INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1993-94
SEMESTRAL-I EXAMINATION
PROBABILITY THEORY IM

Date: 15.11.93 Maximum Marks: 100 Time: 3 Hours

Note: Answer all the questions.

1. Let $x_{11} \dots x_{1m_1}$ and $x_{21} \dots x_{2m_2}$ be two independent samples of size m_1 and m_2 respectively from a $N(\mu,\sigma^2)$ population.

   Define $s_1^2 = \sum_{i=1}^{m_1} (x_{1j}-\bar{x}_1)^2/(m_1-1)$ and $s_2^2 = \sum_{j=1}^{m_2}(x_{2j}-\bar{x}_2)^2/(m_2-1)$

   where $\bar{x}_1 = \sum_{i=1}^{m_1} x_{1i}/m_1 \quad \bar{x}_2 = \sum_{j=1}^{m_2} x_{2j}/m_2$.

   Then show that s_1^2/s_2^2 has a F distribution with (m_1-1, m_2-1) d.f. and derive the corresponding density function.

   [Hint: express s_1^2/s_2^2 in terms of two independent gamma variables] [18]

2. Let $X_{(1)} \dots X_{(n)}$ be the order statistics based on an iid sample of size n from $U(0,1)$ distribution. Derive the density function of the order statistics and show that it is uniformly distributed over the region $E = \{(x_1 \dots x_n) : 0 \le x_1 \le \dots \le x_n \le 1\}$. [15]

3. a) Show that countable union of negligible sets is negligible.

   b) For every fixed n, let $x_1 \dots x_k$ denote the number of elements in a multinomial distribution with k cells $x_1+\dots+x_k = n$. Let n be a poisson (λ) variable. Then show that $x_1 \dots x_k$ are independently distributed. Are these independent when n is fixed?

   c) Let there be (N+1) urns each containing N balls. Urn number k contains k red and (N-k) white balls, (k=0, 1, 2, ...N). An urn is chosen at random and n balls are drawn with replacement from it. Find the probability that (n+1)th draw is a red ball given that the previous balls drawn were red. [3+3+5 = 11]

4. Explain how hypergeometric distribution can be used in a catch - recatch method to estimate the number of unknown fishes in a lake. [10]

p.t.o.

5. Let $X_{(1)} = \min \{X_1, \ldots X_n\}$

and $X_{(n)} = \max \{X_1, \ldots X_n\}$

where $X_1, X_2 \ldots X_n$ are iid $U(0,1)$ variables.

Find $P\{X_{(1)} > a, X_{(n)} < b\}$ where $0 < a < b < 1$.

Hence, or otherwise show that the variables $n X_{(1)}$ and $n(1 - X_{(n)})$ are asymptotically independent.

[Hint: compute $\lim_{n \to \infty} P[n X_{(1)} > c, n(1-X_{(n)}) > d]$]. [10]

6. A group of 2N boys and 2N girls is divided into two equal groups. Find the probability p that each group will be equally divided into boys and girls. Estimate p using Stirling's formula. [7]

7. A man is given n keys of which only one fits his door. He tries them successively (sampling without replacement). This procedure may require 1, 2, ... n trials. Show that each of these outcome has probability $1/n$. [10]

8. Explain the following terms.
   1) Bose-Einstein Statistics.
   2) Multivariate normal distribution.
   3) WLLN. [10]

9. Let X_1, X_2 be iid $U(0,1)$ variable. Find the density of X_1+X_2. [9]

1993-94 524

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1993-94
SEMESTRAL-I EXAMINATION
MATHEMATICAL ANALYSIS IM

Date: 17.11.93 Maximum Marks: 100 Time 4 Hours

Note: a) Give complete answers. You must state any theorem (proved in the class) that you use.
b) The paper carries 117 marks. You may answer as many questions as you wish. Maximum score 100.

1.(a) Show that every open cover of $[a,b]$ admits a finite subscover.

(b) If $f_n : \mathbb{R} \to [0,1]$ is uniformly covergent and $g: [0,1] \to \mathbb{R}$ continuous then show that $g \circ f_n$ is uniformly convergent. [10+7]

2.(a) Find the upper and lower limits of the sequence $\{x_n\}$ defined by

$$x_1 = 0,\ x_{2n} = \frac{x_{2n-1}}{2},\ x_{2n+1} = \frac{1}{2} + x_{2n}.$$

(b) Test for the convergence of the series

$$\Sigma\ n^p \left(\frac{1}{\sqrt{n-1}} - \frac{1}{\sqrt{n}} \right)$$

where p is a real number.

(c) Determine the set of all complex numbers z for which

$\Sigma\ (2^n/n!)\ z^n$ is convergent. [7+7+7]

3.(a) In each of the following cases, determine the intervals in which the function f is increasing or decreasing and find the maxima and minima (if any)

i) $f(x) = 2x^3 - 3x^2 + 6x - 10$

ii) $f(x) = 3x^4 - 4x^3 - 36x^2 - 1$

(b) Let $f : (0,1) \to \mathbb{R}$ be a map s.t $f'(x)$ exists for all $x \neq 1/2$ and $\lim_{x \to 1/2} f'(x)$ exists. If f differentiable at $1/2$? Justify your answer. [(7+7)+7]

p.t.o.

4.(a) Show that the power series $\Sigma a_n z^n$ is continuous on its disc of convergence.

(b) Let $f = u + iv$ be a non-constant map from $\mathbb{C}$ to $\mathbb{C}$ such that $u+v$ is a constant. Can f be holomorphic? Justify your answer. [10+7

5.(a) If $S \subseteq \mathbb{R}^n$ is open, $f : S \rightarrow \mathbb{R}$ has all the first order partial derivatives on S and these are continuous at some $\bar{X} \in S$ then show that f is differentiable at $\bar{X}$.

(b) If $f,g : \mathbb{R}^n \rightarrow \mathbb{R}$ are differentiable at $\bar{X}$ then show that so is $f.g$. [10+7

6.(a) State Inverse function and Implicit function Theorems.

(b) Deduce the implicit function theorem from the inverse function theorem. [4+10

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-stream) I YEAR:1993-94
SEMESTRAL-I EXAMINATION
COMPUTATIONAL TECHNIQUES AND PROGRAMMING

Date: 19.11.93 Maximum Marks: 100 Time: 3 Hours.

Note: The questions carry 110 marks. Maximum you can score is 100. Answer as much as you can.

1. Write short notes on any three: (about 150 words)

   a) Indexed sequential file organisation
   b) Computer storage area
   c) Queue data structure
   d) Linked list [3x5=15]

2.(a) Derive Newton's formula for forward interpolation.

(b) Find $\log_{10}\pi$ from the following table:

log 3.141	=	0.4970679364
log 3.142	=	0.4972061807
log 3.143	=	0.4973443810
log 3.144	=	0.4974825374
log 3.145	=	0.4976206498

Assume $\pi = 3.1415926536$ [8+12=20]

3.(a) Describe Newton-Raphson method for finding real root of an equation $f(x)=0$. Give the geometric interpretation of it.

(b) Find the real root of

$$3x - \cos x - 1 = 0,$$

by Newton-Raphson method. [10+5=15]

4.(a) Given a new term X (X being the information part), write an algorithm that inserts X into a single-link linear list so that it preserves the ordering of the terms in incresing order of their information (INFO) fields. Clearly describe your symbols, notations and steps.

(b) Given X and FIRST, pointer variables whose values denote the address of a node in a single-link linear list and the address of the first node in the list, respectively, write down an algorithm to delete the node addressed by X. Clearly describe your symbols, notations and steps.

(c) Give a suitable scheme for representing a polynomial by a single-link linear list. Explain your scheme with an example. [10+10+5=25]

p.t.o.

5.(a) Derive Simpson's $\frac{1}{3}$rd rule for integration

(b) Compute the value of $\int_{4}^{5.2} \ln x \, dx$ by the above rule.

Divide the interval of integration into six equal parts each of width = 0.2.

(c) Write down the algorithm for Monte Carlo technique for single variable integration.

(d) Write a complete FORTRAN program for the algorithm in (c) above.

[7+8+5+15=35]

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1993-94
SEMESTRAL - I EXAMINATION

Theory and Methods of Statistics I

Date: 22.11.1993 Maximum Marks: 100 Time: $3\frac{1}{2}$ Hours

Note: Answer any FIVE questions. All questions carry equal marks.

1.(a) Let $X \sim Bin(2s, \frac{1}{2})$. Show that if t represents the mode of X, then

$$P(X = t) = \frac{(2s-1) \times (2s-3) \ldots\ldots \times 3 \times 1}{2s \times (2s-2) \ldots\ldots \times 4 \times 2}$$

(b) X is a random variable for which the rth factorial moment defined by $E[X(X-1)(X-2) \ldots\ldots (X-r+1)] = \theta^r$, $r = 1, 2, 3, \ldots\ldots$, where θ is a positive constant.

Find the moment generating function of X. What is the distribution of X ? Justify your answer.

(c) For the random variable X in (b), show that the mean deviation about the mean is $2e^{-\theta}\,\theta^{[\theta]+1}/[\theta]!$, where $[\theta]$ indicates the largest integer $\leq \theta$. Use Stirling's approximation to show that in this case $\frac{\text{mean deviation}}{\text{standard deviation}} \longrightarrow \sqrt{\frac{2}{\pi}}$ as $\theta \longrightarrow \infty$.

Give your comments.

(6+7+7) = [20]

2.(a) Let X be a random variable with distribution function

$$F(x) = \left\{1 + \exp(-(x-\alpha)/\beta)\right\}^{-1}, \quad \beta > 0.$$

Defining $Y = (X - \alpha)/\beta$, show that

$$E(|Y|^r) = 2\,\Gamma(r+1)(1 - 2^{-(r-1)})\,\zeta(r), \quad r > 1$$

$$= 2\,\Gamma(r+1)\sum_{j=1}^{\infty}(-1)^{j-1}\,j^{-r}, \quad r > 0,$$

where $\zeta(r) = \sum_{j=1}^{\infty} j^{-r}$ is Riemann zeta function of order r.

Hence find the mean deviation of X about its mean and variance of X. Also find β_1 and β_2 coefficients of X.
[You can assume $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$].

contd..... 2/-

(b) Consider the random variable X with p.d.f.

$$f(x) = b \exp[-b(x-a)], \quad a < x < \infty$$
$$= 0, \quad \text{otherwise},$$

where $b > 0$.

Find the median of X.

(c) Suppose X is a non-negative random variable whose mean exists. Then show that

$$E(X) = \int_0^\infty [1 - F(x)]\, dx .$$

(10+5+5) = [20]

3.(a) Show that the moments for the two densities of a positive random variable X, viz.,

$$\text{(i)}\ p_1(x) = (2\pi)^{-1/2}\, x^{-1} \exp[-(\log x)^2/2]$$

and $$\text{(ii)}\ p_2(x) = p_1(x)\,[1 + a \sin(2\pi \log x)],\ -1 \leq a \leq 1$$

are the same.

(b) Suggest a distribution which has no moment of any positive order existing i.e., for which, for every real $r > 0$, $E(X^r)$ does not exist.

(c) Find the moment generating function of χ^2 with n d.f.

Considering appropriate series expansions find γ_1 values for (i) $(\chi^2)^{1/2}$ and (ii) $(\chi^2)^{1/3}$, correct to order $1/\sqrt{n}$. Comment on the relative rates of approaches to normality for the variables, χ^2, $(\chi^2)^{1/2}$ and $(\chi^2)^{1/3}$.

(5+5+10) = [20]

4.(a) y_1 and y_2 are independently identically distributed random variables, each following a χ^2 distribution with n d.f.

Show that $\frac{\sqrt{n}}{2} \cdot \frac{y_1 - y_2}{\sqrt{y_1 y_2}}$ follows a t distribution with n d.f.

(b) $X_i \sim G(\alpha, p_i)$, $i = 1, 2, \ldots n$. X_i's are independent.

Find the joint distribution of

$$y_i = \frac{X_i}{X_1 + X_2 + \ldots + X_i}, \quad i = 2, 3, \ldots n$$

and $$S = X_1 + X_2 + \ldots + X_n .$$

contd..... 3/-

(c) Let $X \sim B(a,b)$ and $Y \sim B(c,d)$. X and Y are independent. Find the distribution of XY when $a = c + d$.

(7+7+6) = [20]

5.(a) Let $X_i \sim N(\delta_i, 1)$, $i = 1, 2, \ldots n$. X_i's are independent. Find the distribution of $Y = \sum_{i=1}^{n} X_i^2$.

(b) If x_i's $(i = 1, 2, \ldots n)$ constitute a random sample from $N(\mu, \sigma^2)$, show that

(i) $\bar{x}$ and s^2 are independently distributed.

(ii) $\sqrt{\frac{n}{n-1}}\,(x_1 - \bar{x}) \Big/ \sqrt{\left\{(n-1)s^2 - \frac{n}{n-1}(x_1 - \bar{x})^2\right\}/(n-2)}$

follows a t distribution with df = n-2.

(10+10) = [20]

6.(a) What is a Hermite polynomial of degree n ? If $H_n(x)$ is a Hermite polynomial of degree n, then prove that

(i) $DH_n(x) = n\,H_{n-1}(x)$, $n \geq 1$

and (ii) $\int_{-\infty}^{\infty} H_n(x)\,H_m(x)\,\phi(x)\,dx = 0$, if $m \neq n$
$= m!$, if $m = n$,

where $\phi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$.

(b) Let X be a standardized random variable with p.d.f. f(x), $-\infty < x < \infty$.

Assuming $f(x) = \sum_{i=0}^{\infty} c_i\,\phi_i(x)$, derive Edgeworth's form of Type A series, viz.,

$$f(x) = \exp\left[-\frac{K_3 D^3}{3!} + \frac{K_4 D^4}{4!} + \ldots\ldots\right] \phi(x),$$ where K_n is

the cumulant of order n of X.

(c) Let t follow t distribution with n d.f. Find Edgeworth expansion form for the distribution function of t correct to order 1/n.

(7+6+7) = [20]

:bcc:

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-Stream) I YEAR: 1993-94
SEMESTRAL-I EXAMINATION
LINEAR ALGEBRA

Date: 24.11.93 Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

Note: Answer any two questions. Marks allotted to each question are given within parentheses.

1.(a) Define similarity of matrices. Show that similar matrices have the same determinant, same characteristic equation and same characteristic roots. Do they have the same characteristic vectors?

(b) Show that the following matrix A is similar to a diagonal matrix D

$$A = \begin{bmatrix} 3 & -5 & -4 \\ -5 & -6 & -5 \\ -4 & -5 & 3 \end{bmatrix}.$$

(10+15)=[25]

2.(a) Find a basis for the row space, a basis for the column space and the rank of the matrix A where

$$A = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -2 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -3 & 6 & 6 & 3 \\ 5 & -3 & 10 & 10 & 5 \end{bmatrix}.$$

2.(b) Also find a conditional inverse A^* of the matrix A.

Obtain the characteristic roots of the matrix $(I - AA^*)$.

(10+15)=[25]

3.(a) If $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, show that

$$|\Sigma| = |\Sigma_{11}| \cdot |\Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}|$$

where Σ_{11} is non-singular.

(b) Show that

$$(I + MN)^{-1} = I - M(I + NM)^{-1}N$$

(c) Show that

$$|I - MN| = |I - NM|.$$

(8+12+5) = [25]

p.t.o.

4. Find a real non-singular linear transformation that reduces the quadratic forms

$$6x_1^2+3x_2^2+14x_3^2+4x_2x_3+18x_3x_1+4x_1x_2,$$

$$2x_1^2+5x_2^2+2x_3x_1+4x_1x_2$$

simultaneously to the forms

$$\xi^2 + \eta^2 + \zeta^2,$$

$$\lambda\xi^2 + \mu\eta^2 + \nu\zeta^2$$

respectively, where λ, μ and ν are constants. [25]

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1993-94
SEMESTRAL - I EXAMINATION

Correlation and Regression

Date: 25.11.1993 Maximum Marks: 50 Time: $1\frac{1}{2}$ hours

EITHER

1. Let (X_i, Y_{ij}), $j = 1,2, \ldots, n_i$, $i = 1,2, \ldots, k$ be $n = n_1 + n_2 + \ldots + n_k$ pairs of observations. Let

$$n_i\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}, \quad i = 1,2, \ldots, k.$$

(a) Express the correlation coefficient using the n pairs of observations (X_i, Y_{ij}) in a simplified form.

(b) Express the correlation coefficient using the n pairs of observations $(\bar{Y}_i, Y_{ij})$ in a simplified form; show that it is always non-negative.

(c) Show that

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i)^2 \geq \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - a - bX_i)^2$$

for any a and b.

(d) Hence, or otherwise, establish an inequality between the correlation coefficients computed in (a) and (b).

(4+5+3+6) = [18]

OR

2. Show that

(a) $$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

(b) $$1-R_{1.2\ldots p}^2 = (1-r_{12}^2)(1-r_{13.2}^2) \ldots\ldots (1-r_{1p.2\ldots \overline{p-1}}^2)$$

(Symbols have their usual meanings.)

(9+9) = [18]

EITHER

3. Define rank correlation. Derive an expression for the Spearman's rank correlation coefficient when ranks are not tied.

[12]

contd..... 2/-

OR

4. Let $(X_{1\alpha}, X_{2\alpha}, \ldots, X_{p\alpha})$, $\alpha = 1, 2, \ldots, n$ be n observations on $(X_1, X_2, \ldots, X_p)$. Also let

$$\sum_{\alpha=1}^{n} (X_{1\alpha} - \alpha - \beta_{12.3\ldots p} X_2 - \cdots - \beta_{1p.2\ldots\overline{p-1}} X_p)^2 \geq \sum_{\alpha=1}^{n} E_{1\alpha}^2$$

for any α, $\beta_{1i.23\ldots\overline{i-1}\,\overline{i+1}\ldots p}$, $i = 2, \ldots, p$, where

$$E_{1\alpha} = X_{1\alpha} - a - b_{12.3\ldots p} X_{2\alpha} - \cdots - b_{1p.2\ldots p} - X_{p}.$$

Compute the correlation coefficient using the n pairs $(X_{1\alpha}, E_{1\alpha})$ $\alpha = 1, 2, \ldots, n$, and state the use of this correlation.

[12]

EITHER

5.(a) The correlation coefficient between the heights of father and his eldest son was observed to be 0.14 on the basis of a sample of size 38. Test if the heights are significantly correlated.

(b) In an attempt to fit the linear regression equation

$$Y = \alpha + \beta X$$

on the basis of 16 pairs of observations (X,Y) from a bivariate normal population, the following statistics were computed.

Sample means : $\overline{X} = 2.33$, $\overline{Y} = 0.75$

Sample variances : $S_{xx} = 4.10$, $S_{yy} = 3.20$

Sample covariance : $S_{xy} = 2.85$.

Construct a 95% confidence interval for β.

(8+12) = [20]

OR

6. The following matrix was obtained from the sample variance - covariance matrix of (X_2, X_3, X_4, X_1) by the method of pivotal - condensation. The entries in the parentheses on the diagonal and below the diagonal are the latest elements in the sweep-out process.

	X_2	X_3	X_4	X_1
X_2	1(0.50)	2.00	3.85	-31.12
X_3	0(1.00)	1(0.18)	3.00	-5.54
X_4	0(1.93)	0(0.53)	1(0.045)	10.80
X_1	0(-15.56)	0(-0.97)	0(0.49)	1(3.38)

contd..... 3/-

The sample variance of X_1 was computed to be 498.23.

(a) Write the values of the following directly (from the figures given above):

S_{22}, b_{32}, $b_{43.2}$, $b_{14.23}$

(b) Compute the values of the following :

$R^2_{1.234}$, $b_{13.24}$, $r_{14.23}$

(2 x 4+3+3+6) = [20]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR: 1993-94
SEMESTRAL-I EXAMINATION
ECONOMIC STATISTICS

Date: 26.11.93 Maximum Marks: 100 Time: 3 Hours

Note: Attempt all questions.

1.(i) Show that Laspeyres' and Paasche's Price index numbers can be obtained by the method of averaging price relatives as well as by the aggregative method. State with reasons one advantage of the Laspeyres' index over the Paasche's index in case revisions of an index number are to be made from year to year. [10]

(ii) Show that Laspeyres' formula has an upward bias. [10]

(iii) Using Fisher's 'ideal' formula, calculate the quantity index number from the following data

Commodity	Base year Price (Rs.)	Base year Quantity (Kg.)	Current year Price (Rs.)	Current year Quantity (Kg.)
A	5	50	10	56
B	3	100	4	120
C	4	60	6	60
D	11	30	14	24
E	7	40	10	36

[10]

2.(i) Discuss how you will determine trend in a time series by fitting a Polynomial of a suitable degree. [10]

(ii) Fit a straight line trend to the following data and show how you would obtain the monthly trend values from the trend line fitted to the yearly values, and obtain the trend values for December 1950 and August 1949.

Year	1946	1947	1948	1949	1950	1951	1952	1953	1954
Average monthly profits (million Rs.)	6.3	7.4	9.3	7.4	8.3	10.6	9.0	8.7	7.9

[15]

3.(i) Describe the ratio to moving average method of determining seasonal indices from a series of quarterly figures assuming the multiplicative model for time series. [12]

p.t.o.

3.(ii) Compute the average seasonal movements by the method of quarterly average for the following series of observations.

Total Production of Paper (thousand tons)

Year	Quarters I	II	III	IV
1951	37	38	37	40
1952	41	34	25	31
1953	35	37	35	41

[13]

4. Write short notes on:

i) Chain base method of construction of index numbers.

ii) Agricultural statistics in India. [20]

INDIAN STATISTICAL INSTITUTE
M.STAT.(M-STREAM) I YEAR:1993-94
BACKPAPER SEMESTRAL-I EXAMINATION
COMPUTATIONAL TECHNIQUES AND PROGRAMMING

Date:3.1.94 Maximum Marks:100 Time: 3 Hours

Note: The questions carry 110 marks. Maximum you can score is 100. Answer as much as you can.

1. Write short notes on the following:
   (a) File organisation
   (b) Computer memory
   (c) Stack
   (d) Queue [4x5=20]

2. Define absolute and relative errors in computation. Let $N=f(u_1,u_2,\ldots,u_n)$ denote any function of several independent quantities $u_1,u_2,\ldots u_n$ which are subject to errors $\Delta u_1,\Delta u_2,\ldots,\Delta u_n$ respectively. Derive the general form of absolute and relative errors in N. What will be the general form of errors if N is of the form $N=\frac{K.a^m.b^n.c^p}{d^r e^r}$, K= constant. Apply this result to the fundamental arithmatic operation like addition, subtraction, multiplication and division. [10]

3.(a) Derive Lagrange's interpolation formula.

(b) The following table gives certain values of x and corresponding values of $\log_{10}^{x}$. Compute the value of log323.5.

x	321.0	322.8	324.2	325.0
$\log_{10}^{x}$	2.50651	2.50893	2.51081	2.51188

[10+10=20]

4.(a) Describe Newton-Raphson method for finding real root of an equation f(x)=0. Give the geometric interpretation of it.

(b) Using above method find the real root of

$$x^2+4 \sin x = 0.$$ [10+5=15]

5. Write down the procedures for insertions and deletions in a doubley linked linear list. Write clearly all your symbols, notations and steps. [10+10=20]

p.t.o.

6. A manufacturing firm produces two machine parts using lathes, milling machines and grinding machines. The different machining times required for each part, the machining times available on different machines and the profit on each machine part are given in the following table.

Type of machine	Machining time required for the machine part(min.)		Maximum time available per week(min)
	I	II	
Lathes	10	5	2500
Milling machines	4	10	2000
Grinding machines	1	1.5	450
Profit per unit	Rs.50	Rs.100	

We have to determine the number of parts I and II to be manufactured per week to maximize the profit.

Formulate the above problem as a linear programming problem and solve it graphically. [10

7.(a) Draw a flow-chart for solving a quadratic equation.

(b) Write a complete FORTRAN program for the problem in (a) above. [5+10=15

INDIAN STATISTICAL INSTITUTE
M. STAT. I (M Stream) : 1993 94
SEMESTRAL II EXAMINATION
PROBABILITY II

Date : 25.4.94 **Maximum Score** : 120 **Time** : 4 Hours

Note : This paper carries questions worth a total of **140 points**. Marks for each question are given on the right margin. Answer as much as you can. The maximum you can score is 120 points.

1. (a) State Monotone Class Theorem.

   (b) Let $\Omega = (0,1]$, $\mathfrak{A}$ = Borel σ field on Ω and P = Lebesgue measure on $\mathfrak{A}$. Consider the function f on Ω defined as $f(x) = -\log x$ and denote by μ the measure induced by f.

   (i) Find $\mu(I)$ for intervals $I = (a,b] \subset \mathbb{R}$.

   (ii) Show that $\exists$ a non-negative measurable function ϕ on $\mathbf{R}$ such that for every Borel set B, $\mu(B) = \int_B \phi(u)\, du$.

   (iii) Deduce that for any bounded measurable function h on $\mathbb{R}$,
   $\int_{[0,\infty)} h(u)e^{-u}\, du = \int_{(0,1]} h(-\log x)\, dx$.

   (2+(4+7+7))=[20]

2. (a) Suppose X is a real random variable such that $E(\cos 3X) = 1$. Show that, with probability 1, X must take values in the countable set $\{0, \pm\frac{2\pi}{3}, \pm\frac{6\pi}{3}, \cdots\cdots\cdots\}$.

   (b) State Monotone Convergence Theorem.
   Let X be a real random variable and, for each $n \geq 1$, let $X_n = \min\{X, n\}$. Show that, if EX exists and is $> -\infty$, then X_n is integrable for each n and $EX_n \uparrow EX$ as $n \longrightarrow \infty$.

   (c) Let X be a non-negative random variable with finite expectation. For each $a > 0$, let $\phi(a) = E(e^{-aX})$. Show that $\lim_{n\to\infty} n\,[\,1 - \phi(\frac{1}{n})\,]$ exists and equals $E(X)$. (You may use the inequality : $1 - e^{-y} \leq y$ for $y > 0$.)

   (6+(2+5)+7)=[20]

3. (a) (i) Briefly explain how the product of two probability spaces $(\Omega_1, \mathfrak{A}_1, P_1)$ and $(\Omega_2, \mathfrak{A}_2, P_2)$ is defined. (No Proofs Required!)

(ii) State Fubini's Theorem.

(b) Let X and Y be independent random variables, and $B \subset \mathbb{R}^2$ a Borel set. For each $x \in \mathbb{R}$, let $h(x) = P((x, Y) \in B)$.

(i) Show that $P((X, Y) \in B) = E(h(X))$.

(ii) Deduce that if Y has a non-atomic distribution, then so also does $X + Y$.

((6+3)+(5+6))=[20]

4. (a) For a probability space $(\Omega, \mathfrak{A}, P)$, define the classes L_p, $p \geq 1$.

(b) State Holder's inequality and show that $L_s \subset L_r$ if $r \leq s$.

(c) Prove that if $X \in L_1$, then $\int_A X \, dP \longrightarrow 0$ as $P(A) \longrightarrow 0$.

(3+(3+6)+8)=[20]

5. (a) Define convergence in probability.

(b) Show that if $X_n \xrightarrow{P} X$, and, if $|X_n| \leq 1$ for all n, then $|X| \leq 1$, and, $X_n \xrightarrow{L_1} X$.

(c) Show that if $X_n \xrightarrow{P} X$ and if $f: \mathbb{R} \longrightarrow \mathbb{R}$ is a continuous function, then $f(X_n) \xrightarrow{P} f(X)$.

(3+10+12)=[25]

6. (a) State Kolmogorov's Zero-One Law.

(b) Let $\{X_n, n \geq 1\}$ be a sequence of independent random variables. For each of the following events, say whether it is a tail event or not.

(i) $\{\omega : X_n(\omega) \geq -5$ for infinitely many n's$\}$.

(ii) $\{\omega : X_n(\omega)$ is non-decreasing$\}$.

(iii) $\{\omega : \sum_{i=1}^{n} X_i(\omega)$ converges to $0\}$.

(iv) $\{\omega : \frac{1}{n} \sum_{i=1}^{n} X_i(\omega)$ converges to $0\}$.

(c) (i) State the two Borel-Cantelli Lemmas.

(ii) Let $\{X_n\}$ be a sequence of i. i. d. random variables. Show that if $P(X_1 > 0) > 0$, then $P(\sum X_n$ converges$) = 0$.

(5+(4×4)+6+8)=[35]

INDIAN STATISTICAL INSTITUTE
M.STAT. (M-STREAM) I YEAR
THEORY AND METHODS OF STATISTICS III
SEMESTRAL-II EXAMINATION

Date: 27.4.94 Maximum Marks: 100 Time: $3\frac{1}{2}$ Hours

Note: Answer any four questions. All questions carry equal marks.

1.(a) Let $p_\theta(x)$ represent the p.d.f. or frequency function of a r.v. X. What do you mean by a UMP test of $H_0: \theta \geq \theta_0$ against $H_1: \theta < \theta_0$.

(b) Suppose $p_\theta(x)$ in (a) has montone likelihood ratio in $T(x)$. Show that the size α MP test of $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$, $\theta_1 < \theta_0$ is UMP for the testing problem in (a).

(c) What do you mean by unbiasedness of a test? Show that the UMP test in (b) is unbiased.

(d) What is a UMPU test? Suppose a random sample of size n is drawn from an exponential distribution with mean θ. Find UMPU test of $H_0: \theta = \theta_0$ against the alternative $H_1: \theta \neq \theta_0$. (2+9+4+10)=(25)

2.(a) Define a 'one-parameter exponential family of distributions'.

If $X_1, X_2, \ldots, X_k$ are i.i.d. r.v.'s belonging to a one - parameter ex-ponential family, comment on the distribution of the random variable $X = (X_1, X_2, \ldots X_k)$.

Show that Cauchy distribution with location parameter θ, $-\infty < \theta < \infty$ does not belong to the one-parameter exponential family.

(b) Show that if $p_\theta(x)$ (representing the p.d.f. or frequency function of a r.v. X) belongs to a one-parameter exponential family, there exists a parametric function, say $g(\theta)$ for which an unbiased estimator has the variance given by Cramer-Rao lower bound.

(c) Let Y follow negative binomial distribution given by the frequency function

$$P_p(Y = y) = \binom{m+y-1}{m-1} q^y p^m, \quad y = 0, 1, 2, \ldots\ldots$$
$$= 0, \text{ otherwise}, \quad m \text{ given}.$$

Does the distribution of Y belong to one-parameter exponential family?

Find the uniformly minimum variance unbiased estimator (UMVUE) of $1/p$? Does its variance satisfy Cramer-Rao lower bound? Justify your answer. (8+7+10)=25

3.(a) Let X follow the truncated exponential distribution given by

$$P(X=x) = e^{-\lambda} \frac{\lambda^x}{x!} / (1-e^{-\lambda}), \quad x=1,2,3,\ldots\ldots$$

$$= 0, \text{ otherwise.}$$

Suppose a random sample of size n is drawn from this distribution. Find the uniformly minimum variance unbiased extimator (umvue) of λ.

(b) Consider a random sample of size n from the two parameter exponential distribution given by the p.d.f.

$$f(x) = \beta e^{-\beta(x-\alpha)}, \quad x \geq \alpha$$

$$\beta > 0, \; -\infty < \alpha < \infty$$

$$= 0, \text{ otherwise.}$$

Find a minimal sufficient statistic.

Assuming completeness, find the umvue of (i) α and (ii) $e^{-\beta}$.

(c) Find the UMP test for H_0 $(\alpha = \alpha_0)$ against H_1 $(\alpha \neq \alpha_0)$ assuming β to be known, for the problem in (b). (6+11+8)=(25)

4.(a) Let $p_\theta(x)$ represent the p.d.f. or frequency function of a random variable X, $\theta \in \Omega$. Define a statistic T(X) as follows:

$$T(x) = T(y) \text{ if and only if } \frac{p_\theta(x)}{p_\theta(y)} \text{ is free from } \theta \text{ for all } \theta \in \Omega.$$

Show that T(X) is minimal sufficient for the family of distributions.

(b) A random sample of size n is drawn from a population characterized by the p.d.f. $f_\theta(x) = \frac{1}{2} e^{-|x-\theta|}, \; -\infty < x < \infty, \; -\infty < \theta < \infty$

Find a minimal sufficient statistic for θ.

Hence or otherwise, find a minimal sufficient statistic for the location family of all continuous distributions with location parameter θ. Justify your answer.

(c) For the linear model $Y_{n\times 1} = X_{n\times p}\,\beta_{p\times 1} + \varepsilon_{n\times 1}$, $\varepsilon \sim N_p(0, \sigma^2 I_n)$,

where X is a given non-stochastic matrix and β is a vector of parameters, $\beta \in R^p$, show that

$$y'(I-X(X'X)^- X')y/(n-r) \text{ is the umvue of } \sigma^2, \text{ where } r = \text{Rank}(X)$$

(8+6+11)=(25)

contd. 3/

5.(a) (i) What do you mean by a complete family of distributions?

Suppose P_0 and P_1 are two families of distributions with $P_0 \subset P_1$. Give an example to show that P_0 is complete does not necessarily imply that P_1 is complete..

(ii) Find a minimal sufficient statistic for a random sample of size n from $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, $-\infty < \theta < \infty$

Show that the minimal sufficient statistic is not complete.

(b) (i) What do you mean by consistency of a sequence of estimators.

Show that, if T_n is a sequence of estimators of $g(\theta)$ with

$E_\theta(T_n) - g(\theta) \to 0$ and $Var_\theta(T_n) \to 0$ as $n \to \infty$ for all θ,

then T_n is consistent.

(ii) Find an unbiased estimator of θ based on the minimal sufficient statistic for the problem in a (ii). Prove consistency of the estimator proposed by you. (5+7+6+7) = (25)

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1993-94
SEMESTRAL II EXAMINATION

Applied Stochastic Processes

Date: 29.4.1994 Maximum Marks: 100 Time: 3 hours

1. Let $\{S_n^x\}$ be a simple random walk starting at x.

   (a) Compute $P\{S_n^x = y\}$. [6]

   (b) Let $\eta_x = \inf\{n \geq 1 : S_n^x = x\}$
   $T_y^x = \inf\{n \geq 0 : S_n^x = y\}$.

   Show that $P(\eta_x = k \mid S_1^x = x+1) = P(T_x^{x+1} = k-1)$, $k \geq 2$. [6]

   (c) Show that $P\{\eta_x < \infty\} = 2 \min\{p, q\}$. [7]

2. Let $(X_n)_{n \geq 0}$ be a Markov Chain with transition probabilities $((p_{ij}))$.

   (a) State and prove the Markov property for (X_n). [5]

   (b) Describe p_{ij} in the case of a simple random walk on $S = \{0, 1, 2, \ldots n\}$ with reflecting barriers at 0 and n. Is this chain irreducible ? aperiodic ? Does it have a unique invariant measure ? Justify your answer for each one. (3+4+4) = [11]

3. Let $S = \{0, 1, \ldots n\}$ and let

$$p_{ij} = \binom{n}{j}\left(\frac{i}{n}\right)^j \left(1 - \frac{i}{n}\right)^{n-j} \qquad 0 \leq i, j \leq n.$$

   (a) Show that $((p_{ij}))$ is a transition matrix on S for which 0 and n are absorbing states. [4]

   (b) Show that every state $0 < i < n$ is transient. [5]

   (c) For $0 < i < n$, compute the limits

$$\underset{m \to \infty}{\text{Lt}} \; p_{i0}^{(m)}, \; p_{in}^{(m)}.$$ [10]

4. (a) Define a pure birth process and derive the forward equations for its transition probabilities. [8]

contd..... 2/-

(b) Suppose (X_t^m) is the size of a population at time t, which evolves as a Yule process with parameters $\lambda_n = n\beta$, $\beta > 0$, and $X_0^m \equiv m$. Suppose the evolutions of the offspring of the m individuals at time 0 are independent and without interaction. Compute explicitly $p_{mj}(t)$, $j \geq m$.
(You may use the fact that

$$P(X_t^1 = n) = e^{-\beta t}(1 - e^{-\beta t})^{n-1}, \quad n \geq 1).$$

[10]

5. (a) Compute the stationary distribution of an ∞ - server queue with parameters $\lambda_n = \lambda \ \forall n$ and $\mu_n = n\mu$, $\mu > 0$. [5]

(b) Let (X_t) be a birth and death chain on $S = \{0, 1, 2, \ldots\}$ with parameters λ_m and μ_m. Suppose that 0 is an absorbing state. Let $\tau = \inf\{s > 0 : X_s = 0\}$ and $u_m = P\{X_0 = m, \tau < \infty\}$.

Show that $u_m = \dfrac{\lambda_m}{\lambda_m + \mu_m} u_{m+1} + \dfrac{\mu_m}{\lambda_m + \mu_m} u_{m-1}$. [10]

6. (a) Let $(S_n)_{n\geq 0}$ be a simple symmetric random walk, $S_0 \equiv 0$. Let $X_t^n = \frac{1}{\sqrt{n}} S_{[nt]}$, $t \geq 0$. For fixed $0 \leq t_1 < t_2$ determine the limiting distribution of $(X_{t_1}^n, X_{t_2}^n)$. [10]

(b) State the reflection principle for Brownian motion. If (X_t) is a standard Brownian motion, $X_0 = 0$, $M_t = \max_{0\leq s\leq t} X_s$ and $Y_t = M_t - X_t$, find the joint distribution of (M_t, Y_t)

[10]

:bcc:

1993-94 548

INDIAN STATISTICAL INSTITUTE

M.Stat. I Year (M-stream) : 1993 - 94

SEMESTRAL II EXAMINATION

Linear Statistical Models and Large Sample Statistical Methods

Date: 2.5.1994 Maximum Marks: 100 Time: 3 hours

Note: Answer any FIVE questions choosing AT LEAST TWO questions from each group.

GROUP - A

1. Consider a linear model $(\underset{\sim}{Y}, X\underset{\sim}{\beta}, \sigma^2 I)$ where $\underset{\sim}{Y}$ is an $(n \times 1)$ random vector, X is an $(n \times k)$ matrix of known constants of rank $r(< \min(n, k))$, $\underset{\sim}{\beta}$ is a $k \times 1$ vector of unknown parameters and σ^2 is unknown common variance of the components of $\underset{\sim}{Y}$.

(a) Show that the normal equations for estimating $\underset{\sim}{\beta}$ by the method of least squares admit a solution.

(b) Obtain a necessary and sufficient condition for a linear function $\underset{\sim}{\lambda}'\underset{\sim}{\beta}$ to be estimable and obtain its BLUE.

(8+12) = [20]

2.(a) Let $\underset{\sim}{Y}$ be a $N_n(\underset{\sim}{0}, I)$ random vector and let

$$\underset{\sim}{Y}'\underset{\sim}{Y} = \sum_1^k \underset{\sim}{Y}' A_i \underset{\sim}{Y} = \sum_1^k Q_i \text{ where } R(A_i) = n_i,\ 1 \le i \le k.$$

Show that a necessary and sufficient condition that $Q_1, \ldots, Q_k$ are independent and Q_i has the χ^2 distribution on n_i degrees of freedom for $1 \le i \le k$ is $\sum_1^k n_i = n$.

(b) Hence or otherwise show that a quadratic form $\underset{\sim}{Y}' A \underset{\sim}{Y}$ has the χ^2 distribution with degrees of freedom equal to $R(A)$ if A is idempotent.

(15+5) = [20]

3. In the set up of question (1) suppose $\underset{\sim}{Y}$ is $N_n(X\underset{\sim}{\beta}, \sigma^2 I)$. Let $\Lambda\underset{\sim}{\beta}$, where Λ is an $(m \times k)$ matrix of rank m be a set of estimable parametric functions. Derive a test for the hypothesis $\Lambda\underset{\sim}{\beta} = 0$ against the alternative $\Lambda\underset{\sim}{\beta} \neq 0$. (State precisely the results you use).

[20]

p.t.o.

4.(a) Show that in a usual two way classification layout the presence of interaction can not be tested unless the number of observations is more than 1 in at least one cell.

(b) Derive a test for the hypothesis that there is no interaction between the row effects and column effects in a two way classification with n(> 1) observations in each cell.

(10+10) = [20]

GROUP - B

5.(a) Let $\{X_n\}$ be a sequence of random variables converging in quadratic mean to a random variable X and let $\{Y_n\}$ be a sequence of random variables such that $\{X_n - Y_n\}$ converges in probability to 0. Show that the sequence $\{Y_n\}$ converges in distribution to X.

(b) Show that if a sequence $\{X_n\}$ of random variables converges in probability to X, the sequence would converge in distribution to X. Show that if $\{X_n\}$ converges in distribution to a degenerate random variable $X \equiv C$, then $\{X_n\}$ would converge in probability to C.

(10+10) = [20]

6.(a) Let $\{A_n\}$ be a sequence of independent events and $A = \limsup A_n$ Under suitable non-trivial conditions to be stated by you, show that P(A) = 1.

(b) Let $\{X_n\}$ be a sequence of independent identically distributed random variables with finite mean μ and finite variance σ^2.

Show that $\bar{X}_n = \frac{1}{n} \sum_{1}^{n} X_i$ converges almost surely to μ as $n \longrightarrow \infty$.

(10+10) = [20]

7. A random sample $X_1, \ldots, X_n$ is drawn from a population with density function $f(x, \theta)$, where $\theta \in \Theta \subset \mathbb{R}$ is unknown. Show that, under suitable conditions to be stated by you, that the likelihood equation has a unique consistent solution with probability going to 1 as the sample size tends to ∞.

[20]

:bcc:

INDIAN STATISTICAL INSTITUTE

M.Stat.(M-stream) I Year: 1993-94

Sample Surveys and Design of Experiments

Semestral- II Examination

Date : 4.5.1994 Maximum Marks : 100 Time : 3 Hours

Answer Group A and Group B in separate answer books.

GROUP A : Sample Surveys Max.Marks: 50

Attempt all questions.

1.(a) Distinguish between 'sampling errors' and 'non sampling errors'. Mention briefly the sources of non-sampling errors in a survey of production of mangoes in a district and the steps you would take to asses them.

(4+5+3=12)

(b) A simple random without replacement sample of 34 students was selected from a hostel consisting of 321 students. It was found that 8 of them spend more than 20 hours per week in extra-curricular activities. Estimate the proportion of students in the hostel who belong to this category. Also obtain an estimate of the sampling error.

(3+5 = 8)

2.(a) In stratified simple random sampling with replacement, let C_0 be the overhead cost and C_i be the cost per unit in the ith stratum, i = 1,2,..,k. Derive an optimum allocation of a total sample size to strata such that with a linear cost function of the form $C = C_0 + \sum_1^k c_i n_i$, the variance of the estimated mean $\hat{Y}_{st}$ is a minimum for a specified cost C.
What does this allocation reduce to when C_i are all equal ? How do you actually obtain this allocation in practice ?

(10+2+3 = 15)

3. An experienced tutor makes an estimate of scores in a test of each student in a college of N = 200 students. He finds a total score of X = 11,600. After the test, a simple random (without replacement) of 10 students was selected and the following results were obtained :

p.t.o.

	Student Number									
	1	2	3	4	5	6	7	8	9	10
Actual Score y_i	62	42	50	58	67	45	39	57	71	53
Estimated Score x_i	59	47	52	60	67	48	44	58	76	58

(a) Find the regression estimate $\hat{\bar{Y}}$ of $\bar{Y}$, the average score for the population and estimate its sampling error.

(b) Supposing that the sample has been selected with probabilities of selection of students proportional to their estimated scores x_i and with replacement, obtain an estimate of $\bar{Y}$.

(8+4+3=15)

GROUP B Max.Marks : 50

Design of Experiments

Answer all questions.

1.(a) What are the advantages of partial confounding over total confounding in a factorial experiment ? Explain briefly. An experiment is to be performed to study the effect of acid concentration (A), catalyst concentration (B) and temperature (C) on the rate of a chemical process. A,B,C have 3 levels each and the experiment is to be run in blocks of size 3.

(b) Suggest a suitable confounding scheme for the experiment. For r replications of your design : (c) give the partition of the degrees of freedom in the analysis of variance table. and (d) give the relative efficiencies (w.r.t. an unconfounded design) for each effect.

(3+3+4+4=14)

2. The effect of 5 different mixing methods on the strength (Y) of an alloy is to be studied. There are 5 batches of raw material and 5 chemists to carry out the experiment. It is decided to run the experiment as a Latin Square.

(a) Justify this decision .

(b) Suppose that the first chemist forgot to record the observation from batch 1 when he was working with mixing method 2. Derive a formula for estimating this missing value x so that x will have a minimum contribution to the error sum of squares.

2.(c) Suppose in (a), it is decided to carry out an ANCOVA with the Latin Square design and 2 concomitant variables X and Z. Give an expression (in terms of the observed values of Y,X and Z) for estimating the difference between the effects of method 4 and 5. (no proof required).

(3+9+4 = 16)

3. The following is the layout (in usual notation) of a design when a 2^3 experiment is performed in 2 replicates, each replicate having 2 blocks. The data is given with each treatment combination. Analyse the data.

(20)

Rep. 1.

(1)	-3	a	0
ac	2	b	-1
bc	1	c	-1
ab	2	abc	6

Rep. 2

(1)	-1	a	1
b	0	c	0
ac	3	ab	1
abc	5	bc	1

Given : Total (corrected) SS = 78.00,
Total of 8 observations from Rep 1 = 6.00 and from Rep 2 is 10.00.

$F_{1,5}(0.1) = 16.26$

[1993-94 | 568]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
DEMOGRAPHY
SEMESTRAL-II EXAMINATION

Date: 6.5.94 Maximum Marks: 100 Time: 3 hours

Note: Attempt any five questions.

1.(a) Define a population census.

(b) Discuss the relative advantages and disadvantages of the various methods of enumeration in census operations.

(c) What are special features of the 1991 census of India as compared to the immediately preceding census? [2+10+8=20]

2.(a) Describe various errors and biases in the age returns from a census.

(b) Discuss Chandrasekar's method for evaluating age data given by single years of age.

(c) Define joint scores for evaluating age data given by quinquennial ages and also indicate their possible ranges to measure the degree of accuracy of the data. [5+10+2+3=20]

3.(a) Define and derive the logistic law of population growth.

(b) Interpret the parameter in a logistic law.

(c) Why do we prefer the logistic curve to other curves for describing population growth? [8+8+4]=20

4. Following demographic information is available in a year:

Item	Country A	Country B	Country C
Population by age	P_x^A	P_x^B	P_x^C
Deaths by age	D_x^A	D_x^B	-
Total population	$P^A=\Sigma P_x^A$	$P^B=\Sigma P_x^B$	$P^C=\Sigma P_x^C$
Total deaths	$D^A=\Sigma D_x^A$	$D^B=\Sigma D_x^B$	D^C

4.(a) Estimate crude death rates for these three countries, what can you say about the relative mortalities for these three countries through crude death rates?

(b) Suggest and derive methods for comparison of mortality between these three countries. [3+5+12=20]

p.t.o.

5.(a) Define crude birth rate and total fertility rate.

(b) Find the relationship between these two rates.

(c) Interpret the situations:

(i) NRR=1, (ii) NRR > 1 and (iii) NRR < 1. [4+10+6=20]

6.(a) Show that for five consecutive values of l_x, estimate μ_x by

$$\frac{8(l_{x-1}-l_{x+1})-(l_{x-2}-l_{x+2})}{12l_x}$$

where l_x and μ_x have their usual meanings.

(b) For a certain population

$$l_x = 10.000(121-x)^{1/2}.$$

Find (i) μ_x, (ii) q_x and

(ii) the probability that a person aged 0 will die between ages 21 and 40.

(c) Prove that

$$m_x = -\frac{1}{L_x}\frac{d}{d_x}L_x.$$

using this result, show that

$$m_x \doteq \mu_{x+\frac{1}{2}}.$$

Note: all symbols have their usual meanings. [5+9+6=20]

7.(a) The Force of mortality in a particular population follows the Makeham's Law and $\mu_x = A+BC^x$.

Prove that : $l_x = k\, s^x g^{c^x}$.

where k, s and g are constants.

(b) Describe a procedure to fit the curve:

$$l_x = ks^x g^{c^x}.$$

to observeddata. [10+10=20]

8.(a) Define nuptiality. Describe and derive various columns of a net nuptiality table.

(b) Describe a procedure for projecting regional population of a country. [10+10=20]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year: 1993-94
Applied Stochastic Processes
Semestral-II Backpaper Examination

Date : 20.6.1994 Maximum Marks : 100 Time : 3 Hours.

1. Let $\{S_n^x\}$ be a simple random walk starting at x.
   Let $T_y^x = \inf\{n \geq 0 : S_n^x = y\}$.
   (a) Compute the probabilities $P\{T_d^x < T_c^x\}$ and $P\{T_c^x < T_d^x\}$ for $c \leq x \leq d$ and $p \neq q$. [10]

   (b) If $p = q = 1/2$, then prove that every point $y \in Z$ is recurrent for $\{S_n^x\}$. [10]

2. (a) Define a Markov Chain. [4]
   (b) Construct a Markov Chain which describes the size of a population at times 0,1,2,....., stating your assumptions on the reproductive mechanism. [6]

   (c) Write down the transition probabilities for a simple random walk with absorbing boundaries on S = 0,1,.... n . What are the distinct equivalence classes of essential states ? [6]

3. (a) Let $(X_n)_{n\geq 0}$ be a Markov Chain with transition probabilities $((p_{ij}))$. Let $N_j = \# \{k : X_k = j\}$. Let $G_{ij} = E_i(N_j)$. Show that $G_{ij} = \sum_{n=0}^{\infty} p_{ij}^{(n)}$. [5]

   (b) Show that a state j is recurrent iff $G(j,j) = \infty$ and transient iff $G(j,j) < \infty$. [12]

4. (a) Let $(X_t)_{t\geq 0}$ be a continuous time homogeneous Markov Chain on the countable set S. Let $p_{ij}(t) = P[X_t=j \mid X_0=i]$ and $P(t) = ((p_{ij}(t)))$. Show that $P(t+s) = P(s)P(t)$ and derive the forward equations for P(t). [8]

p.t.o.

4.(b) Let $(X_t)_{t\geq 0}$ be a Poisson process. Let $u < t$ and $k < n$.

Given that n events have occurred during $[0,t]$, find the probability that exactly k of them have occurred during $[0,u]$.

[5]

(c) Let $(X_t)_{t\geq 0}$ be a Markov Chain with $X_0 \equiv i$.

Let $T_0 = \inf\{t > 0 ; X_t \neq i\}$. Find the distribution of T_0.

[8]

5. Let $(X_t)_{t\geq 0}$ be a birth and death chain with parameters λ_n, μ_n.

(a) Derive the forward equations for (X_t).

[10]

(b) Derive an expression for tho stationary distribution of (X_t) in terms of λ_n and μ_n, whenever it exists.

[10]

6. Let $(X_t)_{t\geq 0}$ be a standard Brownian motion, $X_0 \equiv 0$

(a) For $a \neq 0$, let $\tau_\alpha = \inf\{s : X_s = a\}$. Compute the distribution of τ_a.

[8]

(b) Show that $Y_t = tX(1/t)$, $t > 0$ has the same finite dimensional distribution as $(X_t)_{t>0}$.

[8]

1993-94 | 548(b)

INDIAN STATISTICAL INSTITUTE
M.Stat. I Year (M-stream) : 1993-94
BACKPAPER SEMESTRAL II EXAMINATION
Linear Statistical Models and Large Sample Statistical Methods

Date: 22.6.1994 Maximum Marks: 100 Time: 3 hours

Note: Answer any FIVE questions choosing AT LEAST TWO questions from each group.

GROUP - A

1. Consider a linear model ($\underset{\sim}{Y}$, $X\underset{\sim}{\beta}$, $\sigma^2 I$) where $\underset{\sim}{Y}$ is an $n \times 1$ random vector, X is an $(n \times k)$ matrix of known constants $\underset{\sim}{\beta}$ is a $k \times 1$ vector of unknown parameters and σ^2 is the common unknown variance of the components of $\underset{\sim}{Y}$.

(a) To estimate $\underset{\sim}{\beta}$ by the method of least squares, show that the normal equations always admit a solution. When do the normal equations have a unique solution ?

(b) Whether the normal equations have a unique solution or not, show that a solution of the normal equations minimizes the error sum of squares. Obtain an unbiased estimator of σ^2.

(10+10) = [20]

2.(a) In the linear model described in question 1, suppose $R(x) = p$ ($< \min(n, k)$). Show that there are only p linearly independent linear functions of $\underset{\sim}{\beta}$ which are estimable.

(b) Define the error space of the model and show that the least squares estimator of any estimable linear parametric function is uncorrelated with any linear function of observations in the error space. What is the dimension of the error space ? Justify.

(10+10) = [20]

3.(a) Let $\underset{\sim}{Y}$ be $N_n(0, I)$ random vector and $Q = \underset{\sim}{Y}' A \underset{\sim}{Y}$. Show that a necessary and sufficient condition that Q has the χ^2 distribution is that A is idempotent, in which case the degrees of freedom is R(A).

(b) Let $Q_i = \underset{\sim}{Y}' A_i \underset{\sim}{Y}$, $i = 1, 2$, where Q_1 has the χ^2 distribution with a degrees of freedom and Q_2 has the χ^2 distribution with b degrees of freedom. Show that a necessary and sufficient condition that Q_1 and Q_2 are independent is $A_1 A_2 = A_2 A_1 = 0$.

(10+10) = [20]

p.t.o.

4. In order to test the equality of k treatment effects, n_i observations are made on ith treatment $1 \leq i \leq k$. Derive a test for testing the hypothesis stating clearly the assumptions you make and the main results you use. Write also the ANOVA table in this case. [20]

GROUP - B

5.(a) Let $\{X_n\}$ be a sequence of random variables converging in distribution to a random variable X and $\{Y_n\}$ be a sequence of random variables converging in distribution to a constant $c \neq 0$. Show that $\left\{\frac{X_n}{Y_n}\right\}$ converges in distribution to $\frac{X}{c}$.

(b) Show that a sequence $\{X_n\}$ of random variables which converges to a random variable X in rth mean, $r \geq 2$ converges to X (i) in sth mean for $1 \leq s < r$. and (ii) in probability.

(10+10) = [20]

6.(a) Let $X_1, \ldots, X_n$ be independent random variables with $E(X_i) = 0$, $V(X_i) = \sigma_i^2$ $1 \leq i \leq n$. Show that

$$P\left(\max \left| \sum_1^i X_j \right| \geq \delta \right) \leq \frac{1}{\delta^2} \sum_1^n \sigma_i^2 \quad \text{for any } \delta > 0.$$

(b) Let $\{X_n\}$ be a sequence of independent identically distributed random variables with mean 0 and variance σ^2. Show that $\frac{1}{n} \Sigma X_i$ converges to 0 almost surely as $n \to \infty$.

(10+10) = [20]

7. Let $X_1, \ldots, X_n$ be a random sample from a population with density function $f(x, \theta)$, where

$$E_\theta(X) = \theta \varepsilon (H) \quad \mathbb{R} \text{ is unknown.}$$

(a) Show that $\overline{X}_n = \frac{1}{n} \sum_1^n X_i$ converges in P_θ-probability to θ as $n \to \infty$.

(b) Under suitable additional conditions to be stated by you show that the maximum likelihood estimator of θ obtained as the solution of the likelihood equation is asymptotically normal. (You need not show the existence of a solution of the likelihood equation).

(10+10) = [20]

INDIAN STATISTICAL INSTITUTE

M.Stat. I Year (M-stream) : 1993-94

BACKPAPER SEMESTRAL II EXAMINATION

Theory and Methods of Statistics III

Date: 22.6.1994 Maximum Marks: 100 Time: 3 hours

Note: Answer ALL the questions.

1.(a) X is distributed according to a distribution from the family $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$.
Let T(X) be an estimator with $E_\theta(T^2(X)) < \infty$ for all $\theta \in \Omega$ and $\mathcal{U}$ denote the set of estimators u(X) of zero with $E_\theta(u^2(X)) < \infty$ for all $\theta \in \Omega$. Show that a necessary and sufficient condition for T(X) to be umvu estimator of its expectation is that

$$E_\theta(T(X).u(X)) = 0 \quad \text{for all } u(X) \in \mathcal{U} \text{ and all } \theta \in \Omega .$$

(b) Consider the following linear model

$$E(y_{ij}) = 2\theta + 3\phi , \quad i = 1, j = 1, 2, \ldots, n$$
$$= 3\theta - 2\phi , \quad i = 2, j = 1, 2, \ldots, n$$

Let $y' = (y_{11} \ldots y_{1n} \quad y_{21} \ldots y_{2n})$, then $D(y) = \sigma^2 I_{2n}$ and y is normal.

Find umvue's of θ, ϕ and σ^2

(c) Let X be a r.v. with the following distribution:

$$P(X = K) = p , \quad K = -1$$
$$= p^K(1-p)^2, \quad K = 0, 1, 2, \ldots$$
$$= 0, \quad \text{otherwise}, \quad 0 < p < 1.$$

Show that the only parametric functions having umvu estimators are $a + b(1-p)^2$, where a and b are any real numbers.

(7+12+6) = [25]

2.(a) Consider a random sample of size n from the distribution given by the frequency function

$$f_\theta(x) = \frac{1}{\theta}, \quad x = 1, 2, \ldots \theta , \quad \theta \text{ is a positive integer.}$$
$$= 0, \quad \text{otherwise.}$$

contd..... 2/-

Show that the maximum observation $X_{(n)}$ is sufficient for θ

Show that $X_{(n)}$ is also complete

Find a u m v u e of θ.

(b) What is an ancillary statistic ? For a family of distributions $\mathcal{P} = \{ F_\theta, \theta \in \Omega \}$ of a r.v. X, suppose $T(X)$ is a complete sufficient statistic and $U(X)$, an ancillary statistic. Show that $T(X)$ and $U(X)$ are independently distributed.

(c) For a random sample of size n from $U(0, \theta)$, suppose $X_{(1)}$ represent the minimum and $X_{(n)}$ the maximum observation.

Show that $X_{(1)} / X_{(n)}$ is an ancillary statistic.

Show that $X_{(n)}$ and $X_{(1)}/X_{(n)}$ are independently distributed.

(9+7+9) = [25]

3.(a) Let $p_\theta(x)$ represent the p.d.f. or frequency function of a r.v. X. What do you mean by a UMP test of $H_0 : \theta \geq \theta_0$ against $H_1 : \theta < \theta_0$.

(b) Suppose $p_\theta(x)$ in (a) has montone likelihood ratio in $T(x)$. Show that the size α MP test of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$, $\theta_1 < \theta_0$ is UMP for the testing problem in (a).

(c) From a random sample of size n from the distribution given by the p.d.f.

$$f_{(\alpha,\beta)}(x) = \beta e^{-\beta(x-\alpha)} \quad , \begin{cases} x \geq \alpha, & \beta > 0, \\ & -\infty < \alpha < \infty \end{cases}$$
$$= 0, \text{ otherwise}$$

Construct a UMP test for $H_0 \; (\alpha = \alpha_0)$ against $H_1 \; (\alpha \neq \alpha_0)$, assuming β to be known. Derive an expression for the power function of the test.

(3+10+12) = [25]

4.(a) Consider a random sample of size n from $N(\theta, \theta^2)$, $-\infty < \theta < \infty$. Find a minimal sufficient statistic and prove that it is not complete.

(b) From a random sample of size n from $N(\theta, \sigma^2)$ with σ^2 known, $-\infty < \theta < \infty$, find the u m v u e of $\Phi((u-\theta)/\sigma)$, where u is given.

Prove consistency of the u m v u e.

(c) Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. r.v. s with variance $\sigma^2 < \infty$. Show that $s_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is consistent for σ^2.

[State clearly the convergence theorems in probability you may be using in your proof]

(9+9+7) = [25]

:bcc:

1993-94 558(b)

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
SAMPLE SURVEYS AND DESIGN OF EXPERIMENTS
SEMESTRAL-II BACKPAPER EXAMINATION

Date: 24.6.94 Maximum Marks: 100 Time: 3 Hours

Note: Answer Group A and Group B in separate answer scripts.

GROUP-A

Sample Surveys Maximum Marks: 50

Note: Attempt all questions.

1.(a) What are 'non-sampling errors'. Discuss briefly the method of assessment and control of non-sampling errors in a survey relating to the expenditure of students in an educational institution. (4+8)=(12)

(b) From a population of 492 Hatkar males with C.V. of 0.2771 for a characteristic y, how many members should be chosen in order to estimate the population mean $\bar{Y}$ by simple random sampling without replacement design such that the C.V. of $\hat{\bar{Y}}$ is 5%? (8)

2.(a) Describe 'linear systematic sampling' (LSS) and 'Circular systematic sampling' (CSS) procedures. What are the advantages of CSS over LSS ?

(b) Explain why on the basis of a single systematic sample it is not possible to estimate unbiasedly the variance of the estimator of the population mean. Suggest any two procedures for overcoming this difficulty.

(c) Yields based on two independent circular systematic samples of size 4 each were found to be

sample 1 : 247, 238, 359, 125
sample 2 : 256, 214, 368, 141

Estimate the population mean and its sampling error.

3+(5+2)+(2+5) = (15)

p.t.o.

3. An experienced farmer makes an eye estimate of the weight of peaches X_i on each tree in each of the 2 strata of an orchard with 100 trees each. He finds a total weight of X_1=500 and X_2=660 kgs. respectively in each stratum.
The peaches are then picked and weighed on simple random samples (without replacement) from each stratum of 5 trees each and the following results are obtained:

	STRATUM 1					STRATUM 2				
Tree	1	2	3	4	5	1	2	3	4	5
Actual weight y_i	61	42	50	58	67	45	39	57	71	53
Estimated weight x_i	59	47	52	60	67	48	44	58	76	58

Find the regression estimate Y of Y, the total weight of peaches in the population and estimate its sampling error. (8+7)=(15)

GROUP-B : Design of Expt. Maximum Marks: 50

Note: Answer all questions.

1.(a) For a randomised block design with b blocks and t treatments, suppose two observations are missing from the same block. Derive a general formula for estimating this missing observation.

(b) How is the latin square used to eliminate two extraneous sources of variability?

(c) How can you modify the Latin square design if there is one more source of variability? (10+3+2=15)

2.(a) Explain briefly the principle of confounding in a factorial experiment.

(b) Construct a factorial experiment for 2 factors each at 3 levels in blocks of 4 using a suitable confounding scheme.

(c) For r replications of the design constructed in (b), calculate the relative efficiencies of the different effects compared to an unconfounded design. (3+8+4=15)

3. An experiment was carried out as a RBD in 9 blocks with 5 treatments and observations were taken on the variable of interest Y and one concomitant variable X.

(a) Write the appropriate model for analyzing the data obtained from this experiment.

contd. ...3/-

3.(b) The following data were obtained. Complete the analysis of covariance and draw conclusions.

Source of Variation	Sum of Squares and Products		
	Y	X	XY
Blocks	12	2	6
Treatments	8	1	3
Error	14	6	7

(5+15=20)

[1993-94] [518(b)]

INDIAN STATISTICAL INSTITUTE
M.STAT. (M-STREAM) I YEAR
PROBABILITY THEORY II
SEMESTRAL-II BACK-PAPER EXAMINATION

Date: 27.6.94 Maximum Marks: 100 Time: 3 Hours

1. (a) Explain clearly what is meant by a real random variable and the probability distribution of a real random variable.

   (b) Prove that two real random variables having the same distribution function must have the same probability distribution. [8+12]=20

2. (a) When are a finite number of real random variables said to be independent?

   (b) Prove that if X,Y,Z are independent random variables, then X+Y and Z are independent. [4+12]=16

3. Let X be a random variable and F is any probability distribution function.

   (a) Show that, for each real number a, $F(a-X)$ is an integrable random variable.

   (b) Show that $G(a) = E(F(a-X))$ defines a probability distribution function on $\mathbb{R}$. [8+12]=20

4. (a) Define convergence in probability and a.s. convergence.

   (b) Prove that $X_n \xrightarrow{a.s.} X$ if and only if,

   $$\sup_{k \geq n} |X_k - X| \xrightarrow{P} 0$$

   (c) Deduce that a.s. convergence implies convergence in probability. [6+12+6]=24

5. (a) State the two Borel-Cantelli Lemmas

   (b) Show that if $\{X_n\}$ is an i.i.d. sequence of random variables and if $0 < P(X_1 \leq 1) < 1$, then $P(X_n \text{ converges}) = 0$.

   (c) State Kolmogorov's Maximal inequality. [6+10+4]=20