M.Sc. Examination, 2024
Semester-I
Statistics
Course: MSC-11
(Linear Models and Distribution Theory)
Time: 3 Hours **Full Marks: 40**

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **any four** questions

1. (a). Under the assumptions of the Gauss-Markov Model, $\boldsymbol{y} = \boldsymbol{x\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{E(\epsilon) = 0}$, $\boldsymbol{Cov(\epsilon) = \sigma^2 I}$ if $\boldsymbol{\lambda'\beta}$ is estimable, find the BLUE of $\boldsymbol{\lambda'\beta}$.
   (b) Show that the estimator found in (a) is uncorrelated with all unbiased estimators of zero.

   5+5

2. (a). Show that with $\boldsymbol{G}$ a generalized inverse of $\boldsymbol{X'X}$, and $\boldsymbol{H} = \boldsymbol{GX'X}$, then $\boldsymbol{\lambda'\beta}$ is estimable if and only if $\boldsymbol{\lambda'H} = \boldsymbol{\lambda'}$.

   (b) Prove that the BLUE of any linear combination of estimable parametric function is the linear combination of their BLUEs.

   6+4

3. Show that for the linear model $\boldsymbol{y} = \boldsymbol{x\beta} + \boldsymbol{\epsilon}, \boldsymbol{\epsilon} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma^2 I})$

$$\frac{\frac{(\boldsymbol{\Lambda\hat{\beta}} - \boldsymbol{\Lambda\beta})'(\Lambda S^- \Lambda')^{-1}(\boldsymbol{\Lambda\hat{\beta}} - \boldsymbol{\Lambda\beta})}{m}}{\left(\frac{SSE}{n-r}\right)} \sim F_{m,n-r}$$

   where $\boldsymbol{\Lambda}^{m\times p}$ is of rank m, $\boldsymbol{S}^-$ is a generalized inverse of $(\boldsymbol{x'x})$. Discuss the case when $m = 1$. (You need to prove all the results to be used) 8+2

4. (a) Show that the conditional minimum of the sum of squares of the residuals $(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})$, in the model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \boldsymbol{E(\epsilon)} = 0, V(\boldsymbol{\epsilon}) = \sigma^2 I$, subject to m conditions $\boldsymbol{\Lambda\beta} = \boldsymbol{d}$, where $\boldsymbol{\Lambda\beta}$ are estimable and rank$(\boldsymbol{\Lambda}) = m$, exceeds the error sum of squares (SSE).

   (b) If $\boldsymbol{M} \sim \boldsymbol{W}_p(\boldsymbol{\Sigma}, m)$ and $\boldsymbol{B}$ is a $p \times q$ matrix, then show that

$$\boldsymbol{B'MB} \sim W_q(\boldsymbol{B'\Sigma B}, m)$$

   7+3

5. (a) If $\boldsymbol{M} \sim W_p(\boldsymbol{\Sigma}, m), m > p$ then show that the ratio

$\frac{\boldsymbol{a}'\boldsymbol{\Sigma}^{-1}\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{M}^{-1}\boldsymbol{a}}$ has the χ^2_{m-p+1} distribution for any fixed p-vector $\boldsymbol{a}$

(b). Derive the probability density function of a non-central Chi-square distribution.

6+4

6. Consider the following linear model:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \qquad i = 1,2,3; j = 1,2$$

Are the following functions estimable?

(i) μ

(ii) τ_1

(iii) $\mu + \tau_1$

(iv) $\tau_1 - \tau_2$

(v) $\tau_1 - \frac{\tau_2+\tau_3}{2}$

10

M.Sc. Examination, 2024-25
Semester-I
Statistics
Course: MSC-12
(Real Analysis and Measure Theory)
Time: 3 Hours **Full Marks: 40**

Questions are of value as indicated in the margin. Notations have their usual meanings.
Answer any four questions.

1. (a) Define measure and measure space. Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that for any $A, B \in \mathcal{A}$, we have the equality: $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

(b) Define measurable function. If f and g are measurable functions, then prove that $f + g$ is measurable.
(OR)
Define sigma-algebra. Show that if $f_n \xrightarrow{\mu} f$ on D and $g_n \xrightarrow{\mu} g$ on D, then $f_n + g_n \xrightarrow{\mu} f + g$ on D.

$5 + 5$

2. (a) If $g : \mathbb{R} \to \mathbb{R}$ is continuous and $f : \mathbb{R} \to \mathbb{R}$ is measurable, then $g \circ f$ is measurable. Using that (or otherwise), show that if f is measurable, then $|f|$ is measurable. Does the converse hold?

(b) Let $\Omega = \mathbb{R}^2$. Define A_n as the interior of the circle with radius 1 and center $\left(\frac{(-1)^n}{n}, 0\right)$. Find $\limsup_n A_n$ and $\liminf_n A_n$ with suitable explanations. $5 + 5$

3. (a) Define almost sure convergence, probability measures convergence, and distribution convergence. Show that the condition $\lim_{n\to\infty} \mu\{x \in D : |f_n(x) - f(x)| > \epsilon\} = 0$ for all $\epsilon > 0$ implies that $f_n \xrightarrow{\mu} f$ on D. Is the converse true?

(b) State the Borel-Cantelli Lemma. Show that if $\lim_{n\to\infty} \int_D \frac{|f_n - f|}{1 + |f_n - f|} \, d\mu = 0$, then $f_n \xrightarrow{\mu} f$ on D. (Hint: the function $\phi(x) = \frac{x}{1+x}$, $x > 0$, is increasing.) $5 + 5$

4. (a) Show that $|\varphi(t + h) - \varphi(t)| \leq E|e^{ihX} - 1|$ (uniform continuity property), where $\varphi(t)$ is characteristic function of a random variable. Let X be a real-valued random variable following geometric distribution with a characteristic function $\phi_X(t) = \mathbb{E}[e^{itX}]$. Find the expression of $\phi_X(t)$.

(b) Define Power series. Determine the radius of convergence and interval of convergence for either of the following power series: $\sum_{n=1}^{\infty} \frac{2^{2n}}{n}(2x - 4)^n$ (OR) $\sum_{n=1}^{\infty} \frac{(x - 6)^n}{n^n}$ $5 + 5$

5. (a) Prove that (M, d) is a metric space, where $M = \mathbb{R}^n$ and $d(\underline{x}, \underline{y}) = \max(|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|)$ for $\underline{x} = (x_1, x_2, \ldots, x_n), \underline{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$.
Prove that a set $A \subset \Re^n$ is closed if, and only if, it contains all its adherent points.

(b) Let S and T be subsets of $\mathbb{R}^n$. Let S' denote the derived set and $\overline{S}$ the closure of set S, then prove that $\overline{S \cap T} \subseteq \overline{S} \cap \overline{T}$.
(OR)
If S and T are subsets of $\mathbb{R}^n$, prove that $\text{int}(S) \cap \text{int}(T) = \text{int}(S \cap T)$, and $\text{int}(S) \cup \text{int}(T) \subseteq \text{int}(S \cup T)$.

$5 + 5$

6. (a) State and prove the Bolzano-Weierstrass theorem in $\Re^2$.

(b) Define: "covering of a set". State "Lindelof covering theorem" and "the Heine-Borel covering theorem". Define compact set in light of the two theorems. $5 + 5$

M.Sc. Examination, 2024
Semester-I
Statistics
Course: MSC-13
(Statistical Inference-I)
Time: Three Hours **Full Marks: 40**

Questions are of value as indicated in the margin

Notations have their usual meanings

1. Answer any FOUR questions of the following. 4x5=20

   (a) Let $(X_1, X_2, ..., X_n)$ be a random sample from R $(0, \theta)$ distribution with unknown parameter θ. Write $X_{(n)} = \max\ (X_1, X_2, ..., X_n)$. Then, if $L(\theta)$ denotes the likelihood function, show that, for $\theta > 0, L(X_{(n)}) \geq L(\theta)$, and make your comment.

   (b) Suppose $X_i; i = 1(1)n$ follows Bernoulli with parameter θ. The prior distribution of θ is Beta with parameters α and β. Find Bayes estimate of θ under squared error loss.

   (c) Let $(X_1, X_2, ..., X_n)$ be a random sample of size n drawn from Bernoulli population with parameter π. Find an UMVUE of $g(\pi) = 1 + n\pi + \frac{n(n-1)}{2}\pi^2$.

   (d) Describe Kendall's τ and derive its association with U-statistic.

   (e) Describe squared error loss function, absolute error loss function and all-or-nothing loss function. What are the Bayes estimates in these cases? Comment on Bayes estimate of mean of normal distribution.

   (f) Define U-statistic of degree m. Show that U-statistic is an unbiased estimator of population variance.

2. Answer any TWO questions of the following.

   (a) Describe maximum likelihood method of parameter estimation. State its properties. Discuss Minimum Chi-square method of parameter estimation and its limitations compared to maximum likelihood method. 3+3+4

   (b) Find the mean of U-statistic. Derive the limiting form of variance of it. State the result regarding the asymptotic distribution of U-statistic with conditions, if any. 3+4+3

   (c) Show that no unbiased estimator of a real parameter can be Bayes estimator under squared error loss. Let X_1, X_2, …, X_n be samples drawn from a normal distribution with mean μ and variance σ^2, where μ is unknown. The prior distribution of μ is assumed to be normal with mean γ and variance η^2. Derive the Bayesian point estimate of μ under the quadratic (squared error) loss function. 4+6

M.Sc. Examination 2024
Semester-I
Statistics
Course: MSC-14 (Sample Survey (Theory))
Full Marks: 40 Time: 3 Hours

(Answer any four questions.)

1. (a) Define Desh Raj's estimator. Show that it is unbiased for the population total.
   (b) Find the expression of its variance when the sample size is 2.
   (c) Mention a drawback of the estimator and the way of improvement. 3+4+3

2. 'Instead of applying PPSWR directly to draw samples from a population, if we classify the observations through SRSWOR and then select one unit from each group by PPSWR; there might be a gain in the efficiency.'– discuss. 10

3. (a) Describe how the double sampling technique can be applied to the regression method of estimation of the population total.
   (b) Find the approximate expressions for the bias and MSE of the estimator of the population total under this sampling scheme.
   (c) Discuss its efficiency over uni-phase sampling. 2+6+2

4. Given an unbiased estimator t of the population total Y based on a sample s drawn according to a design p, describe a procedure to construct an estimator having variance smaller than that of t. 10

5. (a) Compare the relative merits and demerits of cluster sampling and two-stage sampling.
   (b) Suggest an unbiased estimator of population total under two-stage sampling. Find its variance and an unbiased estimator of the variance. 2+(1+7)

6. (a) Explain the rationale behind Warner's randomized response technique.
   (b) Distinguish between related and unrelated question methods for estimating the proportion of individuals in a population possessing a sensitive characteristic. 5+5

'

M.Sc. Examination, 2024
Semester-I
Statistics
Course: MSC-15 (Practical)

Time: 4 Hours **Full Marks: 40**

One may use computer, if necessary
Questions are of value as indicated in the margin
Notations have their usual meanings

1. The hours per day that a television set was operating was recorded for a randomly selected collection of households, with the results shown in the table.

   3.7, 4.7, 2.1, 3.0, 4.3, 2.5, 4.2, 8.2, 3.6, 3.8, 2.1, 1.3, 1.5, 3.9, 1.1, 2.2, 2.4, 2.8, 3.6, 2.5, 7.3, 4.2, 6.0, 3.0, 5.9, 4.4, 4.2, 3.8, 3.7, 5.6, 3.7, 4.2, 1.5, 3.6, 5.9, 4.7, 8.2, 3.9, 2.5, 4.4, 2.1, 3.6, 1.1, 7.3, 4.2, 3.0, 3.8, 2.2, 4.2, 3.8, 4.3, 2.1, 2.4, 6.0, 3.7, 2.5, 1.3, 2.8, 3.0, 5.6.

Calculate the mean hours and variance hours. Construct a 95% confidence interval estimate for the mean hours that a television set is in operation in all households. 7

2. Consider the problem of point estimation of θ in $N(\theta,1)$. Given that θ belongs to [-0.5, 0.5]. On the basis of a sample of size n, the following estimator has been defined.

$$
\begin{aligned}
T &= -1 \text{ if } \bar{X} < -0.5 \\
&= \bar{X} \text{ if } -0.5 \le \bar{X} \le 0.5 \\
&= 1 \text{ if } \bar{X} > 0.5,
\end{aligned}
$$

$\bar{X}$ being sample mean. Assuming (i) squared error loss and (ii) absolute error loss draw the risk curve of $\bar{X}$ and T over the range $\theta \in [-0.5, +0.5]$ on the same graph paper and comment. Take n=15. 10

3. There are four objects w_1, w_2, w_3, w_4 whose weights are to be determined.

Left pan	Right pan	Weight needed for equilibrium
w_1, w_2, w_3, w_4	-------------	20
w_1, w_2	w_3, w_4	10
w_1, w_3	w_2, w_4	5
w_1, w_4	w_2, w_3	1

(a). Obtain best estimate of all weights

(b). Find their dispersion matrix of the estimators and estimate of the variance of each measurement. 6+3

3. (a) Consider the following regression model

$$y_1 = \beta_1 - 3\beta_2 + 4\beta_3 + \epsilon_1$$

$$y_2 = \beta_1 - 2\beta_2 + \beta_3 + 3\beta_4 + \epsilon_2$$

$$y_3 = \beta_1 - \beta_2 + 2\beta_4 + \epsilon_3$$

$$y_4 = \beta_1 + \beta_3 + \beta_4 + \epsilon_4$$

$$y_5 = \beta_1 + \beta_2 + \epsilon\_5$$

Estimate the following linear parametric functions (if estimable)

$2\beta_1 + \beta_2 + \beta_4$, $2\beta_1 - \beta_2 - \beta_3$, $\beta_1 + \beta_2$, $\beta_1 + \beta_4$ and β_4 where $y_1 = 11, y_2 = 21, y_3 = 13, y_4 = 45, y_5 = 50$.

Also find their standard errors. 9

4. Practical Note Book and Viva-Voce 5

M.Sc. Examination, 2024
Semester I
Statistics(Practical)
Course: MSC-16 (Sample Survey (Practical))
Full Marks :40 Time: 4 Hours

1. At an experimental station, there are 80 fields sown with wheat. Each field was divided into 16 plots of equal size ($1/16^{th}$ hectare). Out of 80 fields, 8 were selected by SRSWOR. From each selected field, 4 plots were chosen by SRSWOR. The yields in kg/plot are given below:

Selected Field	1	2	3	4	5	6	7	8
Plots	Yield							
1	4.32	4.16	3.06	4.00	4.12	4.08	5.16	4.20
2	4.84	4.36	4.24	4.84	4.68	3.96	4.24	4.66
3	3.96	3.50	4.76	4.32	3.46	3.42	4.96	3.64
4	4.04	5.00	3.12	3.72	4.02	3.08	3.84	5.00

(i) Estimate the yield of wheat per hectare for the experimental station along with its standard error.

(ii) How can one estimate obtained from a simple random sample of 32 plots be compared with the estimate obtained in (i)?

(iii) Obtain optimum n and m under cost function $5n + 2\,mn = 100$, where n and m respectively stand for the number of first stage units drawn and the number of second stage units drawn from each sampled FSU by SRSWOR. 5+3+4

2. The following list shows average area (in hectare) under wheat per village and the cluster size for each of the 30 clusters of villages:

Cluster No.	Cluster Size	Cluster Average	Cluster No.	Cluster Size	Cluster Average
1	20	12	16	22	22
2	17	10	17	24	21
3	22	14	18	54	11
4	25	9	19	12	9
5	30	8	20	19	9
6	32	12	21	22	7
7	12	10	22	23	14
8	58	11	23	35	6
9	23	8	24	35	9
10	26	15	25	25	8
11	24	17	26	21	7
12	47	21	27	21	14
13	51	9	28	25	12
14	32	11	29	52	11
15	25	14	30	23	9

(i) Find the sampling variance of the estimator of the average area under wheat per village on the basis of 10 sampled clusters selected without replacement.

(ii) If the average cluster size would have been 27, the within-cluster variance is 9.668, what will be the sampling variance of the estimator of average area under wheat per village with direct SRSWOR sample of 10×27, i.e. 270 villages? 4+4

3. A survey on 28 households was conducted and Warner's randomized response technique (Related Question method with π =0.61) was applied among the heads of the households to ask about the habit of underpaying the income tax. The actual amount of income tax underpaid is given in the following table.

Household		Response	Amount underpaid
Serial No.	Size	Yes(1)/No(0)	
1	3	1	2300
2	2	1	17000
3	5	0	5568
4	1	1	1304
5	3	0	0
6	2	1	0
7	4	0	711
8	3	0	1203
9	2	1	9874
10	4	1	2200
11	4	1	0
12	7	0	12000
13	2	1	1807
14	3	1	1400
15	4	0	708
16	4	1	1500
17	5	1	0
18	2	0	1100
19	1	0	1825
20	4	1	0
21	5	1	1407
22	3	0	713
23	5	0	1822
24	2	1	0
25	3	1	1623
26	5	1	1108
27	6	0	365
28	2	1	0

(i) Take a sample of 4 households using Rao-Hartley-Cochran's sampling scheme.

(ii) Estimate the total amount of income tax underpaid by these 32 households under the above scheme. Also provide an unbiased variance estimate.

(iii) Use the sample to estimate the proportion underpaying the income tax.

(iv) Now take a sample of 4 distinct households using Lahiri's method. Provide an unbiased estimate of the total amount of income tax underpaid under the sampling scheme. 5+3+2+5

4. Practical Note Book and Viva-Voce. 5

M.Sc. Examination, 2023
Semester-I
Statistics
Course: MSC-11
(Linear Models and Distribution Theory)
Time: 3 Hours **Full Marks: 40**

Questions are of value as indicated in the margin
Notations have their usual meanings

Answer **any four** questions

1. Consider the following linear model:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \qquad i = 1,2,3; j = 1,2$$

Are the following functions estimable?

(i) μ

(ii) τ_1

(iii) $\mu + \tau_1$

(iv) $\tau_1 - \tau_2$

(v) $\tau_1 - \frac{\tau_2+\tau_3}{2}$

10

2. (a). Under the assumptions of the Gauss-Markov Model,
$\boldsymbol{y} = \boldsymbol{x\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{E(\epsilon) = 0}$, $\boldsymbol{Cov(\epsilon) = \sigma^2 I}$
if $\boldsymbol{\lambda}'\boldsymbol{\beta}$ is estimable, find the BLUE of $\boldsymbol{\lambda}'\boldsymbol{\beta}$.
(b) Show that the estimator found in (a) is uncorrelated with all unbiased estimators of zero.

5+5

3. (a). Show that if $\boldsymbol{G}$ is a generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$, and $\boldsymbol{H} = \boldsymbol{G}\boldsymbol{X}'\boldsymbol{X}$, then $\boldsymbol{\lambda}'\boldsymbol{\beta}$ is estimable if and only if $\boldsymbol{\lambda}'\boldsymbol{H} = \boldsymbol{\lambda}'$.

(b) Prove that the BLUE of any linear combination of estimable parametric function is the linear combination of their BLUEs.

6+4

4. (a). Show that the general solution of the system of homogeneous equations $A\boldsymbol{x} = \boldsymbol{0}$ can be expressed as

$$\tilde{\boldsymbol{x}} = (I - H)\boldsymbol{z},$$

where $\boldsymbol{z}$ is any arbitrary vector, $\boldsymbol{H} = \boldsymbol{S}^{-}\boldsymbol{S}, \boldsymbol{S} = \boldsymbol{X}'\boldsymbol{X}$
How will you modify this result for the non-homogeneous consistent equations $A\boldsymbol{x} = \boldsymbol{u}$?
(b) Let $\boldsymbol{l}_1'\boldsymbol{\beta}$ and $\boldsymbol{l}_2'\boldsymbol{\beta}$ be two estimable functions and $\boldsymbol{l}_1'\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{l}_2'\widehat{\boldsymbol{\beta}}$ be their least square estimators respectively. Find **Var**$(\boldsymbol{l}_1'\widehat{\boldsymbol{\beta}})$ and **Cov**$(\boldsymbol{l}_1'\widehat{\boldsymbol{\beta}}, \boldsymbol{l}_2'\widehat{\boldsymbol{\beta}})$.

6+4

5. (a) Show that the conditional minimum of the sum of squares of the residuals $(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})$, in the model $\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\epsilon}, \boldsymbol{E}(\boldsymbol{\epsilon})=0, V(\boldsymbol{\epsilon})=\sigma^2 I$, subject to m conditions $\boldsymbol{\Lambda\beta}=\boldsymbol{d}$, where $\boldsymbol{\Lambda\beta}$ are estimable and rank($\boldsymbol{\Lambda}$) $=m$, exceeds the error sum of squares (SSE).

(b) Show that $\frac{SSR(\beta)}{\sigma^2}$ follows a non-central χ^2 distribution. 6+4

6. (a) If $\boldsymbol{M}\sim W_p(\boldsymbol{\Sigma}, m), m>p$ then show that the ratio

$\frac{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{M}^{-1}\boldsymbol{a}}$ has the χ^2_{m-p+1}distribution for any fixed p-vector $\boldsymbol{a}$

(b). Show that for the linear model $\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\epsilon}, \boldsymbol{\epsilon}\sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma^2 I})$

$$\frac{\frac{(\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}-\boldsymbol{\Lambda\beta})'(\Lambda S^-\Lambda')^{-1}(\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}-\boldsymbol{\Lambda\beta})}{m}}{\left(\frac{SSE}{n-r}\right)} \sim F_{m,n-r}$$

where$\boldsymbol{\Lambda}^{m\times p}$ is of rank m, $\boldsymbol{S}^-$ is a generalized inverse of$(\boldsymbol{X}'\boldsymbol{X})$. Discuss the case when $m=1$. (You need to prove all the results to be used) 4+6

M.Sc. Examination, 2023
Semester-I
Statistics
Course: MSC-12
(Real Analysis and Measure Theory)
Time: 3 Hours **Full Marks: 40**

Questions are of value as indicated in the margin.
Notations have their usual meanings

NOTE: There are a total of 6 questions. Answer any four questions.

1. (a) Define measure. Suppose μ is a translation invariant measure on $(\Re, \mathfrak{B})$. Set $\gamma := \mu((0,1])$ and assume $0 < \gamma < \infty$. Show that $\mu((0,1/n]) = \dfrac{\gamma}{n}$ for all $n \in \mathbb{N}$, and $\mu((0,q]) = \gamma q$ for all rational $q > 0$.

(b) Define sigma-algebra and measurable function. Determine whether the following function $f : \mathbb{R} \to \mathbb{R}$ is measurable:

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

5 + 5

2. (a) Prove that, $X_n \xrightarrow{a.s.} X$ implies $X_n \xrightarrow{P} X$. Let $X_1, X_2, \ldots$ be a sequence of independent random variables such that X_n has a uniform distribution on the interval $[0,1]$. Define $Y_n = \max\{X_1, X_2, \ldots, X_n\}$. Show that $Y_n \xrightarrow{a.s.} 1$ and hence, $Y_n \xrightarrow{P} 1$.

(b) State the Borel-Cantelli Lemma. Let $X_1, X_2, \ldots$ be a sequence of random variables such that X_n has a normal distribution with mean μ_n and variance σ_n^2. Suppose $\mu_n \to \mu$ and $\sigma_n^2 \to \sigma^2$ as $n \to \infty$. Show that $X_n \xrightarrow{d} X$ where X has a normal distribution with mean μ and variance σ^2.

5 + 5

3. (a) If X is a random variable with characteristic function $\phi_X(t)$, then show that $\phi_{-X}(t) = \overline{\phi_X(t)}$, where $\overline{\phi_X(t)}$ denotes the complex conjugate of $\phi_X(t)$. Find the characteristic function of the random variable X with probability density function given by:

$$f(x) = \begin{cases} e^{-x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) Determine the radius of convergence and interval of convergence for the following power series: (answer one)

$$(i) \quad \sum_{n=1}^{\infty} \frac{2^{2n}}{n}(2x-4)^n \qquad (OR) \qquad (ii) \quad \sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n}(x+3)^n$$

5 + 5

4. (a) Define Borel sigma-algebra. Let $\Omega = \Re^2$. Define A_n as the interior of the circle with radius 1 and center $\left(\dfrac{(-1)^n}{n}, 0\right)$. Find $\limsup_n A_n$ and $\liminf_n A_n$ with suitable explanations.

(b) Define closed set. Prove that a set $A \subset \Re^n$ is closed if, and only if, it contains all its adherent points.

5 + 5

CONTINUED

5. (a) State and prove the Bolzano-Weierstrass theorem in $\Re^2$.

(b) Define: "covering of a set". State "Lindelof covering theorem" and "the Heine-Borel covering theorem". Define compact set in light of the two theorems.

$5 + 5$

6. (a) Show that the set of accumulation points of a set A is closed.

(b) Prove that every limit point of a set A is also an adherent point of A.

(c) Let $e_n = \cos n$. Determine whether e_n has a convergent subsequence.

(d) Let $f_n(x) = \frac{n}{n+x}$ for $x \geq 0$. Find $\lim_{n\to\infty} \int f_n(x)dx$ and $\int \liminf_{n\to\infty} f_n(x)dx$. Does Fatou's Lemma hold in this case? Explain your answer.

2.5×4

M.Sc. Examination, 2023
Semester-I
Statistics
Course: MSC-13
(Statistical Inference-I)

Time: Three Hours **Full Marks: 40**

Questions are of value as indicated in the margin
Notations have their usual meanings

1. Answer any FOUR questions from the following. 4x5=20

(a) Let $(X_1, X_2, ..., X_n)$ be a random sample from $f(x,\theta) = e^{-(x-\theta)}, x > \theta, \theta > 0$ distribution with unknown parameter θ. Write $X_{(1)} = \min\ (X_1, X_2, ..., X_n)$. Then, if $L(\theta)$ denotes the likelihood function, show that, for $\theta > 0, L(X_{(1)}) \geq L(\theta)$, and make your comment.

(b) Suppose $X_i; i = 1(1)n$ follows Bernoulli with parameter θ. The prior distribution of θ is Beta with parameters α and β. Find the Bayes estimate of θ under squared error loss.

(c) Let $(X_1, X_2, ..., X_n)$ be a random sample of size n drawn from Poisson population with parameter λ. Find an UMVUE of $g(\lambda) = \lambda^2$.

(d) Describe Kendall's τ and its association with U-statistic.

(e) Describe squared error loss function, Absolute error loss function and all-or-nothing loss function. What are the Bayes estimates in these cases? Comment on the Bayes estimate of mean of normal distribution.

(f) Define U-statistic of degree m. Show that U-statistic of degree 2 is an unbiased estimator of population variance.

2. Answer any TWO questions from the following.

(a) Describe Maximum likelihood method of parameter estimation. State its properties. Suppose $(X_1, X_2, ..., X_n)$ is a random sample of size n from a distribution with cumulative distribution function

$$F(x \mid \alpha, \beta) = 0 \quad if\ x < 0$$

$$= \left(\frac{x}{\beta}\right)^{\alpha} \quad if\ 0 \leq x \leq \beta$$

$$= 1 \quad if\ x > \beta$$

where the parameters $\alpha(>0)$ and $\beta(>0)$ are unknown. Obtain the maximum likelihood estimators of α and β. 3+3+4

(b) Find the mean of U-statistic. Derive the limiting form of variance of it. State the result regarding the asymptotic distribution of U-statistic with conditions, if any. 3+4+3

(c) Derive the Bayes estimate of a real parametric function $\gamma(\theta)$ under squared error loss. Let $X_1, X_2, \ldots, X_n$ be a sample drawn from a normal distribution with mean μ and variance σ^2, where μ is unknown. The prior distribution of μ is assumed to be normal with mean γ and variance η^2. Derive the Bayesian point estimate of μ under the quadratic (squared error) loss function. Write it in the form of credible estimate and mention its credibility factor. 4+6

M.Sc. Examination 2023
Semester-I
Statistics
Course: MSC-14 (Sample Survey (Theory))
Full Marks: 40 **Time: 3 Hours**

(Answer any four questions.)

1. (a) Explain the problem you will encounter while asking people directly about their choices on a dichotomous sensitive characteristic.

   (b) Discuss how will you estimate the proportion of individuals belonging to different political parties of West Bengal using unrelated questionnaire method, where the unrelated questions have 'Yes/No' type response. 4+6

2. (a) Is cluster sampling more efficient than uni-stage sampling? Justify.

   (b) What is the difference between cluster sampling and two-stage sampling? Suggest an unbiased estimator of the population mean in two-stage sampling. Also find its variance. 3+(1+2+4)

3. (a) Define Horvitz-Thompson's estimator . Show that it is unbiased for the population total.

   (b) Find its variance and express it in the form of Sen, Yates and Grundy. Also suggest an unbiased estimator of the variance.

   (c) Prove that if there exists an unbiased estimator of the population total, then the inclusion probabilities of each of the population units must be strictly positive. 3+4+3

4. (a) Describe how the double sampling technique can be applied in stratification.

   (b) Suggest an unbiased estimator of the population mean under this sampling scheme and find the variance of the estimator.

   (c) Discuss how you can find the optimum sizes of the first and the second phase samples for a fixed cost. 2+3+5

5. (a) What do you mean by PPS sampling? When should you consider it as an appropriate sampling scheme?

   (b) Estimate the gain in precision using PPSWR instead of SRSWR.

   (c) Describe Rao-Hartley-Cochran's sampling strategy. Suggest an unbiased estimator of the population total under this scheme. 3+4+3

'

6. (a) Define informative and non-informative sampling designs.

   (b) Consider a population containing three units U_1, U_2, U_3 with non-negative variable values Y_1, Y_2, Y_3. A random sample s of two units is selected under the design

(P.T.O)

$$p(s) = \begin{cases} \frac{1}{3} & \text{if } n(s) = 2 \\ 0. & \text{otherwise} \end{cases}$$

To estimate the population mean $\overline{Y} = \dfrac{Y_1 + Y_2 + Y_3}{3}$, the following two estimators are considered

$$t(s) = \begin{cases} \frac{1}{2}(Y_1 + Y_2) & \text{if } s = (U_1, U_2) \\ \frac{1}{2}(Y_1 + Y_3) & \text{if } s = (U_1, U_3) \\ \frac{1}{2}(Y_2 + Y_3) & \text{if } s = (U_2, U_3) \end{cases}$$

$$t'(s) = \begin{cases} \frac{1}{2}Y_1 + \frac{1}{2}Y_2 & \text{if } s = (U_1, U_2) \\ \frac{1}{2}Y_1 + \frac{2}{3}Y_3 & \text{if } s = (U_1, U_3) \\ \frac{1}{2}Y_2 + \frac{1}{3}Y_3 & \text{if } s = (U_2, U_3) \end{cases}$$

i. Show that both t and t' are unbiased for $\overline{Y}$.

ii. Find the condition under which t' is better than t.

(c) Find the expectation and variance of the number of distinct elements in a SRSWR of size 3 from a population of size 10. You need not prove the necessary results.

2+(2+3)+3

M.Sc Semester I Examination, 2023

Statistics

MSC-15(Practical)

Time: Four Hours   Full Marks: 40

One may use computer, if necessary

2. 1. For double genetically data with some value of π, for both parents, following distribution is obtained:

	D1D2	D1R2	D2R1	R1R2
Frequency:	191	36	33	27
Probability:	$\frac{2+p}{4}$	$\frac{1-p}{4}$	$\frac{1-p}{4}$	$\frac{p}{4}$

where $p=(1-\pi)^2$. Find Maximum Likelihood Estimate of π and estimate its standard error. 9

2. Consider the problem of point estimation of θ in $N(\theta,1)$. Given that θ belongs to [-0.5, 1]. On the basis of a sample of size n, the following estimator has been defined.

$$\begin{aligned} T &= -1 \text{ if } \overline{X} < -0.5 \\ &= \overline{X} \text{ if } -0.5 \le \overline{X} \le 1 \\ &= 1 \text{ if } \overline{X} > 1, \end{aligned}$$

$\overline{X}$ being sample mean. Assuming (i) squared error loss and (ii) absolute error loss draw the risk curve of $\overline{X}$ and T over the range $\theta \in [-0.5,1]$ on the same graph paper and comment. Take n=15. 8

3. (i) Suppose that the two observations on each of three treatments are as follows.

Treatments		
t_1	t_2	t_3
8	5	12
6	3	14

Assuming the linear model to be
$y = X\beta + \epsilon$, where $\beta' = (\mu, t_1, t_2, t_3)$, find the BLU estimates of the following functions:

a. μ

b. $t_1 - t_2$

c. $2\mu + t_1 + t_2$

d. $\mu + \frac{t_1+t_2+t_3}{3}$

e. $t_1 + t_2$

(ii) Would you like to reject the null hypothesis $t_1 - t_2 = 7$ at 5% level?

8+6

4. Consider the following data on perspiration in 10 healthy females measured in terms of sweating rate (X_1) along with $X_2 = Na\, content$ and $X_3 = Ka\, content$.

No.	X_1	X_2	X_3
1	3.7	48.5	9.3
2	5.7	65.1	8
3	3.8	47.2	10.9
4	3.2	53.2	12
5	3.1	55.5	9.7
6	4.6	36.1	7.9
7	2.4	24.8	14
8	7.2	33.1	7.6
9	6.7	47.4	8.5
10	5.4	54.1	11.3

Assuming $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, perform a test procedure to test the following hypothesis:

$$H_0: \boldsymbol{\mu} = \begin{pmatrix} 4 \\ 50 \\ 10 \end{pmatrix}$$

4

5. Practical Note Book and Viva-Voce 5

****

M.Sc. Examination, 2023
Semester I
Statistics(Practical)
Course: MSC-16 (Sample Survey (Practical))
Full Marks :40 Time: 4 Hours

(1) The following is the population of M.Sc. (Statistics) students of a certain university, together with the data on two variables, viz. Marks obtained in PG Sample survey course and Marks in UG Sample survey course (both out of 50).

Serial Number	Student Roll no.	Marks in UG Sample Survey	Marks in PG Sample Survey
1	MSC-STAT-01	31	28
2	MSC-STAT-02	26	29
3	MSC-STAT-03	40	42
4	MSC-STAT-04	40	37
5	MSC-STAT-05	36	44
6	MSC-STAT-06	45	45
7	MSC-STAT-07	29	22
8	MSC-STAT-08	21	19
9	MSC-STAT-09	25	21
10	MSC-STAT-10	30	34
11	MSC-STAT-11	39	33
12	MSC-STAT-12	42	43
13	MSC-STAT-13	34	38
14	MSC-STAT-14	30	24
15	MSC-STAT-15	26	28
16	MSC-STAT-16	28	33
17	MSC-STAT-17	45	40
18	MSC-STAT-18	39	43
19	MSC-STAT-19	35	26
20	MSC-STAT-20	41	34

(a) Form 4 groups, each containing 5 students using SRSWOR.
(b) Select 2 groups at random and 2 students from each group at random.
(c) On the basis of your sampling scheme, estimate the average marks in PG sample survey. Find an unbiased estimate of the variance of your estimator.
(d) Next select one member from each group in such a way that the probability of selecting a student is proportional to his/her marks in UG sample survey.
(e) Under this scheme, estimate the average marks in PG sample survey and find an unbiased estimate of the variance of your estimator.
(f) Also select 4 distinct students directly from the list using Lahiri’s method, using the marks in UG sample survey marks as auxiliary information.

(g) Estimate the average marks of PG sample survey based on the samples in question (e). Find an unbiased estimate of the variance of your estimator. Comment whether the use of PPS sampling is appropriate here. 3+2+4+2+5+4+5

(2) Consider the following population:

Unit Number	x-Value	y-Value
1	7	32
2	11	41
3	4	25
4	10	67

Evaluate the variance of Horvitz-Thompson's estimator assuming a PPSWOR sample of size two is drawn. 5

(3) Following a list of 10 units together with the respective size measures taken suitably.

Serial Number	Size Measure
1	542
2	245
3	1032
4	867
5	256
6	1342
7	390
8	604
9	465
10	897

In a without replacement sampling scheme, the first selection is made with probability proportional to size and the remaining units of a sample are selected with equal probability without replacement($n = 3$). Find the first order inclusion probabilities of all units and the second order inclusion probabilities of the following pairs of units : (1,6), (2,9), (4,7), (6,9), (8,10). 5

(4) Practical Note Book and Viva-Voce. 5