Name____________________ Seat No. ____________________

THE MAHARAJA SAYAJIRAO UNIVERSITY OF BARODA

Mid Semester Examination

M.Sc. Statistics Semester I

Date:12.9.2023 Measure Theory (STA 2101C01) Time:

Day: Tuesday 3.00 to 4.00 PM

Note: (i) Write all the answers in the answer book.

(ii) Numbers on the right indicates marks.

1. In each of the following questions there are two statements S_1 and S_2. Read the Statements and write the correct alternative from the following. **(6x2=12)**

(A) Both S_1 and S_2 are correct. (B) Only S_1 is correct

(C) Only S_2 is correct (D) Neither S_1 nor S_2 is correct.

(i) S_1: Every measure is always continuous from below.

S_2 : Every measure is always continuous from above.

(ii) S_1: Every subset of real line is a Borel set.

S_2: Every Borel set is a subset of real line.

(iii) S_1: A monotone field is a σ-field.

S_2 : A σ-field is always a monotone class.

(iv) S_1: $\liminf \mu(A_n) \leq \mu(\liminf A_n)$

S_2 : $\limsup \mu(A_n) \geq \mu(\limsup A_n)$, when μ is finite.

(v) S_1: A finitely additive set function is always countably additive.

S_2: A countably additive set function ψ is finitely additive if $\psi(\phi)=0$.

(VI) S_1: A simple function is always a measurable function.

S_2: A limit of a sequence of simple functions is always a measurable function.

Name________________________ Seat No. ________________________

2. Answer the following.

(i)

$$\text{Let } F(x) = \begin{cases} 0 & \text{if } x < 3 \\ \frac{x}{6} - \frac{1}{8} & \text{if } 3 \leq x \leq 4 \\ 1 & \text{if } x \geq 4. \end{cases}$$

Find μ_F for the following sets

(1) A={3, 4} (2) [3, 4] (3) (3, 4] (4) [3,4). (4)

(ii) Give an example of sequence of open intervals whose limit is a closed Interval. (3)

(iii) Define a σ-finite set function.(2)

(iv) State the two descriptive definitions of a measurable function. Prove their equivalence. (5)

(v) Explain in detail Lebesgue measure. (4)

****

Exam Seat No. 8

Mid Semester Examination
Faculty of Science
Department of Statistics
M.Sc. Semester–I Examination
STA2111C11 : Sampling Theory

Date : 16-09-2023
Day: Saturday
Time : 3.00 to 4.00 pm

Q1. Choose the correct alternative (12x1=12)

1. Under Simple Random Sampling in Ratio method
   (a) $B(\hat{Y}_R)=\left[\frac{N^2(N-n)}{Nn}\right] Y\left\{\left[\frac{S_x^2}{X^2}\right]-\left[\frac{S_{xy}}{XY}\right]\right\}$ (b) $\hat{Y}_R=\frac{\Sigma \quad Y_i}{\Sigma \quad X_i}.X$
   (c) $MSE(\hat{Y}_R)=\left[\frac{N^2(N-n)}{Nn}\right] Y^2\left\{\left[\frac{S_y^2}{Y^2}\right]+\left[\frac{S_x^2}{X^2}\right]-2\left[\frac{S_{xy}}{XY}\right]\right\}$ (d) all three
2. In random group method population is randomly divided in to groups which are
   (a) mutually exclusive but not exhaustive. (b) not mutually but exhaustive
   (c) mutually exclusive and exhaustive. (d) only exhaustive.
3. In random group method
   (a) $V_1E_2\,[\hat{Y}_{RHC}|G_1, \ldots, G_n] > 0$ (b) $V_1E_2\,[\hat{Y}_{RHC}|G_1, \ldots, G_n] < 0$
   (c) $V_1E_2\,[\hat{Y}_{RHC}|G_1, \ldots, G_n] \neq 0$ (d) $V_1E_2\,[\hat{Y}_{RHC}|G_1, \ldots, G_n] = 0$
4. A larger sample size is required when:
   (a) The population of interest for a study is less diverse
   (b) A low level of precision is required
   (c) The population of interest is easily recruited to the study
   (d) A high level of precision is required
5. In Lahiri's method unit is selected finally if
   (a) $R = X_i$ (b) $R \leq X_i$
   (c) $R \geq X_i$ (d) *none of the above*
6. If $Y_i \propto X_i$ in PPS Sampling then
   (a) $V_{pps}(\hat{Y}_{unst}) < 2\, V_{prop}(\hat{Y}_{st})$ (b) $V_{pps}(\hat{Y}_{unst}) = V_{prop}(\hat{Y}_{st})$
   (c) $V_{pps}(\hat{Y}_{unst}) \leq V_{prop}(\hat{Y}_{st})$ (d) $V_{pps}(\hat{Y}_{unst}) \geq V_{prop}(\hat{Y}_{st})$
7. A sample consists of
   (a) all units of the population (b) 50 percent of the population
   (c) any fraction of the population (d) 5 percent unit of the population
8. If larger units have greater probability of their inclusion in the sample, it is known as
   (a) selection with replacement (b) selection with probability proportional to size
   (c) selection with equal probability (d) judgement sampling
9. Under SRS, the product estimator is more precise than the expansion estimator, when the variables X and Y have
   (a) high negative correlation (b) low positive correlation
   (c) high positive correlation (d) none of the above
10. Let $e_o = \frac{\hat{Y}-Y}{Y}$ then e_o satisfied
    (a) $E(e_o) = 0$ (b) $E(e_o^2) = V(\hat{Y})/Y^2$
    (c) both the above are correct. (d) both the above are wrong.
11. Which of the following statement is true
    (a) less the standard error, better it is (b) less the variance, better it is
    (c) both (a) and (b) are correct (d) standard error is always unity
12. Whenever the groups are of same size in Random group method
    (a) $V(\hat{Y}_{RHC}) > V(\hat{Y}_{PPS})$ (b) $V(\hat{Y}_{RHC}) = V(\hat{Y}_{PPS})$
    (c) $V(\hat{Y}_{RHC}) < V(\hat{Y}_{PPS})$ (d) none of the above

P.T.O.

Q2.(a.) What is random group Method? (03)

(b.) Prove that under ppswor, $\hat{Y}_{DR}$ is unbiased for the population total and unbiased estimator of (06)

$$V(\hat{Y}_{DR}) \text{ is } \frac{1}{n(n-1)} \sum_{i=1}^{n} \left(t_i - \underline{t}\right)^2$$

Q3(a.) Explain Lahiri's method with suitable example. In which situation Lahiri method is more suitable than cumulative method? (03)

(b.) In pps sampling with out replacement, $\hat{Y}_{HT}$ is unbiased and its sampling variance is given by $V_{HT}(\hat{Y}_{HT}) = \sum_{i=1}^{N} \frac{(1-\pi_i)Y_i^2}{\pi_i} + \sum_{i}^{N} \sum_{i \neq j}^{N} \frac{(\pi_{ij}-\pi_i\pi_j)}{\pi_i\pi_j} Y_i Y_j$ where π_{ij} is the probability of inclusion of both the ith and jth unit in the sample. Prove it. (06)

OR

(b.) Derive bound for bias of ratio estimator. (06)

----------X--------

① B
② C
③ B
④ D
⑤ C
⑥
⑦ C
⑧ C
⑨ A
⑩ C
⑪ B
⑫ C
⑬
⑭ D

2022
B
B
A
C
D
D

23-6-23

⑭ C

Roll. no. 8
VISHNU SUTHAR.

Date: 15/09/2023

The Maharaja Sayajirao University of Baroda
Faculty of Science, Department of Statistics
M.Sc. Semester – I (Mid-Semester Examination)

Time: 3:00 to 4:00 p.m.
Marks: 30

STA2104C04 – Decision Theory

Note: Answer to all the questions, including Q.1, must be written in the answer book provided ONLY.

Q.1 (A) Choose the correct option and write in the answer book: (1 mark each)

(1) A decision rule is mapping from

(a) Sample space to parameter space
(b) Parameter space to action space
(c) Sample space to action space
(d) None of these

(2) A randomized decision for a decision problem (Θ, a, L), with sample space $\mathcal{X}$ and decision set $\mathcal{D}$, is associated with a probability distribution over

(a) Θ (b) $\mathcal{D}$ (c) $\mathcal{X}$ (d) none of these

(3) Which of the following can be used as a strategy for reducing the number of decision rules out of all possible decision rules?

(a) Principle of unbiasedness
(b) The Bayes Principle
(c) Principle of invariance
(d) Either (a) or (c)

(4) A decision theory problem can be referred to as a problem of point estimation when

(a) $a = (-\infty, \infty)$
(b) a is countably infinite space
(c) a is finite
(d) Any of the above

(5) Let P_1 and P_2 be two randomized actions. Then which of the following will also be a randomized action?

(a) $\frac{1}{2}(P_1 + P_2)$
(b) $0.4\,P_1 + 0.6\,P_2$
(c) Both (a) and (b)
(d) $0.4\,P_1 + 0.4\,P_2$

(6) The Bayes decision rule $d(x)$ for a quadratic loss function can be obtained as

(a) $d(x) = E(X|\theta)$
(b) $d(x) = E(\theta|X = x)$
(c) $d(x) = E(X|a)$
(d) $d(x) = E(\theta|a)$

Q.1 (B) Do as directed: (2 marks each)

(1) A farmer must decide whether or not to plant his crop early. If he plants early and the monsoon occurs timely, then he will gain Rs. 50000 in extra harvest, but if he plants early and the monsoon occurs late, then he will lose Rs. 20000 as the cost of reseeding. If he doesn't plant early, his gain will be Rs. 0. Describe the parameter space, action space and loss table based on this information.

(2) Explain the basic difference between randomized decision rule and behavioral decision rule.

(3) Discuss the decision theory problem as a problem of testing of hypothesis.

Q.2 Attempt Any TWO: (6 marks each)

(a) Define: statistical decision problem, randomized action.

Show that, risk function $R(\theta, \delta)$ is a convex set.

(b) Define: prior distribution, Bayes risk, Bayes decision rule, minimax decision rule.

(c) Define: unbiased decision rule, least favorable prior distribution. State and prove the necessary and sufficient condition under which a prior distribution will be least favorable.

Q.3 Answer Any ONE: (6 marks)

(a) Let $\Theta = (0, \infty)$, $a = \mathcal{R}$ and loss function is squared error. Suppose $(X \mid \theta) \sim Poisson\,(\theta)$ and the prior distribution of θ is

$$g(\theta) = \frac{e^{-\theta/\beta}\,\theta^{\alpha-1}}{\Gamma\alpha\,\beta^{\alpha}} \;;\; \theta > 0 \quad \text{where } \alpha, \beta > 0$$

(i) Show that the posterior distribution of θ given $X = x$ is $G\left(\alpha + x, \frac{\beta}{\beta+1}\right)$.

(ii) show that the Bayes rule is $d_{\alpha,\beta}(x) = \frac{\beta\,(\alpha+x)}{(\beta+1)}$.

(b) An insurance company is faced with taking one of the following 2 actions: decrease sales force by 10% or maintain the existing work force. Depending upon the economy is good, mediocre or bad, the company would expect to lose following amount of money.

		Action taken	
		Decrease sales force by 10%	Maintain the existing work force
State of Economy	Good	-10	-5
	Mediocre	-5	-5
	Bad	1	0

Suppose there is an indicator variable that can be used to have a prediction about the state of economy. Probability that the indicator variable will predict the state of economy correctly is 0.6, where as the probability that the indicator variable will predict the state of economy wrongly is 0.2 for each incorrect prediction. Prepare the risk table for this decision problem.

******

Date:13-9-2023	Department of statistics	Time:3.00PM-4.00PM
Day: Wednesday	M.Sc. Sem I (Mid Sem sept 2023) STA2102C02 -Linear Models(4 credits)	Marks:30

Q-I **Choose the correct alternative** [12]

1. In the usual notation a linear function l'y is said to belong to error if________ irrespective of the values of Θ_1, Θ_2-----Θ_m.
   (A) $E(l'y) > 0$
   (B) $E(l'y) < 0$
   (C) $E(l'y) = 0$.
   (D) None of these

2. In the usual notation a model in which the coefficient a_{ij}'s take indicator variables is called
   (A) Regression model
   (B) Analysis of variance model
   (C) Both (A) and (B)
   (D) Analysis of covariance model

3. In the vector space theory estimation space and error space are ____________ to each other
   (A) Orthogonal
   (B) Perpendicular
   (C) Can be either (A) or (B)
   (D) None of these

4. Suppose y_1, y_2, y_3 are independent random variables with a common variance σ^2 and $E(y_1) = \Theta_1 + \Theta_3$, $E(y_2) = \Theta_2 + \Theta_3$ and $E(y_3) = \Theta_1 + \Theta_3$ then $b_1 \Theta_1 + b_2\Theta_2 + b_3\Theta_3$ is estimable iff
   (A) $b_3 = b_1 + b_2$
   (B) $b_1 = b_2 + b_3$
   (C) $b_2 = b_1 + b_3$
   (D) None of these

5. A linear manifold in a vector space V is any subset of vectors M closed under_______
   (A) Addition
   (B) Scalar multiplication
   (C) Division
   (D) Both (A) and (B)

6. Which of the following is a contrast?

(A) $L = 3X_1 - 3X_2 + X_3$
(B) $L = 2X_1 - X_2 - X_3$
(C) $L = X_1 - X_2 - X_3$
(D) $L = X_1 + X_2 - X_3$

Q-II **Do as Directed** [18]

1. State different linear models along with assumptions [3]

OR

1. Distinguish between fixed effect model and random effect model by giving examples [3]
2. Define (i) estimation space (ii) error space [3]
3. For the following model [4]

$$E(Y_i) = \delta + \beta(x_i - \bar{x}), \ i = 1,2,3\text{-------}n$$
$$D(Y) = \sigma^2 I$$

Obtain the LS estimators of δ and β and show that $cov(\hat{\delta}, \hat{\beta}) = 0$

4. Explain Scheffe's test [4]
5. Write a note on power of F-test [4]

****

2022

7 C
8 B
9 B
10 A
11 D
12 A
13 D
14
15

1 C
2 B
3 A
4 A
5 A
6 D

2017

2018

1
2 B
3 C
4 B
5 D
6 D
7 C
8 D

2019

$x(x'x)^{-1}x'$

$\frac{e_i}{1-h_{ii}}$

$\frac{e_i}{\sqrt{MSSE}}$

2019

1
2
3 D
4 D
5 C

THE MAHARAJA SAYAJIRAO UNIVERSITY OF BARODA

M.Sc. (SEM I) MID- SEM EXAMINATION

DAY: WEDNESDAY

DATE: 27.09. 2023

TIME: 1:15-2:15 PM

SUBJECT/PAPER: STATISTICS/MULTIVARIATE ANALYSIS

PAPER CODE: STA-2103 C 03

TOTAL MARKS: 30

PART-1 (12 MARKS)

1. Suppose that the data is collected on 10 individuals for their 6 characteristics then the data matrix **X** will be of the order

(A) 6X6 (B) 6X10 (C) 10X6 (D) 10X10

2. Consider a random vector **X** of the order 6X1 then its associated variance – covariance matrix ∑ will have

(A) 12 distinct elements (B) 36 distinct elements (C) 21 distinct elements

(D) 18 distinct elements.

3. If the expression of the exponent of a bivariate normal distribution is

$-1/102\ [\ (x+2)^2 - 2.8(x+2)(y-1) + 4(y-1)^2]$ then the mean of X and Y are

(A) 2, 1 (B) -2, 1 (C) -2,-1 (D) -2,-1 respectively

4. In a general if the density function of a multivariate normal density is written as

$\frac{6(6+1)}{2}$ $\frac{\cancel{6}^{3}\times 7}{\cancel{2}}$

$f(\mathbf{X}) = k\, e^{-1/2}\ (\mathbf{X}-\mathbf{B})'\ A\ (\mathbf{X}-\mathbf{B})$ where **X** and **B** are the vectors of order 'p' and A is positive definite matrix then the probability density function of **X** is defined when

(A) $(\mathbf{X}-\mathbf{B})'\ A\ (\mathbf{X}-\mathbf{B}) > 0$ (B) $(\mathbf{X}-\mathbf{B})'\ A\ (\mathbf{X}-\mathbf{B}) < 0$ (C) $(\mathbf{X}-\mathbf{B})'\ A\ (\mathbf{X}-\mathbf{B}) \geq 0$

(D) $(\mathbf{X}-\mathbf{B})'\ A\ (\mathbf{X}-\mathbf{B}) \leq 0$

$p + \frac{p(p+1)}{2}$ $6 + \frac{\cancel{6}^{3}(7)}{\cancel{2}} = 27$

5. Let the random vector **X** of order px1 be distributed as $N_p(0, \Sigma)$ then its characteristic function will be given by

(A) $\exp(-1/2\, t' \Sigma t)$ (B) $\exp(-1/2\, it')$ (C) $\exp(-1/2\, it'\mu)$

(D) $\exp(-1/2\, it'\mu - t' \Sigma t.)$

6. Let $X' = (1, 2)$ $\mu' = (3,4)$ and the variance-covariance matrix be an identity matrix then the value of the Q.F. $(X-\mu)' \Sigma^{-1} (X-\mu)$ will be

(A) 12 (B) 14 (C) 16 (D) 32

7. Consider the partitioning of the mean vector μ and the variance-covariance matrix Σ as $\mu' = (\mu_1, \mu_2)'$ and $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

As the random vector X_{10x1} was partitioned into two sub-sets (X^1 and X^2) one having 4 components and other the remaining ones, then the Marginal distribution of X^1 will be

(A) $N_4(\mu,\Sigma)$ (B) $N_6(\mu_1, \Sigma_{11})$ (C) $N_4(\mu_1,\Sigma_{11})$ (D) $N_6(\mu_2, \Sigma_{22})$

8. Considering the partitioning of Q.7 the correct value of the partial variance – covariance matrix would be

(A) $\Sigma_{11.2} = \Sigma_{22} - \Sigma_{12}\Sigma^{-1}_{22}\Sigma_{21}$ (B) $\Sigma_{11.2} = \Sigma_{22} - \Sigma_{12}\Sigma^{-1}_{11}\Sigma_{21}$

(C) $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma^{-1}_{11}\Sigma_{21}$ (D) $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma^{-1}_{22}\Sigma_{21}$

9.Let $X_1, X_2 \ldots\ldots\ldots X_N$ be a random sample of N vectors from $N_p(\mu, \Sigma)$ then the MLE of Σ when μ is known will be

(A) the sample variance –covariance matrix S (B) depending upon the sample mean vector (C) depending upon the sample mean vector and μ (D) depending upon μ.

10. In usual notations which of the following is the correct expression for the characteristic function of a Wishart distribution?

(A) $1/|I - 2iU\Sigma|$ (B) $1/|I - 2iU\Sigma|^{(N-1)}$ (C) $1/|I - 2iU\Sigma|^{(N-1)/2}$

(D) $1/|I - 2iU\Sigma|^{-(N-1)/2}$

11.In the proposed density of Wishart distribution the value of the constant 'C'

For p=1 is equal to

(A) $1/2^{n/2}$ (B) $1/2^{n/2}\ \Gamma n/2$ (C) $1/\ \Gamma n/2$ (D) $2^{n/2}\ \Gamma n/2$

12. Let the sample variance-covariance matrix 'A' be distributed as $W_p(n,\Sigma)$ then we can obtain chi-square distribution from it by

(A) pre multiplying by a vector L (say) of order px1 (B) post multiplying by a vector L (say) of order px1 (C) pre and post multiplying by a vector L (say) of order px1 (D) none of the above.

PART-2 (18 marks)

1. Attempt ALL questions

2. Figures within parenthesis indicate marks.

1. (A) In a general if the density function of a multivariate normal density is written as

$f(\mathbf{X}) = k\, e^{-1/2}\, (\mathbf{X}-\mathbf{B})'\, A\, (\mathbf{X}-\mathbf{B})$ where **X** and **B** are the vectors of order 'p' and A is positive definite matrix then obtain the value of 'k' so that it represents the pdf of a Multivariate normal density. (3)

1. (B) obtain the characteristic function of a $N_p(\mu, \Sigma)$ distribution. (3)

1. (C) Consider the partitioning of Q.7 then state and prove the necessary condition for the independence of two sub-sets of the vector **X**. (3)

OR

1. (C) Write down the expression of conditional distribution of X^1 given X^2 form this write the expression of the conditional mean and the regression of X^1 on X^2 (3)

2. (A) Let $\mathbf{X}_1,\mathbf{X}_2..........\mathbf{X}_N$ be a random sample of N vectors from $N_p(\mu, \Sigma)$ write down the expression for the likelihood function. What are the MLE^s of μ and Σ when both are unknown.(Do NOT derive) (3)

2. (B) Show that for p=1 the chi-square distribution can be derived from the $W_p(n,\Sigma)$ distribution. (3)

2(C) State and prove the reproductive property of Wishart distribution. (3)

OR

2.(C) Let the matrix A be distributed as $W_p(n,\sum)$ distribution. Then obtain the marginal distribution of any partition of the matrix 'A'. (3)

****

* the problem of classification and Discrimination with the help of a real life example.
* Canonical correalation for a multivariate set up.
* best regression

9 A
11 B
12 A
18 C
19 B
20 D
21