Total No. of Questions : 5]

SEAT No. :

P2493

[Total No. of Pages : 3

[5587]-101
M.Sc. (Semester - I)
STATISTICS
ST - 11 : Mathematical Analysis
(2013 Pattern)

Time : 3 Hours] *[Max. Marks :50*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of scientific calculator and statistical table is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt each of the following :

a) In each of the cases, choose the correct alternative: **[3]**

i) $\lim_{n\to\infty} \log\left(1-\frac{1}{n}\right)^n$ equals

A) 1 B) 0

C) e D) -1

ii) Let f : R→R be a monotone function. Then

A) f has no discontinuities

B) f has only finitely many discontinuities

C) f can have at most countably many discontinuities

D) f can have uncountably many discontinuities

iii) If $\Sigma\, a_n$ is convergent but not absolutely convergent and $\Sigma\, a_n = 0$.

Let s_n be partial sum $\sum_{i=1}^{n} a_i$, n = 1, 2, 3,... Then

A) $s_n = 0$ for infinitely many n

B) $s_n > 0$ for finitely many n

C) $s_n > 0$ for infinitely many n

D) $s_n > 0$ for all but a finite number of values of n

P.T.O.

b) State whether the following statements are true or false. Justify your answer **[3]**

i) Set $E = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right\}$ is closed set

ii) If F = {0, 1}. Then (F, +, .) is field.

iii) Every bounded sequence is convergent.

c) Define the following terms **[4]**

i) Field of real numbers ii) Metric Space

Q2) Attempt any two questions of the following: **[2 × 5 = 10]**

a) If $x \in$ R, $y \in R$ and $x > 0$, then show that there is a positive integer n such that $nx > y$.

b) Let A be a nonempty set of real numbers which is bounded below. Let $-A$ be the set of all numbers of the form $-x$, where $x \in$ A, then prove that inf A = - sup($-A$).

c) Give an example of a set satisfying the condition in each of the following (Specify metric space):

i) Set is neither Closed nor Open set.

ii) Bounded set which is Closed set as well as Open set.

iii) Unbounded set which is Closed set as well as Open set

iv) Bounded set which is Closed but not Open set

v) Unbounded set which is Closed set but not Open set

Q3) Attempt any two questions of the following: **[2 × 5 = 10]**

a) If $S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!}$, for n = 1, 2, 3, ... Prove that $2 < \lim_{n\to\infty} s_n < 3$.

b) State and prove Ratio test for convergence of series.

c) If $\{A_n; n \geq 1\}$ be a collection of countable sets. Then show that countable union of $\{A_n, n \geq 1\}$ is countable set.

Q4) Attempt any two questions of the following: **[2 × 5 = 10]**

a) Consider a function f defined on [a, b]. If f has a local maxima at a point $x \in (a, b)$, and $f'(x)$ exists, then show that $f'(x) = 0$. Also state analogous statement for local minima and prove it.

b) State and prove Fundamental Theorem of Calculus.

c) Obtain the radius of convergence for the following series.

i) $\sum \frac{n^3}{3^n} z^n$ ii) $\sum \frac{2^n}{n!} z^n$

Q5) Attempt any one question of the following:

a) Prove or disprove the following statements: **[2 Each]**

i) Countable Union of Closed set is Closed.

ii) Every convergent sequence is bounded sequence.

iii) If series $\Sigma\, a_n$ is convergent then series $\Sigma\, |a_n|$ is also convergent.

iv) Every neighborhood is an open set.

v) Set of irrational numbers is uncountable.

b) Solve the following examples: **[2 Each]**

i) Find limit superior and limit inferior of $t_n = (-1)^n \left(1 + \frac{1}{n}\right)$.

ii) Find *lub* and *glb* of the set $\{x \in R \mid x^2 - 5x + 6 < 0\}$.

iii) Discuss convergence of the series $\sum \frac{1}{2+3n^2}$.

iv) If I is set of integers, then check whether $(I, +, .)$ is a field.

v) Show that d is metric, if $\underline{x} = (x_1, x_2)$, $\underline{y} = (y_1, y_2)$ are any two points in R^2, define $d(\underline{x},\underline{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$

❖ ❖ ❖ ❖

Total No. of Questions : 3]

SEAT No. :

P2494

[Total No. of Pages : 2

[5587]-102

M.Sc. STATISTICS

ST - 12 : Integral Calculus and Statistical Computing

(2013 Pattern) (Semester - I)

Time : 1½ Hours] *[Max. Marks :25*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of Scientific calculator and statistical table is allowed.***
4) ***Symbols and abbreviations have their meaning.***

Q1) Attempt all the questions given below :

a) Define the terms with an illustration each : Improper integral of type I and type II. **[2]**

b) Explain how Rieman integral is particular case of Riemann-Stiltje integral. **[1]**

c) State the necessary and sufficient condition for the existence of Rieman integral. **[1]**

d) Suggest the value of 'c' so as to get random integers between 0 and 10, using the linear congruential generator is $X_{n+1} = (5X_n + 3) \bmod c$. **[1]**

Q2) Attempt any two of the following :

a) i) If m = inf f(x) and M=sup f(x) for $a \le x \le b$ then show that

$$m(b-a) \le L(p,f) \le U(p, f) \le M(b-a)$$ **[2]**

ii) Evaluate $\int_a^b \frac{dx}{x^2}$ as a limit of Rieman sum when the norm of partition tends to zero. **[3]**

b) Define beta integral and discuss its convergence. **[5]**

P.T.O.

c) i) Explain the Monte-Carlo method of simulation to evaluate

$$\int_0^{\infty} \sin(x) e^{-x} dx.$$ **[2]**

ii) Using Newton-Raphson iterative method. Solve the following simultaneous equations.

$$x^2 + y^2 - 4 = 0$$

$$xy - 1 = 0$$

Take (2, 0) as the initial solution. Carry out two iterations. **[3]**

Q3) Attempt any Two of the following :

a) i) If $\alpha(x)$ is differentiable with respect to x the show that

$$\int_a^b f(x) d\alpha(x) = \int_a^b f(x) \frac{d\alpha(x)}{dx} d(x).$$ **[2]**

ii) Derive trapezoidal rule to evaluate $\int_a^{a+2h} \int_b^{b+2k} f(x, y) dy dx$. **[3]**

b) i) Discuss the convergence of $\int_a^b \frac{dx}{(x-a)^p}$. **[2]**

ii) Describe the Hooke's and Jeeve's method to minimize $f(x, y)$. State its advantage as compared to steepest descent method of minimizing $f(x, y)$. **[3]**

c) i) Show that $\Gamma n = (n-1)\Gamma(n-1)$, hence show that for n non negative integer $\Gamma n = (n-1)!$ **[2]**

ii) The following table gives the values of B(m,n). Interpolate B(1.2, 2.2) using Newton-Raphson method **[3]**

m / n	1	2	3
1	1	1/2	1/3
2	1/2	1/6	1/12
3	1/3	1/12	1/30

❖ ❖ ❖ ❖

Total No. of Questions : 5]

SEAT No. :

P2495

[Total No. of Pages : 3

[5587]-103
M.Sc. - I (Semester - I)
STATISTICS
ST - 13 : Linear Algebra
(2013 Pattern)

Time : 3 Hours] ***[Max. Marks :50***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Attempt any FIVE questions from the following questions: **[2 Marks each]**

a) Given that A is a square matrix of order n, show that A'A = 0 implies that A = 0, where 0 is null matrix.

b) Define rank of a matrix. Is this true that row rank and column rank of a matrix are equal? Justify.

c) Define idempotent matrix. If A and B are two idempotent matrices of order n. Check whether (A–B) is idempotent matrix.

d) Give one example of a set of independent vectors and set of dependent vectors with justification.

e) Define algebraic multiplicity and geometric multiplicity of a matrix with one example each.

f) Suppose V_1 and V_2 are n-dimensional subspaces of a vector space V. Check whether $V_1 \cup V_2$ is a subspace or not.

g) Examine the nature of the following quadratic form.

$$3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$$

P.T.O.

Q2) Attempt any TWO questions from the following questions :**[5 Marks each]**

a) Define an orthogonal matrix. Show that the columns of an orthogonal matrix are linearly independent and converse of this is not true.

b) Define Vector space and its basis with an example and prove that the number of members in any one basis of a subspace is the same in any other basis.

c) Solve the following system of linear equation and obtain all possible solutions, if it is consistent.

$$X_1 + 2X_2 - 3X_3 - 4X_4 = 0$$

$$X_1 + 3X_2 - X_3 + 2X_4 = 0$$

$$2X_1 + 5X_2 - 2X_3 - 4X_4 = 0$$

Q3) Attempt any TWO questions from the following questions:**[5 Marks each]**

a) Prove or disprove that:

i) Subset of linearly dependent set of vectors is linearly dependent.

ii) If $\underline{\eta}$ is linear combination of the set $\{\underline{x}_1,\underline{x}_2,\underline{x}_3,\ldots,\underline{x}_n\}$ then the set $\{\underline{\eta},\underline{x}_1,\underline{x}_2,\underline{x}_3,\ldots,\underline{x}_n\}$ is linearly dependent.

b) Use Gram-Schmidt orthogonalisation procedure on the following vectors and get an orthogonal basis for the vector space spanned by these vectors.

$\alpha_1 = [1, 1, 1, 1]$ $\alpha_2 = [1, 2, 3, 4]$ $\alpha_3 = [1, 2, 2, 1]$ $\alpha_4 = [1, 1, -2, 3]$

c) Prove that for a real symmetric matrix, eigen values are real.

Q4) Attempt any TWO questions from the following questions:**[5 Marks each]**

a) State and prove Cayley-Hamilton theorem.

b) State and prove a necessary and sufficient condition for a quadratic form to be non-negative definite.

c) Obtain a Moore-Penrose generalized inverse of matrix A, where A is

given by $A = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2 & 2 & 4 \\ 3 & 6 & 1 & 2 \\ 4 & 8 & 2 & 4 \end{bmatrix}$.

Q5) Attempt any ONE questions from the following questions :

a) i) Prove or disprove :

A) If A is symmetric matrix then its left eigen vector is same as its right eigen vector.

B) If t is eigen value of matrix A then (t + c) is eigen value of (A + cI), where c is a constant.

ii) Using spectral decomposition of a matrix, obtain A^8 if A is given by

$$A = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

[4+6]

b) i) Prove or disprove : Let A^- is generalized inverse of a matrix A then

A) AA^- is idempotent matrix

B) Rank (A) $\leq$ Rank(A^-)

ii) Prove that if A is a real symmetric matrix of order n then there exists a real orthogonal matrix P such that P'AP is diagonal matrix with diagonal entries $\lambda_1,\lambda_2,.....,\lambda_n$ where $\lambda_1,\lambda_2,.....,\lambda_n$ are eigen values of A.

iii) Prove or disprove: if A is orthogonal matrix then quadratic form corresponding to A is positive definite.

[4+4+2]

❖ ❖ ❖ ❖

Total No. of Questions : 3]

SEAT No. :

P2496

[Total No. of Pages : 3

[5587]-104
M.Sc. (Semester - I)
STATISTICS
ST - 14 : Probability Distributions - I
(2013 Pattern)

Time : 1½ Hours] *[Max. Marks :25*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of scientific calculator and statistical table is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) A) Choose the correct alternative in each of the following. **[3]**

i) If a continuous random variable X has distribution function F(X) then Y = –log(1–F(X)) follows __________.

a) Normal distribution b) Uniform (0, 1)

c) Laplace distribution d) Exponential distribution

ii) A probability generating function (PGF) of a random variable X denoted by $P_X(S)$ is well defined if $\Sigma s^x p(x)$ is convergent for ____.

a) $S \in R$ b) $S < 1$

b) $|S| \leq 1$ d) $S < \infty$

iii) The random variables X and Y are independent if and only if

a) Corr (X, Y) = 0

b) E(XY) = E(X).E(Y)

c) $E(e^{tXY}) = E(e^{tx}).E(e^{ty})$ for all $t \in R$

d) $E[I_{[x \leq t]}. I_{[y \leq s]}] = E[I_{[x \leq t]}] .E[I_{[y \leq s]}]$ for all t and s in R. Where I_A is a indicator function of the set A.

b) If $f(x)$ is a distribution function then show that $[f(x)]^2$ is a distribution function. **[2]**

P.T.O.

Q2) Attempt any two of the following :

a) i) Define random variable having probability distribution symmetric around a constant 'a'.

ii) Let X be a random variable with p.d.f.

$$f(x) = \frac{3x^2}{2} \quad , \text{if } -1 \le x \le 1$$
$$= 0 \quad , \text{otherwise}$$

Examine whether X has symmetric probability distribution around zero. Find E(X). Verify whether mean, mode, median coinside. **[1+4]**

b) Let X be a random variable with distribution function

$$F(x) = 0 \qquad x < 0$$
$$= \frac{x^2}{4} + \frac{1}{5} \qquad 0 \le x < 1$$
$$= x - \frac{x^2}{4} - \frac{1}{5} \qquad 1 \le x < 2$$
$$= 1 \qquad x \ge 2$$

Decompose F(x) as a mixture of discrete and continuous distribution functions. Also find E(X). **[5]**

c) Let (X,Y) be a bivariate r.v. with joint p.m.f **[5]**

$$P(x,y) = \frac{e^{-1}}{y!(x-y)!} P^y (1-P)^{x-y} \text{ if } x = 0, 1, 2, \ldots\ldots\ldots\ldots\ldots\ldots$$
$$y = 0, 1, 2, \ldots\ldots\ldots\ldots\ldots\ldots, x$$

Find i) the marginal distributions of X and Y

ii) the conditional distribution of x given Y=y

iii) E(X|Y=y) and verify that E[E(X|Y=y)] = E(X).

Q3) Attempt any two of the following :

a) If $X_1, X_2, \ldots\ldots, X_N$ are independent and identically distributed Bernoulli(P) random variables and N is Binomial (m, p') then find the probability distribution of $S_N = X_1 + X_2 + \ldots\ldots + X_N$.
State $E(S_N)$ and $var(S_N)$. **[5]**

b) Suppose X and Y are independent and identically distributed $B\left(2, \frac{1}{2}\right)$ random variables. Find the probability generating function of U=X–Y. Hence find the probability distribution of U. Is U symmetric around zero? Justify. **[5]**

c) If (X, Y) is a random vector with joint probability density function.

$$f(x, y) = e^{-(x+y)} \quad ; x > 0, y > 0$$
$$= 0 \quad ; \text{otherwise}$$

obtain the joint moment generating function (m.g.f.) of (X, Y) Hence obtain the m.g.f. of

i) $U = X + Y$

ii) $V = X - Y$

Identify the probability distributions of U and V. **[5]**

❖ ❖ ❖ ❖

Total No. of Questions : 5]

SEAT No. :

P2498

[Total No. of Pages : 3

[5587]-106
M.Sc. (Semester - I)
STATISTICS
ST - 16 : SAMPLING THEORY
(2013 Pattern)

Time : 3 Hours] *[Max. Marks :50*

Instructions to the candidates:

1) *All questions are compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of scientific calculator and statistical table is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) A) Choose the correct alternative for each of the following:

i) The most important factor in determining the size of a sample is **[1]**

a) the availability of resources b) purpose of the survey

c) heterogeneity of population d) Sample frame

ii) The variance of the mean of cluster means in a simple random sample without replacement of n clusters from N clusters of size M each is proportional to **[1]**

a) S_i^2 b) S^2

c) S_w^2 d) S_b^2

B) In each of the following, state whether the given statement is true or false. Justify your answer.

i) The discrepancy between estimates and population parameters is known as sampling error. **[1]**

ii) In the method of stratified sampling with proportional allocation of sample sizes is better than Neyman's allocation for estimating population mean. **[1]**

C) Define the following terms with an illustration.

i) Sampling Frame. **[1]**

ii) Elementary unit. **[1]**

D) i) Differentiate between the methods of SRS and varying probability scheme. **[2]**

ii) State ratio estimator for sample mean. Also explain the situation where ratio estimator is used. **[2]**

P.T.O.

Q2) Attempt any two of the following : **[2 × 5 = 10]**

a) Define inclusion probabilities π_i and π_{ij} for a sample of size n where i and j denote units in the population of N units. Obtain π_i and π_{ij} for Simple Random Sampling Without Replacement (SRSWOR) and Simple Random Sampling With Replacement (SRSWR). Derive the expression for variance of sample mean when SRSWOR is used.

b) Explain cumulative total method and Lahiri's method of selecting a sample of size n by SRSWR with Probability Proportional to Size (PPS) with illustration.

c) Obtain Hurvitz-Thompson (HT) estimator of population mean when Simple Random sampling without Replacement (SRSWOR) with varying probabilities of selection is used. Show that it is unbiased and derive expression for its variance.

Q3) Attempt any two of the following : **[2 × 5 = 10]**

a) Derive an unbiased ratio type estimator for population total. Also find variance of your estimator.

b) Explain the method of 'Post Stratification' giving an illustration. State the estimator of population mean in this method and verify whether it is unbiased estimator of population mean.

c) Find the condition under which systematic sampling is more efficient than SRSWOR of the same size for estimating population mean. Express variance of the mean of a systematic sample in terms of intra class correlation coefficient.

Q4) Attempt any two of the following : **[2 × 5 = 10]**

a) Obtain the efficiency of estimator of population mean in cluster sampling with clusters of equal sizes over corresponding estimator of population mean used in SRSWOR in terms of intra class correlation coefficient.

b) Show that regression estimator for separate strata is more efficient than regression estimator for combined strata if $\beta_1 = \beta$ for all i = 1,2,....,k.

c) Explain Warner's Randomized Response Technique with an illustration.

Q5) Attempt any one of the following :

a) i) Write short note on 'Observational errors and non-sampling errors'. Discuss mathematical model for measurement of observational errors

ii) Explain the methods of drawing a sample by 'Circular systematic sampling' and 'Two dimensional systematic sampling', Compare them. **[5+5]**

b) Derive the expression for construction of strata. **[10]**

❖ ❖ ❖ ❖

Total No. of Questions : 4]

SEAT No. :

P2500

[Total No. of Pages : 2

[5587]-202
M.Sc. (Semester - II)
STATISTICS
ST - 22 : Limit Theorems and Convergences
(2013 Pattern) (3 Credits)

Time : 2¼ Hours] ***[Max. Marks : 38***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of scientific calculator and statistical tables is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Define the following terms : **[2 each]**

a) Independence of two classes.

b) Class of independent events.

c) π - System

d) λ - System

Q2) Attempt any two of the following :

a) $X_n \xrightarrow{P} 0$ if and only if $E\left(\frac{|X_n|}{1+|X_n|}\right) \to 0$ as $n \to \infty$. **[5]**

b) Define convergence in probability and convergence in distribution. Show that convergence in probability implies convergence in distribution. **[5]**

c) Let $\{X_n, n \geq 1\}$ and $\{Y_n, n \geq 1\}$ be sequences of random variables. Let $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{L} C$, then show that **[5]**

i) $X_n + Y_n \xrightarrow{L} X + C$

ii) $X_n Y_n \xrightarrow{L} CX$

P.T.O.

Q3) Attempt any two of the following :

a) If $|X_n| \le Y$ a.s., Y is integrable, then show that $X_n \xrightarrow{P} X$ implies $E(X_n) \to E(X)$. **[5]**

b) Let $P[X_n = 2^{nc}] = 2^{-n}$, $P[X_n = -2^{nc}] = 2^{-n}$, and $P[X_n = 0] = 1-2^{-n+1}$, Then show that $X_n \xrightarrow{P} 0$ but X_n does not converge in mean. **[5]**

c) State and prove Khintchin's Weak Law of Large Numbers (WLLN).**[5]**

Q4) Attempt any one of the following :

a) Solve the following examples :

i) Let $\{X_n, n \ge 1\}$ be iid with finite mean and variance. Then show that $\frac{1}{n}\sum_{i=1}^{n} X_i$ converges to $E(X_1)$ in mean square. **[3]**

ii) Let $\{X_n, n \ge 1\}$ be a sequence of iid random variables with common mean μ and common variance σ^2. Let $S_n = \sum_{i=1}^{n} X_i$ prove that in the long run $E(S_n)$ behaves like a median of S_n. **[3]**

iii) Let X_λ be Poisson with parameter $\lambda > 0$. Show that $\frac{(X-\lambda)}{\sqrt{\lambda}} = Y_\lambda$ tends to N(0, 1) as $\lambda \to \infty$. **[4]**

b) i) State and prove multiplication theorem for n random variable.**[6]**

ii) If $Y \le X_n$; Y is integrable then show that $E \underline{\lim} X_n \le \underline{\lim} E X_n$.**[4]**

❖ ❖ ❖ ❖

Total No. of Questions : 5]

SEAT No. :

P2501

[Total No. of Pages : 3

[5587]-203

M.Sc. (Semester - II)

STATISTICS

ST - 23 : Regression Analysis

(2013 Pattern) (4 Credits)

Time : 3 Hours] *[Max. Marks : 50*

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) Consider a simple linear regression model $y = \beta_0 + \beta_1 x + \in$ with $\in \sim IND\ (0, \sigma^2)$

a) Obtain the least squares estimate $\hat{\beta}_1$. **[2]**

b) Obtain the least squares estimate $\hat{\beta}_0$. **[2]**

c) Show that Cov $\left(\hat{\beta}_0, \hat{\beta}_1\right) = \dfrac{-\bar{x}\sigma^2}{S_{xx}}$. **[2]**

d) Show that E $(MS_{Res}) = \sigma^2$. **[2]**

e) Show that Cov $\left(\bar{y}, \hat{\beta}_1\right) = 0$. **[2]**

Q2) Attempt any two questions :

a) Consider the simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \in_i$, $i = 1,2,....n$ with $\in_i \sim NID\ (0, \sigma^2)$. Obtain the $(1-\alpha)100\%$ confidence intervals on β_0, β_1 and σ^2. **[5]**

P.T.O.

b) Consider the multiple linear regression model $y = X\beta + \in$ with $\in \sim NID(0, \sigma^2 I)$ **[5]**

i) Prove that $\sum_{i=1}^{n} \mathrm{Var}(\hat{y}_i) = \rho\sigma^2$

ii) Prove that $\mathrm{Var}(\hat{y}) = \sigma^2 H$

c) Consider the simple linear regression model $y_{ij} = \beta_0 + \beta_1 x_i + \in_{ij}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n_i$ with $\in_{ij} \sim NID(0, \sigma^2)$. Obtain the test of lack of fit. **[5]**

Q3) Attempt any two questions :

a) Explain the Gauss-Newton iteration method of parameter estimation in nonlinear regression models. **[5]**

b) Explain the logistic regression and obtain the fitted logistic regression model. **[5]**

c) Prove that the residual sum of squares for ridge regression is

$$SS_{Res} = (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\hat{\beta}_R - \beta)'X'X(\hat{\beta}_R - \hat{\beta}).$$ **[5]**

Q4) Attempt any two questions :

a) Describe the Box-Cox transformation method. **[5]**

b) Explain the concept of orthogonal polynomial and obtain first five orthogonal polynomial. **[5]**

c) Consider the multiple linear regression model $y = X\beta + \in$ with $\in \sim NID(0, \sigma^2 I)$. Obtain the measure of DFFITS and DFBETAS. Also, obtain the relationship with R-student residual of each measure. **[5]**

Q5) Attempt any one question :

a) i) State and prove Gauss-Markov theorem. **[6]**

ii) Consider the maximum-likelihood estimator $\tilde{\sigma}^2$ of σ^2 in the simple linear regression model. Compute the amount of bias in $\tilde{\sigma}^2$. **[4]**

b) i) Consider the multiple linear regression model **[5]**

$y_{ij} = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \in_{ij}$, i = 1, 2, 3 j = 1,2,......n with $\in_{ij} \sim NID(0, \sigma^2)$. Suppose that the indicator variables x_1 and x_2 are defined as

$$x_1 = \begin{cases} 1, \text{if observation is from treatment } 1 \\ -1, \text{if observation is from treatment } 3 \\ 0, \quad \text{otherwise} \end{cases}$$

$$x_2 = \begin{cases} 1, \text{if observation is from treatment } 2 \\ -1, \text{if observation is from treatment } 3 \\ 0, \quad \text{otherwise} \end{cases}$$

Show that the relationship between the parameters in the regression and analysis of variance model is

$$\beta_0 = \frac{\mu_1 + \mu_2 + \mu_3}{3} = \bar{\mu}$$

$$\beta_1 = \mu_1 - \bar{\mu}$$

$$\beta_2 = \mu_2 - \bar{\mu}$$

ii) Consider the Generalized linear regression model $y = f(x, \beta) + \in$. Where $f(x, \beta)$ is continuously differentiable function of β. Obtain the $(1-\alpha)100\%$ confidence interval on the true response at the point x_0. **[5]**

❖ ❖ ❖ ❖

Total No. of Questions : 5]

SEAT No. :

P2502

[Total No. of Pages : 3

[5587]-204
M.Sc. (Semester - II)
STATISTICS
ST - 24 : Parametric Inference (Estimation)
(2013 Pattern) (4 Credits)

Time : 3 Hours] ***[Max. Marks :50***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) a) Define the following terms. Give one example each **[1 each]**

i) Complete statistic

ii) Loss function

b) Define Fisher Information Matrix. State the regularity conditions. **[2]**

c) Define sufficient statistic. State factorisation criterion for a sufficient statistic. **[2]**

d) State Bhattacharya lower bound of order k for the variance of an unbiased estimator of parameter θ. Also state the regularity conditions to be held. **[2]**

e) What is pivotal quantity? Explain its use in construction of confidence interval. **[2]**

Q2) Attempt any two of the following: **[5 each]**

a) When the parametric function $\psi(\theta)$ is said to be estimable? Let random variable X follow truncated Poisson distribution with parameter θ truncated to the left at $x = 0$. Let (i) $\psi(\theta) = \theta$ and (ii) $\psi(\theta) = \frac{1}{\theta}$. Check estimability of $\psi(\theta)$ in (i) and (ii)

b) Define minimal sufficient statistic. Show that likelihood equivalence leads to minimal sufficient statistic.

c) Let $X \sim N(\theta, 1)$. Let the prior distribution of θ be $N(\mu, 1)$, where μ is known. Obtain posterior distribution of θ.

P.T.O.

Q3) Attempt any two of the following : **[5 each]**

a) Let $X_1, X_2, \ldots\ldots, X_n$ be i.i.d. $B(1, \theta)$. Find unbiased estimators of (i) $\psi(\theta) = \theta$ and (ii) $\psi(\theta) = \theta\,(1-\theta)$. Further, Rao-Blackwellise each to obtain MVUE.

b) Let $\{f(x, \theta), \theta \in \Theta\}$ be a family of probability density functions such

that $$f(x, \theta) = \frac{u(x)}{v(\theta)}, \quad a(\theta) < x < b(\theta)$$
$$= 0, \quad \text{Otherwise}$$

Find minimal sufficient statistic for θ, if

i) $a(\theta) = a$ (constant)

ii) $a(\theta)$ is increasing and $b(\theta)$ is decreasing function of θ.

c) Let $X_1, X_2, \ldots\ldots, X_n$ be a random sample from $\{f(x, \theta), \theta \in \Theta\}$. Define uniformly most accurate confidence bounds for the parameter θ. Let $X \sim U(0, \theta)$. Obtain uniformly most accurate confidence bounds for θ of level $(1-\alpha)$.

Q4) Attempt any two of the following : **[5 each]**

a) Let $$f(x, y, \lambda, p) = \binom{x}{y} p^y (1-p)^{x-y} \frac{e^{-\lambda}\lambda^x}{x!},$$
$$y = 0, 1, 2, \ldots\ldots\ldots, x;$$
$$x = 0, 1, 2, \ldots\ldots\ldots;$$
$$\lambda > 0;\ 0 < p < 1$$
$$= 0, \quad \text{otherwise}$$

be joint p.m.f. of (X, Y). Show that it belongs to two parameter exponential family. Also obtain minimal sufficient statistic $(T_1, T_2)'$ for $(p, \lambda)'$.

b) Let $X_1, X_2, \ldots\ldots, X_n$ be i.i.d. $N(\theta, 1)$. Construct shortest expected length confidence interval of level $(1-\alpha)$ for the parameter θ.

c) State and prove Basu's theorem.

Q5) Attempt any one of the following :

a) i) State and prove Rao-Blackwell theorem. **[6]**

ii) Let $X_1, X_2, \ldots\ldots, X_n$ be i.i.d. with p.d.f. given by $f(x, \theta) = \dfrac{\theta}{x^{\theta+1}}$, $x \geq 1; \theta > 0$.

Find sufficient statistic using likelihood equivalence. **[4]**

b) i) State and prove necessary and sufficient condition for existence of MVUE. Show that MVUE is unique if it exists. **[7]**

ii) Let $X_1, X_2, \ldots\ldots, X_n$ be i.i.d. $P(\theta)$. Check whether MVBUE exists if $\psi(\theta) = e^{-\theta}$. **[3]**

❖ ❖ ❖ ❖

Total No. of Questions : 2]

SEAT No. :

P2503

[Total No. of Pages : 2

[5587]-205
M.Sc. (Semester - II)
STATISTICS
ST - 25 : Testing of Hypothesis
(2013 Pattern) (Credit System)

Time : 1 Hour] *[Max. Marks :13*

Instructions to the candidates:

1) *Question 1 is compulsory.*
2) *Figures to the right indicate full marks.*
3) *Use of calculator and statistical tables is allowed.*
4) *Symbols and abbreviations have their usual meaning.*

Q1) a) Using suitable examples, illustrate the terms, level of a test and size of a test. **[1]**

b) Define Monotone Likelihood Ratio (MLR) property. Show that {poisson (θ), $\theta > 0$} satisfies this property. **[2]**

Q2) a) Let X be a random variable with p.m.f. under H_0 and H_1 given by **[3]**

x	1	2	3	4
$P(x\|H_0)$	0.45	0.05	0.05	0.45
$P(x\|H_1)$	0.2	0.35	0.25	0.2

consider (i) $\phi_1(x) = 1$......if $x = 2$ (ii) $\phi_2(x) = 1$......if $x = 3$

$= 0$ ....o.w. $= 0$ ....o.w.

Find size and power of these tests. Which one will you prefer? Justify.

b) Define a Most Powerful (MP) test. Show that for a MP level α test, power $\geq$ level. **[3]**

OR

c) Show that any test $\phi_k(\underline{x})$ of the form

$$\phi_k(\underline{x}) = \begin{cases} 1 \text{.........} if & L_1(\underline{x}) > k\, L_0(\underline{x}) \\ \gamma(\underline{x}) \text{......} if & L_1(\underline{x}) = k\, L_0(\underline{x}) \\ 0 \text{.........} if & L_1(\underline{x}) < k\, L_0(\underline{x}) \end{cases}$$

P.T.O.

for some $0 \leq k < \infty$ and $0 \leq \gamma(x) \leq 1$, is MP test of size $E(\phi_k(\underline{x})|H_0)$ for testing

H_0 : $\underline{X}$ has probability distribution $L_0(\underline{x})$ against

H_1 : $\underline{X}$ has probability distribution $L_1(\underline{x})$. **[4]**

OR

a) Let $x_1, x_2, \ldots\ldots. x_n$ be a random sample from $\{f(x, \theta), \theta \in \Theta\}$ which belong to one parameter exponential Family. Obtain uniformly Most Powerful (UMP) level α test for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$. **[6]**

b) Show that power function for the test in (a) above is monotone. **[4]**

❖ ❖ ❖ ❖

Total No. of Questions : 3]

SEAT No. :

P2504

[Total No. of Pages : 2

[5587]-206
M.Sc. (Semester - II)
STATISTICS
ST - 26 : Exploratory Multivariate Analysis
(2013 Pattern) (2 Credits)

Time : 1½ Hours] ***[Max. Marks :25***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of scientific calculator and statistical table is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) A) Choose the correct alternative for each of the following: **[3]**

i) Dendogram plot is used for

a) Making clusters b) Analyze clusters

c) Both (a) and (b) d) Discriminating observations.

ii) Consider two statements

I) Principle components are orthogonal to each other.

II) The variation explained by first principal component is maximum.

a) I true II false b) I false II true

c) both true d) both false

iii) A factor loading is

a) empirically based hypothetical variable consisting of items which are strongly associated with each other.

b) the correlation between binomial variable and variable which has a continuous distribution of scores.

c) the correlation of a variable with a whole score.

d) a correlation coefficient between a variable and a factor.

B) State whether following statements are **TRUE** or **FALSE** : **[2]**

i) Rotation technique is used to provide a simpler and interpretable picture of the relationships between factors and variables.

ii) Cluster analysis does not classify variables dependent and independent.

P.T.O.

Q2) Attempt **any two** of the following : **[2 × 5 = 10]**

a) Let random vector $\underline{X}' = [X_1,X_2,X_3,X_4]$ has mean vector $\underline{\mu} = (3,2,-2,0)$

and $\sum = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ Let $A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}$

i) Find E $(A\underline{X})$

ii) Find variance covariance matrix of $A\underline{X}$.

b) Define orthogonal factor model (state assumption clearly). Explain the terms communalities and Specific variances.

c) Let $\underline{X}$ of order $p\times1$ has mean μ and covariance matrix Σ. Obtain first k-principal components of standardized vector $\underline{Z}$ of $\underline{X}$, where $k < p$.

Q3) Attempt **any two** of the following : **[2 × 5 = 10]**

a) Differentiate hierarchical and non-hierarchical cluster analysis. Explain k-means clustering method in brief.

b) Explain the following terms

i) Factor loadings

ii) Factor scores

iii) Sample variance

c) Define canonical correlations and canonical variables. Derive the first canonical correlation coefficient and the corresponding canonical variables. Also show that first canonical correlation explains maximum correlation.

❖ ❖ ❖ ❖

Total No. of Questions : 4]

SEAT No. :

P2505

[Total No. of Pages : 3

[5587]-207

M.Sc. (Semester - II)

STATISTICS

ST - 27 : Inference in Multivariate Analysis

(2013 Pattern)

Time : 2¼ Hours] ***[Max. Marks :38***

Instructions to the candidates:

1) ***All questions are compulsory.***
2) ***Figures to the right indicate full marks.***
3) ***Use of statistical tables and scientific calculator is allowed.***
4) ***Symbols and abbreviations have their usual meaning.***

Q1) State whether the following statements are true or false, with justification. **[2 each]**

a) A joint vector constructed from any two univariate normal random variables has bivariate normal distribution.

b) If $X \sim N_p(\mu, \Sigma)$ with $|\Sigma| \neq 0$, then the maximum likelihood estimate of Σ is always non-singular.

c) The distribution of Hotelling's T^2 test statistic is multivariate normal.

d) The low value of the likelihood ratio test statistic results in the rejection of the null hypothesis.

Q2) Attempt any two questions from the following:

a) Observations are to be classified into two multivariate normal populations, $N_p(\mu_1, \Sigma)$ and $N_p(\mu_2, \Sigma)$ using Fisher's discriminant function. Compute the Mahalanobis distance between the two populations. Classify the new observations $X = (-1, 1)'$ and $Y = (1, -1)'$ by using Fisher's discriminant function. The parameters of the two population are given by, $\mu_1 = (0.1, -0.1)'$, $\mu_2 = (-0.1, 0.1)'$ and

$\Sigma = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. **[5]**

P.T.O.

b) Derive the test procedure for testing the equality of two variance-covariance matrices. Clearly state the assumptions. State one application of this procedure (i.e., describe a situation where such a test is needed). **[5]**

c) The daily maximum temperature was recorded for four metro cities in India for a period of one year. Derive a likelihood ratio based test procedure to examine if all the daily maximum temperatures from the four cities are uncorrelated with each other and have a common variance. Clearly state your assumptions. **[5]**

Q3) Attempt any two questions from the following.

a) Show that for a sample from a multivariate normal population, sample mean $\bar{X}$ and the corrected sum of squares and sum of products (CSSSP) matrix S are independent. State the major results that you use in the proof. **[5]**

b) Let X_1, X_2, X_3 denote sales of three different types of vehicles. Further, let X_4, X_5, X_6 denote the corresponding sales of the same categories after launching an attractive loan scheme. The management is of the opinion that the new loan scheme has increased the sales by 20%. Construct an appropriate null hypothesis to test this claim and describe the test procedure for testing the same. Clearly state the results and the assumptions that you need to use for this derivation. **[5]**

c) Define Wishart distribution. State and prove its two properties. **[1+4]**

Q4) Attempt any one question out of the following.

a) i) Define Multivariate Normal distribution. **[2]**

ii) If $X \sim N_p(\mu, \Sigma)$, derive the distribution of $Y = AX$, where A is a $q \times p$ matrix. **[2]**

iii) Let X_1, X_2, X_3 and X_4 be independent $N_p(\mu, \Sigma)$ random vectors.

- Find the marginal distributions of each of the random vectors **[3]**

$$V_1 = \frac{1}{4}X_1 - \frac{1}{4}X_2 + \frac{1}{4}X_3 - \frac{1}{4}X_4$$

$$V_2 = \frac{1}{4}X_1 + \frac{1}{4}X_2 - \frac{1}{4}X_3 - \frac{1}{4}X_4$$

- Find the joint density of V_1 and V_2. **[3]**

b) i) Let $X = (X^{(1)}, X^{(2)}) \sim N_p(\mu, \Sigma)$, where $X^{(1)} \in \mathbb{R}^k$ and $X^{(2)} \in \mathbb{R}^{p-k}$ Define

$X^{(2.1)} = X^{(2)} - \Sigma_{21}\Sigma_{11}^{-1}X^{(1)}$, where $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ with Σ_{11}, Σ_{22} being square symmetric matrices of orders k and p – k respectively. Derive the distribution for $X^{(2.1)}$. Further, prove that the two vectors $X^{(2.1)}$ and $X^{(1)}$ are independent. **[5]**

ii) Describe any method for examining the hypothesis of multivariate normality in the given data set. **[5]**

❖ ❖ ❖ ❖