 0.$$

Do you think multicollinearity is present in the model ? If yes, give reasons. Further show that as θ increases, $V(\hat{\beta}_i)$ also increases.

(f) Explain demand function. Discuss Pigou's method of deriving demand curves from time series data. Indicate the assumptions made. Give a criticism of the method.

6. (a) Give the various steps for finding the variance of the random component using variate difference method. Is F-test used for testing the significance of homogeneity of two successive estimates of variance ? If yes, ok, otherwise which test can be used ? Discuss.

(b) Explain the errors in the measurement of price and quantity index number.

(c) Discuss second order autoregressive series. For this series, obtain complementary function (CF) only.

(d) Explain curve of concentration. Obtain income concentration for Pareto's law of income distribution. *10×4=40*

7. (a) For generalized least square linear model $\underline{Y} = X\underline{\beta} + \underline{e}$, write assumptions. Obtain $V(\hat{\beta})$ and estimate $\hat{\sigma}^2$ provided given another linear model is $Y^* = X^* \beta + e^*$, where $X^* = T^{-1} X$, $Y^* = T^{-1} Y$, $e^* = T^{-1} e$, $\Omega = TT'$ and $e^* \sim N(0, \sigma^2)$.

(b) If $U = c\, x^{\alpha} y^{\beta}$ is an individual's utility function of two goods, show that the demand for the goods is

$$x = \frac{\alpha}{\alpha + \beta} \frac{\mu}{p_x} \quad \text{and} \quad y = \frac{\beta}{\alpha + \beta} \frac{\mu}{p_y}$$

where p_x and p_y are the fixed price and μ be the individual fixed income.

(c) For which simultaneous equation, is indirect least square method of estimation used ? Discuss a method to estimate parameters using indirect least square estimator.

(d) Discuss economic forecasting of one single future observation on Y_T. *10×4=40*

8. (a) For the auto-regressive scheme

$$U_{t+2} + aU_{t+1} + bU_t = e_{t+2},$$

show that if e is a random variable and the series is long, then

$$\frac{\text{Var}(U)}{\text{Var}(e)} = \frac{(1+b)}{(1-b)[(1+b)^2 - a^2]}$$

and hence show that, variance of the generated series may be much greater than that of e itself.

(b) Explain the formulation of the problem of distribution of income. Further discuss its mathematical formulation in order to find number of persons with income x or more.

(c) Discuss the method of estimating the parameters of simultaneous equation model using two stage least square method.

(d) Explain rank condition of identification problem of simultaneous equation model. *10×4=40*