IES/ISS Exam, 2013

A-HRR-TUBC

STATISTICS III

Time Allowed : Three Hours	Maximum Marks : 200

INSTRUCTIONS

Candidates should attempt FIVE questions in ALL including Question Nos. 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.

All parts and sub-parts of a question are to be attempted together in the answer book.

Attempts of a part/question shall be counted in chronological order. Unless struck off, attempt of a part/question shall be counted even if attempted partly. Any page or portion of the page left blank in the answer book must be clearly struck off.

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 of the following : *8×5=40*

(a) In simple random sampling without replacement (SRSWOR), the sample mean $\bar{y}$ is an unbiased estimator of population mean $\overline{Y}$. Obtain the variance of the estimator of the population total Y. *8*

(b) Describe the principles of stratification. Explain how the allocation of sample size in different strata is made. *8*

(c) Discuss a procedure, using Sen – Midzuno scheme, to obtain the inclusion probability of n units. *8*

(d) Identify the design given below, whose blocks are : (1, 2, 3, 4); (2, 3, 4, 1) and (4, 1, 2, 3).

Write its parameters. Suppose a treatment '3' belonging to 3^{rd} row and 4^{th} column is missing. How can you estimate that missing treatment ? Estimate that value. Write degrees of freedom for error and total variation. *8*

(e) Construct a key block of 2^5 confounded factorial experiment conducted into a block of size 8. Discuss how to construct other blocks. Write its confounded interactions. *8*

(f) Discuss the concept of Balanced incomplete block design (BIBD). How does this design differ from randomized block design ? For a BIBD with parameters v, b, r, k and λ, show that $b \geq v$. *8*

2. (a) Give the names of various procedures for selecting a sample using probability proportional to size (PPS) scheme. Obtain an unbiased estimator of the population mean under PPS. *10*

(b) Show that ratio estimator is unbiased if the line of regression of Y on X is a straight line and passing through origin. *10*

(c) Discuss two-stage sampling. Let n first stage units be selected from a population of N units and m second stage units from each chosen first stage unit be selected using simple random sampling without replacement. Obtain the variance of sample mean $\bar{y}$. *10*

(d) Explain circular systematic sampling. Show that minimum value of ρ is $\frac{-1}{(n-1)}$, when $V_{sy}(\bar{y}) = 0$, where ρ is intraclass correlation and $\bar{y}$ is the sample mean. *10*

3. (a) Explain the need of Factorial experiments. Develop a method to estimate all main effects and interaction effects of 2^3-factorial experiment. Give its ANOVA table. *10*

(b) In a two-way classified data, the incidence matrix of the design is given below. Analyze the data.

Block

Treatment	1	2	3	4	5	6	7
1	1	0	0	0	1	0	1
2	1	1	0	0	0	1	0
3	0	1	1	0	0	0	1
4	1	0	1	1	0	0	0
5	0	1	0	1	1	0	0
6	0	0	1	0	1	1	0
7	0	0	0	1	0	1	1

N =

Show that the variance of the estimates of all treatment contrasts $\left(\hat{t}_i - \hat{t}_j\right)$ is constant for all $i \neq j = 1, 2, ..., v$. *10*

(c) What is a symmetrical BIBD ? If N is the incidence matrix of SBIBD and λ is the number of treatments common in any two blocks, then establish a relation between N and λ. *10*

(d) For a PBIB design with parameters v, b, r, k, λ_i, n_i, show that $\sum_{i=1}^{m} n_i \lambda_i = r(k-1)$. *10*

4. (a) In simple random sampling without replacement, obtain the sampling variance of the regression estimator $\bar{y}_l$. Show that the estimator is superior to SRSWOR. *10*

(b) Describe a procedure for systematic sampling under linear trend in the population. Obtain the expressions of population mean and population variance. *10*

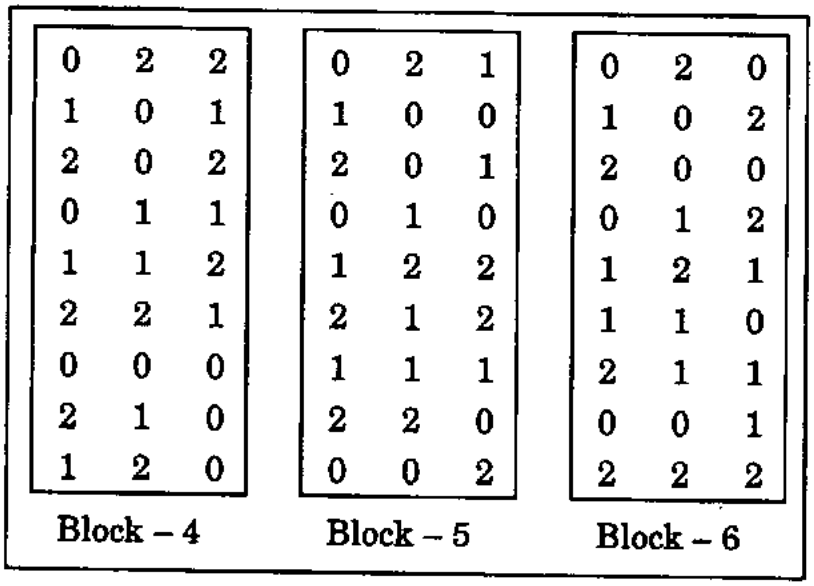
(c) Explain split-plot designs. Considering main plot treatment α, sub-plot treatment β and the number of replications in the RBD for main plot treatments be r, write the ANOVA table of main plot totals. *10*

(d) For the following experiment :

Replication – 1

Block – 1			Block – 2			Block – 3		
0	2	2	1	1	1	0	2	1
0	0	0	0	0	1	0	0	2
0	1	1	2	0	2	1	1	2
2	2	0	0	2	0	2	2	2
1	2	1	2	2	1	2	0	0
2	1	2	1	0	0	1	2	0
1	0	2	0	1	2	1	0	1
2	0	1	1	2	2	2	1	1
1	1	0	2	1	0	0	1	0

Replication – 2



Block – 4		
0	2	2
1	0	1
2	0	2
0	1	1
1	1	2
2	2	1
0	0	0
2	1	0
1	2	0

Block – 5		
0	2	1
1	0	0
2	0	1
0	1	0
1	2	2
2	1	2
1	1	1
2	2	0
0	0	2

Block – 6		
0	2	0
1	0	2
2	0	0
0	1	2
1	2	1
1	1	0
2	1	1
0	0	1
2	2	2

identify the experiment. Find the confounded experiment. Write the type of confounded interactions. Further write its generalized confounded interaction, if available. Give your comments. *10*

SECTION B

5. Answer any ***five*** parts of the following : *8×5=40*

(a) Discuss various steps of finding adjusted monthly indices of seasonal variations using link relative method. *8*

(b) Obtain the complementary function and particular integral of first order regressive model. Show that U_t is a moving average of random elements with weights 1, a, a^2, ... a^{t-1}. *8*

(c) Discuss simple aggregate method of Index number. Why is weighted average of price relative index preferred over simple aggregate method ? Show that weighted average of price relative index number is same as that of Paasche's index number. *8*

(d) Let under usual linear model set-up, $\hat{\beta}$ is OLS estimator given by $\underline{\hat{\beta}} = (X'X)^{-1} X' \underline{Y}$. Let $\hat{\hat{\beta}}$ be a new linear estimator obtained by augmenting p'y in the coefficient of Y. Under what conditions, $\hat{\hat{\beta}}$ is unbiased estimator of β ? Further obtain its variance. Mention under what condition $V(\hat{\beta})$ and $V(\hat{\hat{\beta}})$ will be same. *8*

(e) Consider a hypothetical data

$X_1 = (10, 15, 18, 24, 30)$; $X_2 = (50, 75, 90, 120, 150)$ and $X_3 = (52, 75, 97, 129, 152)$.

Find the type of multicollinearity between X_1 and X_2 and X_2 and X_3. Discuss how to tackle the multicollinearity problem using a prior estimator. *8*

(f) Explain heteroscedasticity. Show that under heteroscedastic model, OLS estimator is less efficient than the weighted least square estimator. *8*

6. (a) Why do the errors occur in the measurement of price and quantity index numbers ? Give the names of those errors. *10*

(b) Discuss the relevance of variate difference method in time series analysis data. Show that $V(\Delta^k U_t) = E[\Delta^k \varepsilon_t]^2$. *10*

(c) Explaining Pareto's law of income distribution, show that the elasticity decreases in the number of persons while passing to a higher income class. *10*

(d) Show that for a series generated by an m-points simple moving average of random components, the correlogram consists of a straight line joining the points (m, 0) and (0, 1) together with k-axis from the point (m, 0) onwards. *10*

7. (a) For a model $\underline{Y} = X\underline{\beta} + \underline{e}$, if $\underline{e} \sim N(0, \sigma^2 \Omega)$, where Ω is a positive semi-definite matrix, can we use OLS method of estimation ? If not, why ? Using which method of estimation, can $\hat{\beta}$ be estimated ? Obtain its variance. *10*

(b) Consider a simple regression model $y_t = \beta X_t + U_t$. Let $U_t = \rho U_{t-1} + e_t$, with $E[e_t] = 0$, $E[e_t, e_{t-1}] = 0$, $V(e_t) = \sigma^2$. If U_t are AR(1) errors, show that $\rho_j = \rho^j$, where ρ is autocorrelation coefficient. *10*

(c) What is the rank condition of identification problem ? For the simultaneous equation model :

$$Y_1 = \Gamma_{21} Y_2 + \beta_{11} X_1 + e_1 \ldots\ldots\ldots\ldots (1)$$

$$Y_1 = \Gamma_{22} Y_2 + \beta_{22} X_2 + \beta_{32} X_3 + e_2 \ldots (2)$$

discuss a method to identify the equation (1) only. *10*

(d) Consider measurement errors in one variable are in the limit uncorrelated with measurement errors in other variable. Again measurement errors in both variables are uncorrelated with their true values, then show that under such situations, the OLS estimator $\hat{\beta}$ is inconsistent for β. *10*

8. (a) Discuss Engel's curve. Explain the Engel's curve for constant price of a demand function $d = f(\mu, p)$, where μ is national income and p is price. *10*

(b) Explain the following : *10*

(i) Weighted Aggregates method

(ii) Criteria of a good index number

(c) Explain the effect of Ridge estimator in multicollinearity. Show that Ridge estimator is biased. Obtain its bias and variance. *10*

(d) Explain Keysian income determination simultaneous equation model. Show that simultaneous equation model violates the assumptions of OLS estimator. *10*