IES/ISS Exam, 2013

A-HRR-TUBB

STATISTICS – II

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.

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.)

Any essential data assumed by candidates for answering questions must be clearly stated.

(Contd.)

Section 'A'

1. Attempt any *five* of the following : 8×5=40

(a) Assume a full rank model $Y = X\beta + \in$, where $\in \sim N_n\left(0,\ \sigma^2 I_n\right)$. $X(n\times p)$ is the design matrix and $\underset{\sim}{\beta}(p\times 1)$ is the parameter vector. Construct a test for $H_0 : \underset{\sim}{\beta} = \underset{\sim}{0}$ against $H_1 : \underset{\sim}{\beta} \neq \underset{\sim}{0}$.

(b) Consider the linear model $Y = X\beta + \in$, where X is $(n\times p)$, $\underset{\sim}{\beta}$ is $(p\times 1)$ and $\in \sim N_n\left(0,\ \sigma^2 I_n\right)$. State and prove a necessary and sufficient condition for any linear parametric function $l'\beta(\underset{\sim}{l} \neq 0)$ to be estimable.

(c) Define Fisher Information. Illustrate its application in estimation theory. Derive Fisher Information about the parameter θ in a random sample of size n from an exponential distribution with parameter θ.

(d) Let $g(\theta)$ be a real valued estimable function of an unknown parameter θ and T be a complete sufficient statistic. Prove that an unbiased estimator of $g(\theta)$ based on T is enough for finding the UMVUE of $g(\theta)$.

(e) Show that the S.S. due to a testable hypothesis in a linear model, can be expressed as the difference between conditional and unconditional S.S. due to error.

(f) State and prove Rao-Blackwell theorem.

2. (a) In a general linear model setup $(Y, X\beta, V)$, derive the covariance between the BLUEs of two estimable linear parametric functions $p'\beta$ and $q'\beta$. 10

(b) Consider a two way fixed effects model with m observations per cell :

$$y_{ijk} = \mu_{ij} + e_{ijk};\ i = 1....p;\ j = 1....q;\ k = 1....m$$

with

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

where the parameters are defined with respect to a system of weights; $\{e_{ijk}\}$ are random errors. Show that the above representation of μ_{ij} in terms of μ, α_i, β_j and γ_{ij} is unique for a given system of weights. 10

(c) Define (i) Efficiency and (ii) UMVUE. Let T_0 and T_1 be two unbiased estimators for θ with efficiency e_θ. Then prove that no linear combinations of T_0 and T_1 can be a MVUE. 10

(d) Distinguish between confidence interval and tolerance limits.

Based on a random sample of size n from a Gamma (α, β), obtain the moment estimators of α and β. Suggest a large sample confidence interval for α when $\beta = 1$. 10

3. (a) Consider the linear model

$$E(Y_i) = \mu_i - \mu_{i+1},\ i = 1 \ldots\ldots n-1$$
$$E(Y_n) = \mu_n - \mu_1$$
$$E\big(Y_i - E(Y_i)\big)\big(Y_j - E(Y_j)\big) = \sigma^2 \text{ if } i = j$$
$$= 0 \quad i \neq j$$

(i) Show that $\overline{Y} = \frac{1}{n}\left(Y_1 + \ldots\ldots + Y_n\right)$ is the only error function.

(ii) Find the BLUE of $\mu_j - \mu_{j+1}$ for $j = 1, 2, \ldots., n$ with $\mu_{n+1} = \mu_1$.

(iii) Find the BLUE of $\sum_{j=1}^{n} t_j \mu_j$ with

$$\sum_{j=1}^{n} t_j = 0$$

(iv) Find an estimate of σ^2. 10

(b) 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 its distribution under additivity. Construct a test for additivity for such a model. 10

(c) State and Prove Cramer-Rao inequality with usual notations. 10

(d) State the properties of a maximum likelihood estimator (MLE). Based on a random sample of size n from the distribution with pdf

$$f(x;\, \alpha,\, \beta) = \frac{1}{\beta - \alpha} \qquad \alpha \leq x \leq \beta$$

obtain a MLE of α and β. 10

4. (a) Discuss the following statements :

(i) A statistic can be minimal sufficient without being complete

(ii) A maximum likelihood estimate is not necessarily unique 10

(b) For a two-way ANOVA with K treatments and n blocks, show that

$$E\left[\frac{K\sum_{j=1}^{n}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2}{n-1}\right]=\sigma^2+\frac{K\sum_{j=1}^{n}\beta_j^2}{n-1}$$

where β_j is the jth block effect. 10

(c) If $X_1, X_2, \ldots, X_n$ constitute a random sample from uniform $(0, \theta)$, then show that

$\left(\frac{n+1}{n}\right)X_{(n)}$ is an unbiased estimator of θ

where $X_{(n)}$ denotes the largest observation. Also show that

(i) $2\overline{X}$ is unbiased for θ

(ii) Compare the efficiency of these two estimators. 10

(d) Define a family of uniformly most accurate confidence sets for an unknown parameter θ, with confidence level $(1-\alpha)$ where $0 < \alpha < 1$. Give an example to show that $(1-\alpha)$ confidence interval is not unique. 10

Section 'B'

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

(a) Define a similar test. Obtain a similar test for $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 < \sigma_0^2$ when the samples are drawn from $N(\mu, \sigma^2)$, where μ is unknown.

(b) State monotone likelihood ratio property of a family of distributions. Use the property to test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ for a random variable having a distribution function $F(x; \theta)$, $\theta \varepsilon \Omega$.

(c) Define :

(i) randomized decision rules

(ii) admissibility of a decision rule

(iii) Complete class of decision rules

(iv) Minimax decision rules

(d) Let X be a p-dimensional random vector with mean μ and dispersion Σ of rank $r(\leq p)$. Show that with probability 1, $X = \mu + LY$, where $L(p \times r)$ is of rank r. Hence discuss the possibility when $r = 1$.

(e) Distinguish between defects and defectives. Discuss the control charts for defects and defectives clearly mentioning their importance and specify their control limits (no proof is required).

(f) Let $X(p \times 1)$ be a random vector with dispersion $\Sigma(> 0)$. Find the principal components of X and their variances when

(i) $\Sigma = \sigma^2 I_p$ and

(ii) $\Sigma = \sigma^2 \left[(1-\rho) I_p + \rho \underset{\sim}{1}_p \underset{\sim}{1}'_p \right]$

where $\underset{\sim}{1}_p$ is a $p \times 1$ vector with unit elements and σ^2, $\rho(> 0)$ are known constants.

6. (a) Define a simple and a composite hypothesis. State and prove (necessary part only) Neyman Pearson Fundamental lemma for testing a simple hypothesis against one sided alternatives. 10

(b) In estimating the parameter p of a Binomial $(2, p)$, random variable Y, find, under squared error loss, the risk when (i) $d_1(y) = \frac{y}{2}$ and (ii) $d_2(y) = \frac{y+1}{y}$ where $d_1(y)$ and $d_2(y)$ are the estimators (decision rules) based on $Y = y$. 10

(c) In an examination each student had to answer p questions carrying equal marks which were examined independently by two examiners. The problem is to test the hypothesis that there is no significant difference between the average marks awarded by the two examiners on each question on the basis of marks obtained by a random sample of N students, $2(N-1) \geq p$. Assuming that the marks distribution for the two examiners are multivariate normal with a common unknown positive definite dispersion matrix, suggest and justify a sized unbiased test for the problem, $0 < \alpha < 1$. 10

(d) Obtain the control limits for the following data providing the number of coloured threads (considered as defects) in 15 pieces of cloth and comment : 10

7, 12, 3, 20, 5, 4, 3, 10, 8, 0, 9, 6, 21, 7, 20.

7. (a) Define size of a test. Show that the power of a Neyman-Pearson test exceeds its size.

The specifications for a certain kind of wire has mean breaking strength of 185 pounds. If five pieces randomly selected from different rolls have breaking strength of 171·6, 191·8, 178·3, 184·9 and 189·1 pounds, test the null hypothesis $H_0 : \mu = 185$ against $H_1 : \mu < 185$ at 5% level of significance.

[Given $t_{\cdot 05, 4} = 2 \cdot 132$] 10

(b) Discuss Fisher-Behren's problem for testing two multivariate population means. 10

(c) Consider the problem of classifying a random vector into one of the two populations $N_p(\mu_1, \Sigma)$ and $N_p(\mu_2, \Sigma)$ where μ_1, μ_2, Σ (positive definite) are means and dispersion matrix and all the parameters are known. Assuming a simple loss function, derive the Bayes rule for the problem and show that the rule is admissible. 10

(d) What is Statistical Quality Control and what is Statistical Process Control ? Discuss various types of variations encountered in a process and their methods of eliminations. 10

8. (a) Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables having pmf

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

Propose a SPRT for testing $H_0 : p = p_0$ against $H_1 : p = p_1$. Also find the ASN functions of the test. 10

(b) Discuss the general idea of classifications problem. Also discuss the procedure for classifying into one of two populations with known probability distribution. 10

(c) For the problem of estimating μ on the basis of a random sample from $N(\mu, 1)$ with absolute error loss $(-1 \leq \mu \leq 1)$, show that the sample mean is an inadmissible estimator. 10

(d) Write short notes on :

(i) Canonical Correlation Analysis

(ii) Bootstrapping the standard error of a trimmed mean. 10
