1.ES./ISS-09

1382

C-JTT-J-TB

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.

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.

https://bdust.com

SECTION A

1. Attempt any *five* sub-parts.

(a) Let A be any $n \times m$ matrix and K be any $m \times m$ non-singular matrix and let $B = AK$. Show that

$$BB^{-} = AA^{-}$$

where A^{-} denotes g-inverse of matrix A. *8*

(b) Let $l_1' \hat{\beta}$ and $l_2' \hat{\beta}$ be the least squares estimators of two estimable linear functions $l_1' \beta$ and $l_2' \beta$ respectively. Prove that

$$\text{Var}(l_1' \hat{\beta}) = \sigma^2 \, l_1' C^- l_1,$$

$$\text{Var}(l_2' \hat{\beta}) = \sigma^2 \, l_2' C^- l_2,$$

$$\text{Cov}(l_1' \hat{\beta}, \, l_2' \hat{\beta}) = \sigma^2 \, l_1' C^- l_2,$$

where C^- denotes a g-inverse of $C = A'A$. 8

(c) Consider three independent random variables y_1, y_2, y_3 having a common variance σ^2 and expectations $E(y_1) = \beta_1 + \beta_2$, $E(y_2) = \beta_1 + \beta_3$, $E(y_3) = \beta_3 - \beta_2$. Find the BLUE of the function $\beta_1 - \beta_2 + 2\beta_3$. 8

https://bdust.com

(d) Let $x_1, x_2, \ldots, x_n$ be independent observations from a rectangular population in the range 0 to θ. Obtain the best estimate of θ, in the sense of its important properties. 8

(e) Let $x_1, x_2, \ldots, x_n$ be a random sample drawn from the population

$$f_\theta(x) = \begin{cases} \theta, & \text{if } x = 1 \\ \dfrac{(1-\theta)}{(\lambda-1)}, & \text{if } x = 2, 3, \ldots, \lambda \\ 0, & \text{otherwise,} \end{cases}$$

where θ and λ are unknown.

Find maximum likelihood estimators of θ and λ. 8

(f) Explain the concept of completeness of a statistic. Show that

$$T = \sum_{i} X_i$$

is a complete sufficient statistic for the parameter θ in a random sample $X_1, X_2, \dots, X_n$ drawn from the population

$$f_\theta(x) = \theta^x (1 - \theta)^{1 - x}, \quad x = 0, 1.$$ *8*

2. (a) Prove that a generalized inverse of any matrix always exists, but not unique. *10*

(b) Let S_L denote the sum of squares with r degrees of freedom due to the set L_y of linear functions $l'y$, where $y = A\beta + \varepsilon$, $E(\varepsilon) = 0$, $D(\varepsilon) = \sigma^2 I$.

Show that S_L/σ^2 is distributed as non-central chi-square with r degrees of freedom and non-centrality parameter $\delta = S_{L}, \eta/\sigma^2$, where

$$S_{L, \eta} = \sum_{i=1}^{r} (l_i' \eta)^2 .$$ *10*

(c) Define Minimum Variance Bound (MVB) estimator.

Let $x_1, x_2, \dots, x_n$ be a random sample from a population with pdf

$$f(x, \theta) = \begin{cases} \theta(1+x)^{-(1+\theta)}, & x > 0, \theta > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Obtain MVB estimator of $1/\theta$ and e^θ. *10*

(d) When would you use minimum chi-square method of estimation ? Describe the methods of minimum chi-square and modified minimum chi-square and compare them with the maximum likelihood method. *10*

3. (a) Define Estimation Space and Error Space. When is a linear function $l'Y = \sum_{i} l_i Y_i$ said to belong to error ?

If $Y = C'Y$ be any linear function of observations $Y_1, Y_2, ..., Y_n$ and if C_0 be a resolved part of C along the estimation space, show that

$$E(Y) = E(C_0' Y).$$

10

(b) Consider the model

$$Y_1 = \alpha_1 + \varepsilon_1 ,$$

$$Y_2 = 2\alpha_1 - \alpha_2 + \varepsilon_2 ,$$

$$Y_3 = \alpha_1 + 2\alpha_2 + \varepsilon_3 ,$$

where $\varepsilon \sim N(0, 1)$.

Derive an appropriate test to test the hypothesis $H_0 : \alpha_1 = \alpha_2$. *10*

https://bdust.com

(c) Find the MLE of θ for the exponential family of distributions

$$f_\theta(x) = \exp\{C(\theta) + Q(\theta)\,T(x) + H(x)\}, \quad \theta \in \Theta \subset R^1$$

and prove that there exists a solution of the likelihood function, if and only if T(x) and $-C'(\theta)/Q'(\theta)$ have the same range of values. *10*

(d) Let $x_1, x_2, \ldots, x_n$ be a random sample from the uniform distribution $U(0, \theta)$. Obtain $100(1-\alpha)\%$ confidence interval for θ and explain how the length of the interval be made shortest. *10*

4. (a) For the linear model

$$Y = A\beta + \varepsilon$$

with $E(\varepsilon) = 0$

$$D(\varepsilon) = \sigma^2 I,$$

show that F-test can be applied to test the hypothesis $H_0 : \beta = 0$. *10*

(b) Let T_1 and T_2 be MVU estimators of a parameter θ. Show that no unbiased linear combination of T_1 and T_2 can be an MVU estimator of θ. *10*

(c) Discuss in detail, under appropriate assumptions, the analysis of a balanced two-way cross classification under random effects model. *10*

(d) Define uniformly most accurate (UMA) and uniformly most accurate unbiased (UMAU) confidence sets.

Using the relationship between confidence intervals and hypothesis testing, find UMAU confidence set of level $1 - \alpha$ for σ^2 in random sampling from $N(\theta, \sigma^2)$ population, θ being assumed known. *10*

SECTION B

5. Attempt any ***five*** sub-parts.

(a) Which of the two types of error is considered to be more serious in testing statistical hypotheses ? Explain with appropriate examples. How is the problem tackled ? State the major steps involved in testing statistical hypotheses. *8*

(b) Define unbiasedness of a test and explain why such a test is most desirable.
Let the power function of every test ϕ of $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta_1$ be continuous in θ. Then prove that a UMP α-similar test is UMP unbiased. *8*

(c) Explain the terms : stopping rule, decision rule, stopping region, stopping variable with examples in sequential test procedure. *8*

https://bdust.com

(d) Explain chance causes and assignable causes of variation in quality manufactured product. Assuming that the characteristic variable follows normal distribution with unknown mean and unknown s.d., give the central line and the control limits for $\overline{X}$ and R-charts. *8*

(e) Let $X \sim N_p(\mu, \Sigma)$ and suppose that X, μ and Σ are partitioned as

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

where $X^{(1)}$ has l elements, $l < p$.
Under what condition are $X^{(1)}$ and $X^{(2)}$ stochastically independent ? *8*

(f) Let $A \sim W(p, n, \Sigma)$ and let A and Σ be partitioned into q and $(p - q)$ rows and columns

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

where A_{11} and Σ_{11} are each $q \times q$ and $W(\circ)$ denotes Wishart statistic.

Prove that $A_{11} \sim W(q, n, \Sigma_{11})$. *8*

6. (a) Let $X_1, X_2, \ldots, X_n$ be independent random variables each distributed as exponential distribution with pdf

$$f(x, \theta) = \frac{1}{\lambda}\, e^{-(x-\theta)}, \quad x > \theta,$$

λ being known. Apply likelihood ratio method to test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$. *10*

https://bdust.com

(b) Define O.C. and ASN functions. Derive formulae for these functions of Wald's sequential probability ratio test. *10*

(c) What do you mean by ASN and ATI in acceptance sampling plans ? Show their relationship, if any. Obtain ASN and ATI functions of Double Sampling plan by attribute. *10*

(d) Explain Discriminant function and state its uses. Describe how discrimination between two multivariate normal populations can be tested. *10*

7. (a) Let $x_1, x_2, ..., x_n$ be a random sample from the population

$$f(x, \theta) = \frac{1}{\theta} e^{-x/\theta}, \quad 0 \le x \le \infty, \ \theta > 0.$$

Find α-size UMP critical region to test $H_0 : \theta \le \theta_0$ against $H_1 : \theta > \theta_0$. *10*

(b) Define Bayes decision rule.

Let $X \sim N(\mu, 1)$ and let a priori distribution of μ be N(0, 1). Assuming quadratic loss function

$$L(\mu, \delta) = [\mu - \delta(X)]^2,$$

find Bayes estimate of μ and the corresponding Bayes risk. *10*

(c) What is a sequential sampling plan ? When would you advocate the use of such a plan ? Determine the acceptance and rejection lines of a sequential plan for fraction defectives. *10*

(d) State the reproduction property of Wishart distribution.

Let $A \sim W(p, n, \Sigma)$. Consider the transformation $A = CBC'$, where C is a non-singular $p \times p$ matrix. Show that B is distributed as $W(p, n, \wedge)$, where $\wedge = C^{-1} \Sigma C'^{-1}$. *10*

8. (a) Let $X \sim N(0, 1)$ under H_0 and $X \sim C(1, 0)$ under H_1, where $C(\cdot)$ denotes Cauchy variate. Find α-size MP test for testing H_0 against H_1. *10*

(b) Let X_1, X_2, ..., X_n be independent and identically distributed random variables each having pdf

$$f_\theta(x) = \begin{cases} \theta^x (1-\theta)^{1-x}, & x = 0, 1 \\ 0, & \text{otherwise.} \end{cases}$$

Test $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ by SPRT. Also find the OC function of the test. *10*

(c) Explain how principal components can be extracted from a covariance matrix.

Let $X' = [X_1, X_2, ..., X_k]$ have a covariance matrix Σ, with eigenvalue-eigenvector pairs (λ_i, e_i), $i = 1, 2, ..., k$, where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k \geq 0$.

Let $Y_i = e_i' X$, $i = 1, 2, ..., k$, be the principal components. Show that

$$\sigma_{11} + \sigma_{22} + ... + \sigma_{kk} = \sum_{i=1}^{k} \text{var}(X_i) = \sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \text{var}(Y_i).$$

10

(d) Define canonical variables and canonical correlations. Show that the canonical correlation is invariant under scale transformation. *10*

https://bdust.com

https://bdust.com