IES/ISS EXAM, 2015

Sl. No. 0007310

A-LVV-O-UVB

STATISTICS

Paper II

Time Allowed : Three Hours	Maximum Marks : 200

INSTRUCTIONS

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

There are EIGHT questions divided under TWO sections.

Candidate has to attempt SIX questions in all.

Question No. 1 and 5 are compulsory and out of the remaining, FOUR are to be attempted choosing at least TWO from each Section.

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.

Candidates should attempt questions/parts as per the instructions given in the Section.

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

Attempts of questions shall be counted in chronological 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 answer book must be clearly struck off.

Answers must be written in ENGLISH only.

Section – A

1. Answer ***all*** of the following : 5×8=40

(a) Let X be a random variable having probability density function

$$f_X(x,\ \theta) = \frac{1}{\theta}\ ,\quad 4 < x < 4+\theta$$
$$= 0\ ,\quad \text{otherwise.}$$

Find an unbiased estimator of θ^2 based on a sample of size 1.

(b) Consider the matrix

$$X' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 \end{pmatrix}$$

Find a basis of the column space (X) and obtain a g-inverse of $X'X$.

(c) Consider the Gauss-Markov model

$$E(y_1) = 2\beta_1 + \beta_2$$
$$E(y_2) = \beta_1 - \beta_2$$
$$E(y_3) = \beta_1 + \alpha\beta_2,$$

with usual assumptions. Determine α so that the best linear unbiased estimators (BLUEs) of β_1 and β_2 are uncorrelated.

(d) Let y_1, y_2, ... y_n be a random sample from $N(\mu, \sigma^2)$ where μ and σ^2 are both unknown. Obtain a confidence interval for μ with confidence coefficient $(1-\alpha)$.

(e) Show that, if T is a sufficient statistic for θ, any solution of the likelihood equation will be the function of this statistic. State true or false. A minimal sufficient statistic is always sufficient.

(f) Stating the regularity conditions, give the Cramer-Rao lower bound for the variance of an unbiased estimator of a parameter. Give an example, each, of a situation where the regularity conditions (i) does not hold (ii) holds.

(g) Show, in the context of Gauss Markov model $(Y, X\beta, \sigma^2 I)$, that the projection of an unbiased estimator of an estimable linear parametric function $\lambda'\beta$ is also an unbiased estimator of $\lambda'\beta$.

(h) Consider an unbalanced one way fixed effects model

$$y_{ij} = \mu + \alpha_i + e_{ij},\ i = 1 \ldots k,\ j = 1 \ldots n_i$$

where $E(e_{ij}) = 0$ for all i, j. Obtain the constraint needed for estimability of the parameters. Discuss the analysis of variance in a situation where the model is applied.

2. Answer ***all*** of the following : 10×3=30

(a) Consider the linear model $\underset{\sim}{y} = X\beta + \underset{\sim}{e}$, $E(\underset{\sim}{e}) = 0$, disp $(\underset{\sim}{e}) = \sigma^2 V$, V known and positive definite and σ^2 is unknown. Show that the estimator $l'\underset{\sim}{y}$ is the BLUE for $E(l'\underset{\sim}{y})$ iff $l'\underset{\sim}{y}$ is uncorrelated with all unbiased estimators of zero.

(b) Let $X_1, X_2 \ldots X_n$ be a random sample from the probability distribution with density

$$f_X(x, \theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad 0 < x < \infty$$
$$= 0 \quad , \text{ otherwise}$$

where $0 < \theta < \infty$. Show that $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ is a minimum variance bound estimator and has variance $\frac{\sigma^2}{n}$.

(c) Write Tukey's non-additive model for a set of two-way classified data with one observation per cell. Suggest an estimator of the non-additivity parameter and find the distribution under additivity. Derive a test for additivity for such a model.

3. Answer ***all*** of the following : 10×3=30

(a) The observations

3·9, 2·4, 1·8, 3·5, 2·4, 2·7, 2·5, 2·1, 3·0, 3·6, 3·6, 1·8, 2·0, 4·0, 1·5

are a random sample from a rectangular population with pdf

$$f(x;\, a,\, b) = \begin{cases} \dfrac{1}{b-a} & , a \leq x \leq b \\ 0 & , \text{otherwise} \end{cases}$$

Estimate the parameters by the method of moments.

(b) Explain how the Rao-Blackwell theorem helps one to find a uniformly minimum variance unbiased estimator (UMVUE) of an unknown parameter. What is the relevance of the Lehman-Scheffe theorem in this scenario ? If $X_1, X_2 \ldots X_n$ are Bin $(1, p)$ variates, find the UMVUE of p.

(c) For a completely balanced two-way random effects model, find unbiased estimators of different variance components. Explain how to obtain the variance of an unbiased estimator of any one variance component corresponding to a main effect. Find an approximate $100(1-\alpha)\%$ confidence Interval to any arbitrary linear function of variance components.

4. Answer *all* of the following : 10×3=30

(a) Consider the following cross classified model without replication.

$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij},\ E(e_{ij}) = 0,\ i = 1,\ 2.$
$j = 1,\ 2,\ 3.$

Where $\underset{\sim}{y} = (y_{11}\ \ y_{12}\ \ y_{13}\ \ y_{21}\ \ y_{22}\ \ y_{23})'$

$\beta = (\mu,\ \ \alpha_1,\ \ \alpha_2,\ \ \beta_1,\ \ \beta_2,\ \ \beta_3)$ and

$$X = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Write down the normal equations and find all solutions. Show that $\alpha_1 - \alpha_2$ and $\beta_1 - 2\beta_2 + \beta_3$ are estimable and give their least squares estimators.

(b) Obtain the sufficient statistics for the following distributions :

(i) $f(x,\ \theta) = \frac{1}{\theta}\, e^{-\frac{x}{\theta}},\ 0 < x < \infty;\ \theta > 0.$

(ii) $f(x,\ \theta) = (1-\theta)^x \theta,\ x = 0,\ 1,\ 2\ \ldots\ ;$
$0 < \theta < 1.$

(c) Let $X_1, X_2 \ldots X_n$ be a random sample from the binomial distribution with probability mass function.

$$f(x, \theta) = \theta^x (1-\theta)^{1-x} \quad x = 0, 1;\ 0 < \theta < 1$$
$$= 0 \qquad \text{otherwise}$$

Examine whether the statistic $T = \sum_{1}^{n} X_i$ is complete for this distribution.

Section – B

5. Answer ***all*** of the following : 5×8=40

(a) Let $X = (X_1, X_2, X_3)'$ be distributed as $N_3(\mu, \Sigma)$ where $\mu' = (2, -3, 1)$ and

$$\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

(i) Find the distribution of $3X_1 - 2X_2 + X_3$.

(ii) Find a 2×1 vector $\underset{\sim}{a}$ such that X_2 and $X_2 - a'\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ are independent.

(b) Find the maximum likelihood estimators of the 2×1 mean vector $\underset{\sim}{\mu}$ and the 2×2 covariance matrix Σ based on the random sample

$$X = \begin{bmatrix} 3 & 4 & 5 & 4 \\ 6 & 4 & 7 & 7 \end{bmatrix}$$

from a bivariate normal population.

(c) Explain the notion of unbiasedness with regard to a test of a hypothesis. Examine the validity of the statement. A Most Powerful (MP) test is invariably unbiased.

(d) To test the hypothesis $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ for the distribution

$$f(x, \theta) = \theta^x (1-\theta)^{1-x}, \; x = 0, 1; \; 0 < \theta < 1$$
$$= 0 \qquad \text{otherwise,}$$

develop the sequential probability ratio test.

(e) Distinguish between the single sampling plan and double sampling plan. Discuss how the O.C curves can be used for comparing two sampling plans.

(f) Let $\underset{\sim}{X} = (X_1, X_2, X_3)'$ be distributed as $N_3(\underset{\sim}{\mu}, \Sigma)$ where $\underset{\sim}{\mu} = (10, -7, 2)'$ and

$$\Sigma = \begin{pmatrix} 2 & -1 & 3 \\ -1 & 5 & -3 \\ 3 & -3 & 5 \end{pmatrix}.$$

Find the partial correlation between X_1 and X_2 given X_3.

(g) Show that T^2 statistic is invariant under changes in the units of measurements for a $p \times 1$ random vector X of the form $\underset{\sim}{Y} = C\underset{\sim}{X} + \underset{\sim}{d}$ where C is a $p \times p$ nonsingular matrix, $\underset{\sim}{d}$ is a $p \times 1$ vector.

(h) Distinguish between control chart for variables and control chart for attributes. Give an example for each. When will you say that a process is in control? Suppose that all the points in a control chart falls above the central line. What will be your conclusion?

6. Answer ***all*** of the following : 10×3=30

(a) Derive the likelihood ratio test for comparing the means of k independent homoscedastic normal populations.

(b) Let $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ be a random sample from an $N_p(\underline{\mu}, \Sigma)$ population with Σ, a positive definite matrix. Derive $100(1-\alpha)\%$ simultaneous confidence intervals for $\underline{l}'\underline{\mu}$ for all $\underline{l} \in \mathbb{R}^p \setminus \{\underline{0}\}$.

(c) Samples of size $n = 5$ are taken from a manufacturing process every hour. A quality characteristic is measured, and $\overline{X}$ and R are computed for each sample. After 25 samples have been analyzed, we have $\sum_{j=1}^{25} \overline{x}_j = 662 \cdot 50$ and $\sum_{i=1}^{25} R_i = 9 \cdot 00$. Assume that the quality characteristic is normally distributed.

(i) Find the control limits for the $\overline{X}$ and R charts.

(ii) Assume that both charts exhibit control, if specifications are $26\cdot40 \pm 0\cdot50$, estimate the fraction nonconforming. Express your answer in terms of CDF of $N(0, 1)$ random variable.

[For $n = 5$, $A_2 = 0\cdot577$, $A = 1\cdot342$, $A_3 = 1\cdot427$, $D_1 = 0$, $D_2 = 4\cdot918$, $D_3 = 0$, $D_4 = 2\cdot115$ and $d_2 = 2\cdot326$]

7. Answer ***all*** of the following : 10×3=30

(a) Find a Most Powerful (MP) test for testing the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against the simple alternative $H_1 : \sigma^2 = \sigma_1^2$ based on n random observations from $N(\mu, \sigma^2)$ where μ is known. Show that this MP test is UMP (uniformly most powerful).

(b) (i) Let A_i be distributed as Wishart $W_{m_i}(A_i \mid \Sigma)$, $i = 1, 2$ and A_1, A_2 be independent. Show that $A_1 + A_2$ is distributed as $W_{m_1 + m_2}(A_1 + A_2 \mid \Sigma)$.

(ii) If A is distributed as $W_m(A \mid \Sigma)$, then CAC' is distributed as $W_m(CAC' \mid C\Sigma C')$ where C is a nonsingular matrix of order m.

(c) (i) Explain the terms average outgoing quality (AOQ) and average total inspection (ATI).

(ii) Let $N = 10{,}000$, $n = 89$ and $C = 2$. Determine P_a, the probability of acceptance and use it to determine AOQ and ATI.

8. Answer ***all*** of the following : 10×3=30

(a) For a sequential probability ratio test of strength (α, β) and stopping bounds A and B $(B < A)$, show that

$$A \leq \frac{1-\beta}{\alpha} \text{ and } B \geq \frac{\beta}{1-\alpha}.$$

(b) Let

$$\Sigma = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 5 \end{pmatrix}$$

Determine

(i) the principal components y_1, y_2 and y_3.

(ii) the proportion of variance explained by each one of them.

(iii) correlation between the first principal component y_1 and the third original random variable.

(c) What is meant by acceptance sampling by attributes ? Outlining the criteria for the goodness of sampling inspection plan, give briefly the steps in the Dodge inspection plan.