3 (Sem–4/CBCS) STA HC 1

2023

STATISTICS

(Honours Core)

Paper : STA–HC–4016

(Statistical Inference)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following as directed : 1×7=7

   (a) Sample median is _____ estimator for the mean of normal population.

   (Fill in the blank)

   (b) Unbiased estimators are necessarily consistent.

   (State True ***or*** *False)*

(c) Area of critical region depends on

(i) number of observations

(ii) value of the statistic

(iii) size of type I error

(iv) size of type II error

(Choose the correct option)

(d) For a certain test if $\alpha = 0.05$, $\beta = 0.10$, then the power of the test is

(i) 0.95

(ii) 0.90

(iii) 0.05

(iv) 0.10

(Choose the correct option)

(e) Sample moments are _____ estimators of the corresponding population moments. *(Fill in the blank)*

(f) Suppose we put forward an interval which we expect to include the trees parameter value, then the process is called _____ estimation.

(Fill in the blank)

(g) The N-P lemma proceeds the best critical region for testing _____ hypothesis against _____ alternative hypothesis. *(Fill in the blanks)*

2. Answer the following questions : 2×4=8

(a) If $x_1, x_2, ...x_n$ is a random sample from a normal population $N(\mu,1)$, then show that $T = \sum_{i=1}^{n} x_i^2$ is an unbiased estimator of $\mu^2 + 1$.

(b) Find the maximum likelihood estimator of θ for the following probability distribution :

$f(x,\theta) = \theta e^{-\theta x}$, $x > 0, \theta > 0$

(c) State the Neyman-Pearson lemma.

(d) Give example of a maximum likelihood estimator which is not unbiased.

3. Answer ***any three*** questions from the following : 5×3=15

(a) Obtain the M.L.E. of α and β for the rectangular distribution

$$f(x:\alpha,\beta)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \alpha<x<p\\ u \quad , & \text{elsewhere}\end{cases}$$

(b) Show that, if a sufficient estimator exists, it is a function of the M.L.E.

(c) What is meant by statistical hypothesis? Explain the concept of type I and type II error with example. What is the power of a test ?

(d) Let X have the p.d.f. of the form

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

Find the most powerful test to test the simple hypothesis

$$H_0:\theta=1$$

against the alternative hypothesis

$$H_1:\theta=2$$

by means of a single observation X. What would be the size of type I and type II error, if you choose the interval

(i) $x\geq 0.05$

(ii) $x\geq 1.5$

as critical region ?

(e) Let $x_1, x_2, \ldots.. x_n$ be a random sample from a distribution with p.d.f.

$$f(x,\theta)=e^{-(x-\theta)},\ \theta<x<\infty$$
$$-\infty<\theta<\infty$$

Obtain a sufficient statistic for θ.

4. Answer ***any three*** questions from the following : 10×3=30

(a) What do you mean by MP and UMP tests ? Show that the most powerful test is necessarily unbiased.

(b) State the Cramer-Rao inequality. What are the conditions for equality sign in C-R inequality ? Show that,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

in random sampling from

$$f(x,\theta) = \begin{cases} \frac{1}{\theta}\exp\left(-x/\theta\right), & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

where, $0 < \theta < \infty$ is an MVB estimator of θ and has variance $\frac{\theta^2}{n}$.

(c) Define consistent estimator. State and prove the sufficient condition for consistency of an estimator.

(d) Show that with the help of example,

(i) an MLE is not unique;

(ii) an MLE may not exist.

(e) What is likelihood ratio test ? Show that likelihood ratio test for testing the variances of two normal population is the usual F-test.

(f) (i) Describe the method of moments for estimating parameter.

(ii) Show that in sampling from Cauchy population,

$$f(x,\theta) = \frac{1}{\pi\left[1+(x-\theta)^2\right]}, \quad -\infty < x < \infty \quad \theta > 0$$

is not sample mean, but sample median is a consistent estimator of θ.

———

3 (Sem–4/CBCS) STA HC 2

2023

STATISTICS

(Honours Core)

Paper : STA–HC–4026

(Linear Models)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following questions as directed : 1×7=7

   (a) In regression analysis, the variable that is being predicted is

   (i) the independent variable

   (ii) the dependent variable

   (iii) usually denoted by x

   (iv) usually denoted by r

   (Choose the correct option)

(b) The coefficient of determination is

(i) equal to zero

(ii) the ratio of explained and total variation

(iii) usually less than zero

(iv) 100% of $(1-r^2)$

(Choose the correct option)

(c) In least square estimation, which of the following is not a required assumption about the error term ?

(i) The expected value of the error term is one

(ii) The variance of the error term is the same for all values of x

(iii) The values of the error term are independent

(iv) The error term is normally distributed

(Choose the correct option)

(d) If the regression equation is equal to $Y = 23.6 - 54.2X$, then 23.6 is the _____ while -54.2 is the _____ of the regression line.

(i) slope, intercept

(ii) slope, regression coefficient

(iii) intercept, slope

(iv) radius, intercept

(Choose the correct option)

(e) Analysis of variance is a statistical method of comparing the _____ of several populations.

(i) standard deviations

(ii) variances

(iii) means

(iv) None of the above

(Choose the correct option)

(f) The sum of squares due to _____ measures the variability of the observed values around their respective treatment means

(i) treatment
(ii) error
(iii) interaction
(iv) total

(Choose the correct option)

(g) All OLS estimators are linear estimators.
(Write True or False)

2. Answer the following questions briefly : 2×4=8

(a) State some applications of the analysis of variance.

(b) What do you understand by components of variation ?

(c) Define estimability of linear parametric functions.

(d) Define R^2 in the context of a linear model.

3. Answer ***any three*** of the following questions : 5×3=15

(a) What is a linear model ? Discuss different types of linear models.

(b) A sample of 20 observations on X and Y gave the following data :

$$\sum Y = 21.9 \qquad \sum (Y - \bar{Y})^2 = 86.9$$
$$\sum X = 186.2 \qquad \sum (X - \bar{X}) = 215.4$$
$$\sum (X - \bar{X})(Y - \bar{Y}) = 106.4$$

Estimate the regression equation of Y on X and X on Y.

(c) Consider the one-way AOV model $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, for $i = 1, 2$ and $j = 1, 2, 3$

Examine if μ, α_1, α_2 are estimable without any constraints.

(d) In what respects do AOV, regression analysis and AOCOV differ ? Discuss briefly.

(e) Write a note on the technique of hypothesis testing in case of simple regression models.

4. Answer ***either*** *(a)* ***or*** *(b)* : 10

(a) State and prove the Gauss-Markov theorem.

(b) What is analysis of variance (AOV) ? What are the basic assumptions associated with it ? What are the remedies, if the assumptions are violated ?

5. Answer ***either*** *(a)* ***or*** *(b)* : 10

(a) Define a linear regression model. Write the basic assumptions of the linear model. Estimate the parameters of the model.

(b) Give linear model (fixed effect) for two-way classification (one observation per cell) and state its assumptions. Derive the analysis of variance of two-way classification through the method of least squares.

6. Answer ***either*** *(a)* ***or*** *(b)* : 10

(a) Using the following data

Y :	65	57	57	54	66
X :	26	13	16	−7	27

estimate the regression line $Y = \alpha + \beta X$, test the hypothesis that $\beta = 0$ against the alternative $\beta < 0$ at 5% level of significance, also construct 95% confidence interval for β.

(Given $t_{0.05,3} = 2.353$)

(b) Derive the 'analysis of covariance' for a one-way layout (with one consistent variable only).

———

3 (Sem–4/CBCS) STA HC 3

2023

STATISTICS

(Honours Core)

Paper : STA–HC–4036

(Statistical Quality Control)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following as directed : 1×7=7

   (a) The variation due to _____ factors is tolerable. *(Fill in the blank)*

   (b) Which one of the following is not a control chart for variable ?

   (i) $\overline{X}$ - chart
   (ii) σ - chart
   (iii) R - chart
   (iv) C - chart

   (Choose the correct option)

(c) In case of large samples _____ charts should preferably be used.

(Fill in the blank)

(d) In the construction of a control chart the extreme control limits are fixed at a distance of

(i) σ

(ii) 2σ

(iii) 3σ

(iv) 1.96σ

(Choose the correct option)

(e) Define OC curve.

(f) In SQC, when is $\overline{X}$ - chart used ?

(g) Control chart for fraction defective is a type of control chart for variables.

*(State True **or** False)*

2. Answer the following questions : 2×4=8

(a) What are the control limits for R- chart ?

(b) Mention *two* utilities of SQC technique in industrial production.

(c) Write down the control limits in P- chart if 50 mobiles are found defective in a consignment of 200 mobiles.

(d) Distinguish between product control and process control in SQC.

3. Answer ***any three*** of the following questions : 5×3=15

(a) Write a note on criterion for detecting lack of control in $\overline{X}$ - chart.

(b) Explain the basic principles underlying the construction of control charts bringing out the difference between 'natural tolerance limits' and 'specification limits'.

(c) Explain in brief the purpose and advantages of C - chart.

(d) Explain briefly the overview of six-sigma limit.

(e) Explain the following terms :

(i) Lot Tolerance Proportion Defective (LTPD)

(ii) Acceptance Quality Level (AQL)

4. What do you understand by sampling inspection plan ? Explain the concept of producer's risk and consumer's risk in such plan. Describe briefly the single sampling inspection plan. 2+4+4=10

Or

What are chance causes and assignable causes in SQC ? Explain the concepts of product control and process control. Describe briefly the double sampling inspection plan. 3+3+4=10

5. What are the $\overline{X}$ and R charts ? What purpose do they serve ? What are their advantages over the P chart ? 4+4+2=10

Or

What is Average Sample Number (ASN) and Average Total Inspection (ATI) ? Explain the method of their calculation for single sampling plan. Why are ASN and ATI calculated ? 4+5+1=10

6. Explain how a control chart helps to control the quality of a manufactured product. Justify for using the 3σ-limits in the control charts irrespective of the actual probability distribution of the quality characteristic. 5+5=10

Or

What is statistical process control ? Describe seven tools of it. 3+7=10
