1985-86 | 382(b)

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1985-86

NONPARAMETRIC AND SEQUENTIAL METHODS

SEMESTRAL-II BACKPAPER EXAMINATION

Date :4.7.86. Maximum Marks : 100 Time : 3 Hours.

Note : Show all your work. Attempt all questions.
22 marks have been allotted to each question.

1. Compute ARE (B, W^+) for Pitman alternatives, where B is the sign-test and W^+, the Wilcoxon - signed rank test statistic, in the one-sample location setting with one-sided alternative.

2. (a) Derive the asymptotic null distribution of the test-statistic for the sign-test.

   (b) Write a note on the property of "Robustness" for the t-test.

3. Prove that the SPRT terminates with prob.1 under H_0 and H_1 (State and prove all results you need).

4. Show by using Stein's method, that an interval estimator of a <u>given length</u> can be constructed for the mean of a normal distribution, even when the variance is unknown, by sampling in two stages.

5. Assignments - 12 marks.

1985-86 | 302

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1985-86

ELECTIVE-5 : BIOLOGICAL SCIENCES
SEMESTRAL-II EXAMINATION

Date : 16.5.86. Maximum Marks : 100 Time: 3 Hours.

Note : Answer five questions.

1.(a) Construct a simple deterministic model of DNA, RNA and Protein Synthesis during embryogenesis using the principle of lateral inhibition. Discuss the conditions of stability of the system.

(b) Find the condition of the existence of an oscillatory solution of the above system and deduce the explicit solution (oscillatory) of RNA and Protein. [10+10-[illegible]

2.(a) Construct the Goodwin - type model of [illegible] [illegible]tein and Enzyme synthesis during [illegible]genesis using the principle of feedback [illegible]ion.

(b) Discuss the stability properties of the above model.

(c) Establish the Griffith condition of oscillatory instability of the model at the equilibrium when $\rho > 8$, ρ being the Hill Coefficient. [5+5+10 = 20]

3.(a) Construct a Stochastic model of the two sex population process.

(b) Find out the probability distribution of the females in the population at any time t from the above model. [10+10=20]

4. Write down the Lotka-Volterra pre-predator model.

Deduce the solution of the above system near equilibrium point and draw the trajectories. [5+15=20]

5. Construct the deterministic model of general Epidemic. Discuss the properties of the above model, considering all the threshold phenomena. [5+15=20]

p.t.o.

6. What is Catastrophe ? Give two physical examples.

If the state of a physical system can be expressed in terms of the variables $x_1, x_2, \ldots, x_n$ and some parameters $\alpha_1, \alpha_2, \ldots, \alpha_k$; demonstrate when catastrophe takes place in the system.

[5+15=20]

7. Give the axiomatic definition of measure of Information.

State the requirements to be satisfied by the logarithmic form for the Entropy. Prove that if the function $H(p_1, p_2, \ldots\ldots, p_n)$ satisfies these requirements for any values of $p_k (k = 1, 2, \ldots\ldots, n$ then

$$H(p_1, p_2, \ldots, p_n) = -\lambda \sum_{i=1}^{n} p_i \log p_i$$

where $\lambda > 0$, $\sum_{i=1}^{n} p_i = 1$

Establish the uniqueness of the Entropy function H.

[4+4+12 =20]

8.(a) Explain the following with suitable diagrams.

(i) " Premitive Brain"

(ii) " Self Re-exciting neural system"

(iii) " Delay Network"

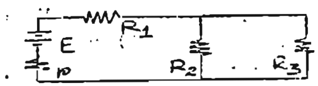
(b) Suppose a cold object is held to the skin for a moment and removed, a sensation of heat is felt; if it is applied for larger time, the sensation will be only of cold, with no preliminary warmth at all. Design this behaviour by suitabl neural network. [3+3+2+12 = 20]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

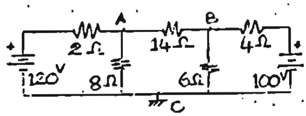
ELECTIVE-5 : PHYSICAL AND EARTH SCIENCES
SEMESTRAL-II EXAMINATION

Date : 16.5.86. Maximum Marks : 100 Time: 3 Hours.

Note : You can answer all the questions, but maximum marks allotted 100. Each question carries 20 marks. Whenever necessary draw the circuit diagram and related graphs.

1. 

State Thevenin's theorem. Using this, find the current flowing through the resistance R_3 of the given circuit. [10]



Calculate the magnitude and direction of current flowing through 14Ω resistance of the circuit. [10]

2. A coil, of self-inductance L and an appreciable resistance R, is being put across a battery of V volts (dc). Derive an expression for the growing current. With time t, sketch the graph of potential difference across and the current flowing through the inductance. [10]

After the inductive current reaches its maximum value, the coil is short-circuited. Find out the nature of the decaying current. What will be the value of potential difference across L at time $t = 0$ and 5τ ? [10]

3. Using resistance R, and capacitance C, draw the diagram of a differentiating circuit. Mention the initial assumptions. Derive an expression for current flowing through it. Give two application of this circuit. [10]

What is an integrating circuit ? Draw a circuit diagram of it using R and C. What are the necessary initial assumptions to derive output voltage of this circuit ? Find the output voltage. [10]

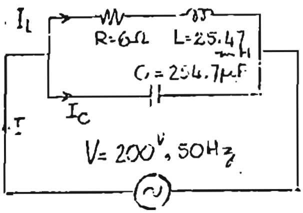
p.t.o.

4. A coil, of appreciable resistance R and self-inductance L, is joined along a capacitor C and is allowed to put across an alternating supply voltage V. After necessary deductions, draw the impedance triangle. What will be the magnitude of impedance ? [3]

Clearly show the relation between the applied voltage V and circulating current i, when the net reactance is i) greater than, ii) equal to, and iii) smaller than, zero. [7]

Under what condition this circuit is in resonance ? Derive an expression for this resonant frequency. [5]

5. Across a capacitor C, a coil of inductance L and resistance R(in series) is joined. An alternating voltage V is applied across them. Find the condition of electrical resonance and derive an expression for the resonant frequency.



Mathematically show that this parallel a.c. circuit is in electrical resonance. Find the overall line(I) current and compare it with the two branch currents I_L and I_C. [Given R = 6Ω, L = 25.47 mH C = 254.7 μF and V = 200 volt, 50 Hz] [10]

6. How does a high vacuum thermoionic diode operate ? Explain with suitable sketches. What is a rectifier ? [8]

Starting from a single phase alternating current, how do you obtain almost smooth direct current ? Explain with suitable diagrams describing each stage. [12]

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1985-86

DESIGN OF EXPERIMENTS

SEMESTRAL-II EXAMINATION

Date : 14.5.86. Maximum Marks : 100 Time: 4 Hours.

Note : Answer all questions.

1. Consider the following design with 4 treatments in 5 blocks :

1 2 3
1 2 4
1 3 4
2 3 4
1 2 3 4

(i) Write down the incidence matrix N.
(ii) Compute the C matrix.
(iii) Is the design connected ? Why ?
(iv) Is the design balanced ? Give reasons. [15]

2. Construct a BIBD with the following parameters :

$v = 25 \quad b = 30 \quad r = 6 \quad k = 5 \quad \lambda = 1$

State any theorem or result clearly that you are using to construct the design. [15]

3. Prove the following inequality for a resolvable BIBD :

$$b \geq v + r - 1$$ [15]

4. Construct two replicates of a 2^7 factorial experiment in 2^3 blocks, so that all the effects are estimable and all the main effects and first order interactions are estimable with full precision. [15]

5. There are 7 kinds of seeds of paddy A,B,C,D,E,F,G. The seeds are grown on 7 kinds of lands, each with 3 plots. The application of seeds in plots and production (in tonnes) per acre are given below.

Kinds of lands	Treatments and production		
I	B (5)	C (8)	E (3)
II	C (4)	D (7)	F (4)
III	D (5)	E (8)	G (2)
IV	E (2)	F (6)	A (6)
V	F (4)	G (3)	B (2)
VI	G (2)	A (9)	C (4)
VII	A (7)	B (8)	D (3)

Let $\tau_A, \tau_B, \ldots \tau_G$ be the treatment effects.

(i) Test the hypothesis : $\tau_A = \tau_B = \ldots = \tau_G$

(ii) Estimate : $(\tau_A - \tau_B + \tau_C - \tau_D)$ and $(\tau_A + \tau_B - 2\tau_C)$.

[12+8=20]

6. Assignments. [20]

1985-86 392

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1985-86

OPTIMIZATION TECHNIQUES

SEMESTRAL-II EXAMINATION

Date : 12.5.86. Maximum Marks : 100 Time: 4 Hours.

Note : At each step of any algorithm which you use, you have to describe the step in words. The paper carries 115 marks. The maximum you can score is 100.

1. Using Max flow Min cut theorem prove the theorem on system distinct representatives. [2

2. Player I holds cards marked 1,2 and 3 and player II holds marked 1 and 2. They play a game as follows : Each player one card from the set of his/her cards. If the sum of the values of the picked cards is even (odd) player II(I) pays amount to player I(II).
Formulate the pay off matrix, find the optimal (mixed) strategies for the two players (describe them) and find the value the game. [20]

3. Show that for any matrix A and any vector b , either $x^T A = b$ has a nonnegative solution or $Ay \geq 0$, $b^T y < 0$ has a solution and that exactly one of these two holds. [20]

4. Use Gomory's cutting plane algorithm to

$$\text{Minimize } x_1 - 2x_2$$

$$\text{Subject to } \quad 2x_1 + x_2 \leq 5$$

$$-4x_1 + 4x_2 \leq 5$$

$$x_1, x_2 \geq 0 \text{ and integral.}$$ [20]

5. A factory manager has the following table which shows how mu profit is accomplished in an hour when each of five men operate each of five machines. Find out which machine has to operated by which man so that the total profit in an hour is maximized

		Men 1	2	3	4	5
Machines	1	2	4	6	3	5
	2	5	2	0	4	3
	3	1	5	3	6	1
	4	2	2	3	1	7
	5	4	2	1	0	3

[2

6. Assignments [

1985-86 | 382

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1985-86

NONPARAMETRIC AND SEQUENTIAL METHODS

SEMESTRAL-II EXAMINATION

Date : 9.5.86. Maximum Marks : 100 Time: 3 Hours.

Note : Attempt any two questions from Part A and any two question from Part B. 22 marks are allotted to each question and 12 marks are allotted to Assignments of Part C.
Show all your work.

PART A

1.(a) Define and explain the concept of Asymptotic Relative Efficiency (ARE) of two sequences of test statistics.
State Noether's Theorem on ARE.

(b) Stating explicitly all assumptions you make, derive Pitman's ARE of the W^+, Wilcoxon signed rank statistic relative to the T^+, Student's t-statistic in the one-sample location setting with one-sided alternative.

2.(a) Deduce the test statistic for the LMP rank test corresponding to the two-sample location problem.

(b) Discuss the concept of permutation distribution. Give an example, with necessary derivations, to illustrate the advantage of application of this concept.

3.(a) Let X_i, $i=1,\ldots,n$ be a random sample from a symmetric population. Let $V(X_1,\ldots,X_n)$ be an odd translation statistic and $W(X_1,\ldots,X_n)$ an even translation-invariant statistic. Then, show that, if it exists, $Cov[V,W]=0$.

(b) Derive the expression for the Hodges-Lehmann estimator that is associated with the signed-rank test statistic.

(c) Briefly discuss the concept of M-estimators.

PART B

4. Let $X \sim N(\theta,\sigma^2)$, σ^2 known. Let $H_0 : \theta = \theta_0$, $H_1: \theta = \theta_1 > \theta_0$.
Consider the SPRT and following $\bar{X}$-test, both of strength (α = .05, β = .05). $\bar{X}$-test based on fixed sample size N gives
Reject H_0 if $\sqrt{N}(\bar{X}-\theta_0) \geq d_\alpha \sigma$
where $P[Z > d_\alpha] = .05$, $d_\alpha = 1.6449$, $Z \sim N(0,1)$.
Show that the " saving" in the observations by the sequential method is of the order of 50% under both H_0 & H_1.

p.t.o.

5. Show that, compared to any sequential test procedure which terminates with prob. 1, the ASN(H_0) for the SPRT is (approximately) the least.

6.(a) Prove the Funadmental Identity of sequential analysis.

(b) Compute the (approximate) O.C. and ASN for the normal distribution, $N(\theta,\sigma^2)$ for the SPRT for testing H_0: $\theta = \theta_0$, against H_1: $\theta = \theta_1 > \theta_0$, σ^2 known.

PART C

7. Assignments. [12]

1985-86 | 382

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1985-86

NONPARAMETRIC AND SEQUENTIAL METHODS

SEMESTRAL-II EXAMINATION

Date : 9.5.86. Maximum Marks : 100 Time: 3 Hours.

Note : Attempt any two questions from Part A and any two question from Part B. 22 marks are allotted to each question and 12 marks are allotted to Assignments of Part C.
Show all your work.

PART A

1.(a) Define and explain the concept of Asymptotic Relative Efficiency (ARE) of two sequences of test statistics.
State Noether's Theorem on ARE.

(b) Stating explicitly all assumptions you make, derive Pitman's ARE of the W^+, Wilcoxon signed rank statistic relative to the T^+, Student's t-statistic in the one-sample location setting with one-sided alternative.

2.(a) Deduce the test statistic for the LMP rank test corresponding to the two-sample location problem.

(b) Discuss the concept of permutation distribution. Give an example, with necessary derivations, to illustrate the advantage of application of this concept.

3.(a) Let X_i, $i=1,..,n$ be a random sample from a symmetric population. Let $V(X_1,..,X_n)$ be an odd translation statistic and $W(X_1,..., X_n)$ an even translation-invariant statistic. Then, show that, if it exists, $Cov[V,W]= 0$.

(b) Derive the expression for the Hodges-Lehmann estimator that is associated with the signed-rank test statistic.

(c) Briefly discuss the concept of M-estimators.

PART B

4. Let $X \sim N(\theta,\sigma^2)$, σ^2 known. Let $H_0 : \theta = \theta_0$, $H_1: \theta = \theta_1 > \theta_0$.
Consider the SPRT and following $\bar{X}$-test, both of strength ($\alpha = .05$, $\beta = .05$). $\bar{X}$-test based on fixed sample size N gives,
Reject H_0 if $\sqrt{N}(\bar{X}-\theta_0) \geq d_\alpha \sigma$
where $P[Z > d_\alpha] = .05$, $d_\alpha = 1.6449$, $Z \sim N(0,1)$.
Show that the "saving" in the observations by the sequential method is of the order of 50% under both H_0 & H_1.

p.t.o.

5. Show that, compared to any sequential test procedure which terminates with prob. 1, the ASN(H_0) for the SPRT is(approximately) the least.

6.(a) Prove the Fundamental Identity of sequential analysis.

(b) Compute the (approximate) O.C. and ASN for the normal distribution, $N(\theta,\sigma^2)$ for the SPRT for testing H_0: $\theta = \theta_0$, against H_1: $\theta = \theta_1 > \theta_0$, σ^2 known.

PART C

7. Assignments. [12]

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1985-86

MULTIVARIATE DISTRIBUTIONS AND TESTS

SEMESTRAL-II EXAMINATION

Date : 5.5.86. Maximum Marks : 100 Time : 3 Hours.

Note : Answer as many questions as possible. Submission of your practical records is compulsory. This question paper carries a total of 115 marks. Marks allotted to individual question are shown in [ ].

1. (a) Let $\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim N \left[\begin{pmatrix} \mu \\ \nu \end{pmatrix}, \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \right]$, $i = 1,2$ be independent.

   Find the joint distribution of 4 variables. [7]

   (b) Find the joint distribution of $\left(\frac{X_1+X_2}{2}, \frac{Y_1+Y_2}{2} \right)$ [8]

2. (a) Let $X \sim N_3(\mu, \Sigma)$. Find the conditional distribution of X_3 given X_1 and X_2. [5]

   (b) Let $\hat{X}_1 = \beta_1 + \sum_{s=2}^{p} \beta_s X_s$, where $\beta_1, \ldots, \beta_p$ are such that $\sum(X_1 - \hat{X}_1)^2$ is minimized. Show that

$$\rho^2_{1(2\ldots p)} = \sigma^2_{\hat{X}_1} / \sigma^2_{X_1}$$

   where $\rho^2_{1(2\ldots p)}$ is the multiple correlation coefficient of X_1 on $X_2, \ldots, X_p$ and $\sigma^2_X = \text{var}(X)$. [10]

3. (a) Prove the following theorem :

   Let $Y \sim N_p(0, \pi)$, Γ being non-singular. Then $Y' \Gamma^{-1} Y$ is distributed as χ^2.

   Give the appropriate degrees of freedom. [5]

   (b) Based on the theorem in (a) construct a test statistic for the hypothesis $H : \mu = \mu_0$ (given) for a known Σ based on a random sample $(X_1, \ldots, X_n)$ from $N_p(\mu, \Sigma)$. [5]

   (c) Find the likelihood ratio test for the above hypothesis and show that it is a function of the test statistic you proposed. [10]

4. Let A be a Wishart matrix of order p with degrees of freedom = k, variance-covariance matrix = Σ and noncentrality parameter = 0. Prove the following properties.

   (a) $\frac{\sigma^{pp}}{a^{pp}} \sim \chi^2_{k-p+1}$ and is independent of (a_{ij}), $i,j=1,\ldots,p-1$, where $\Sigma^{-1} = (\sigma^{ij})$ and $A^{-1} = (a^{ij})$. [7]

p.t.o.

4.(b) $\frac{\ell' \Sigma^{-1} \ell}{\ell' A^{-1} \ell} \sim \chi^2_{k-p+1}$ for any vector ℓ. [7]

(c) Let $\Sigma=I$, prove that,

(i) $a_{ii} \sim \chi^2$,

(ii) $\Sigma a_{ii} \sim \chi^2$, and

(iii) $\sum_i \sum_j a_{ij} \sim \chi^2$.

Specify the degrees of freedom of the Chi-square variables above. [2+2+2]

5.(a) Define Hotelling's generalized T^2 and show that it is a constant multiple of an appropriate F random variable. [2+4]

(b) Give a test in the form of T^2 based on a random sample $(X_1, \ldots, X_n)$ from $N_p(\mu, \Sigma)$ for the hypothesis $h: \mu=\mu_0$ (given) when Σ is unknown. [4]

(c) Derive the likelihood ratio test of the hypothesis in (b) and show that it is function of T^2 you defined in (b). [5]

6.(a) Describe the multivariate linear model and show that the ℓ.s. estimators obtained from the component models are b.ℓ.u.e. [5]

(b) Derive the likelihood ratio test for a general linear hypothesis in the above model. Show that the test statistic is a function of the roots of the equation

$$|R_0 - \lambda R_1| = 0,$$

where the matrices R_0 and R_1 are appropriately defined. [5]

(c) Let $A \sim W_2(n, I)$. Describe the rectangular coordinates of A and find their distribution. [5]

7. Submit your practical records. [15]

1985-86 | 381

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year: 1985-86

ELECTIVE - 5: BIOLOGICAL SCIENCES
PERIODICAL EXAMINATION

Date : 5.3.86. Maximum Marks : 100 Time: 3 Hours

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

1.(a) Discuss Michaelis - Menten Mechanism and construct the Mathematical Model of this mechanism.

(b) Find out the Michaelis constant of the above system.

[10+10]

2.(a) What is the pseudo-steady-state hypothesis of Briggs and Haldane ?

(b) What is the trouble with the pseudo-steady-state hypothesis ?

[4+16]

3.(a) Express the Michaelis-Menten system in the non-dimensional form.

(b) Analyse the above non-dimensional form using the pseudo-steady-state hypothesis and discuss the disadvantages of this hypothesis. [10+10]

4.(a) Construct a stochastic model of the uni-molecular irreversible reaction $A \rightarrow B$, treating this mechanism as a Markov Process.

(b) Solve the above model to find out $p_{ik}(t)$, i.e., the probability of transition from the state i to state k in time t.

[10+10]

5.(a) Explain " intrinsic growth rate".

(b) Deduce Malthus Model for a single species population.

(c) What are the disadvantages of this model ? Draw the curve when " intrinsic growth rate ", r, say, is less than, equal to and greater than zero. [2+10+4+4]

.(a) What is the Environmental Carrying Capacity ?

(b) Deduce Pearl-Verhulst model. [4+16]

.(a) Describe a population model in which the number of males and females maintain a constant ratio, assuming that the number of contact between males and females is proportional to the product of their individual numbers and the number of births during one contact may be greater than one. [20]

1985-86 | 391

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year: 1985-86

OPTIMIZATION TECHNIQUES

PERIODICAL EXAMINATION

Date : 3.3.86. Maximum Marks : 100 Time: 3 Hours

Note : The paper carries 107 marks. The maximum you can score is 100. Answer as many questions as you can.

1. Use the simplex algorithm to solve

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 + x_3 \\ \text{subject to} \quad & -x_1 + 2x_2 + x_3 \leq 1 \\ & -x_1 + 2x_3 \geq 4 \\ & x_1 - x_2 + 2x_3 = 4 \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

[10]

2. An oil refinery uses three different processes to produce petrol. Each process produces varying amounts of three grades (I, II and III) of petrol. These amounts in thousands of liters per day of operation, and the cost of per day of operation are given in the following table.

	grade I	grade II	grade III	Cost
process A	3	4	2	1600 Rs.
process B	6	6	8	4000
process C	6	3	4	3000

The refinery is expected to produce at least 3600 liters of grade I, at most 2000 liters of grade II and at least 3000 liters of grade III per week. Using the simplex algorithm determine the operation of the refinery that satisfies the conditions of production and minimizes the cost. [20]

3. Use the Dual Simplex algorithm (and not by other methods) and solve the LP

$$\begin{aligned} \text{minimize} \quad & 3x_1 + x_2 \\ \text{subject to} \quad & x_1 + x_2 \geq 1 \\ & 2x_1 + 3x_2 \geq 2 \\ & x_1, x_2 \geq 0. \end{aligned}$$

[10]

p.t.o.

4. Solve the above problem (of (3)) geometrically. [10]

5.(a) Show from the fundamentals that the dual of the LP

$$\left.\begin{array}{ll} \text{maximize} & \underset{\sim}{c}^T \underset{\sim}{x} \\ \text{subject to} & A\underset{\sim}{x} = \underset{\sim}{b} \quad \underset{\sim}{x} \geq 0 \end{array}\right\} (*)$$

is

$$\left.\begin{array}{ll} \text{minimize} & b^T \underset{\sim}{y} \\ \text{subject to} & A^T \underset{\sim}{y} \geq \underset{\sim}{c} \end{array}\right\} (**)$$

(b) Assume the Duality theorem and prove the following result. If $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are feasible solutions of (*) and (**) respectively then $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are optimal iff $x_i = 0$ whenever $\underset{\sim}{a}_i^T \underset{\sim}{y} > c_i$ where $\underset{\sim}{c}^T = (c_1 \dots c_n)$ and $\underset{\sim}{a}_i$ is the ith column vector of A.

(c) Use the result of (b) to show that $(0, \frac{4}{7}, \frac{12}{7}, 0, 0)$ is an optimal solution of

$$\begin{array}{ll} \text{maximize} & -x_1 - 6x_2 + 7x_3 - x_4 - 5x_5 \\ \text{subject to} & 5x_1 - 4x_2 + 13x_3 - 2x_4 + x_5 = 20 \\ & x_1 - x_2 + 5x_3 - x_4 + x_5 = 8 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{array}$$

[Hint: Construct the Dual and a feasible solution for the dual given by the necessary conditions of (b)). [10+10+10]

6. <u>Without</u> using the simplex algorithm find the optimal value (not the optimal vector) of the following LP if it exists.

$$\begin{array}{ll} \text{minimize} & 2x_1 - x_2 - 2x_3 + x_4 \\ \text{subject to} & x_1 + 2x_2 + x_3 - x_4 = 0 \\ & 2x_1 - x_2 + 3x_3 - 2x_4 = 0 \\ & x_1 + 4x_2 - x_3 + 4x_4 = 0 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{array}$$

[8]

7. Resilve the degeneracies by using the Lexicographic ordering and solve by the simplex method the following LP

$$\begin{array}{ll} \text{minimize} & 2x_1 - x_2 - 2x_3 \\ \text{Subject to} & -x_1 + 2x_2 + x_3 + x_4 = 0 \\ & 2x_1 - x_2 + 3x_3 + x_5 = 0 \\ & x_1 + 4x_2 - x_3 + x_6 = 0 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{array}$$

[12]

8. Assignment. [7]

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year: 1985-86

DESIGN OF EXPERIMENTS
PERIODICAL EXAMINATION

Date : 28.2.86. Maximum Marks : 100 Time: 3 Hours

1. Three treatments A,B,C are to be compared by a completely randomized design where A is replicated a times, B b times and C c times. The yields of the plots where B or C is applied have variance σ^2 whereas that, when A is applied, is $2\sigma^2$, where σ^2 is an unknown constant.
   (i) Work out the estimates: $\hat{A} - \hat{B}$, $\hat{B} - \hat{C}$, $\hat{C} - \hat{A}$
   (ii) Work out $V(\hat{A} - \hat{B})$, $V(\hat{B} - \hat{C})$, $V(\hat{C} - \hat{A})$
   (iii) If $\frac{1}{3}[V(\hat{A} - \hat{B}) + V(\hat{B} - \hat{C}) + V(\hat{C} - \hat{A})]$ is to be minimized, then how would you choose a,b,c given that a+b+c=N, a fixed number. [6+6+8 = 20]

2.(i) Define a BIBD with parameters (v,b,r,k,λ)
  (ii) Prove Fisher's inequality
  (iii) Construct an SBIBD with parameters (13,4,1). [4+8+8=20]

3. Consider a latin square design where the latin square is in the standard form. After the experiment was performed, it was found that no yield in the first row has been reported. While analysing the data, the model is assumed to be

$$y_{ij(k)} = \mu + \alpha_i + \beta_j + \tau_k + e_{ij(k)}$$

$$i = 2,3, \ldots t$$
$$j = 1,2, \ldots t$$
$$k = 1,2, \ldots t$$

with constraints $\sum_{i=2}^{t} \alpha_i = 0$, $\sum_{j=1}^{t} \beta_j = 0$, $\sum_{k=1}^{t} \tau_k = 0$.

Obtain the estimates for μ, α_i's, β_j's and τ_k's. [20]

p.t.o.

4. An experiment was conducted to test whether the average I.Q. of 4 different human races (R_1, R_2, R_3, R_4) differs significantly. 2 persons from each race were taken. There are two different types of I.Q. test, namely A and B. The following table gives the data obtained from the experiment

Table

Races \ Tests	A	B
R_1	105	95
R_2	110	120
R_3	80	85
R_4	145	135

(i) Considering it to be a RBD, analyse the data to test whether the effects of different races is significantly different.

(ii) Arrange the races in decreasing order of effects on I.Q. Also test whether I.Q's of races R_1 and R_2 differ significantly. [10+10 = 20]

5. Assignments. [20]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1985-86

MULTIVARIATE DISTRIBUTION AND TESTS
PERIODICAL EXAMINATION

Date : 26.2.86. Maximum Marks : 100 Time: 3 Hours

Note : Attempt any four questions.

1.(a) Let $\underset{\sim}{X}$ be a random vector. Define the dispersion matrix $\mathscr{D}(\underset{\sim}{X})$ of $\underset{\sim}{X}$. Show that

$$\mathscr{D}(\underset{\sim}{X}) = \mathscr{E}(\underset{\sim}{X}\underset{\sim}{X}') - \underset{\sim}{\mu}\underset{\sim}{\mu}'$$

where $\underset{\sim}{\mu} = \mathscr{E}(\underset{\sim}{X})$. (2+4)

(b) Suppose that the first p-1 components of $\underset{\sim}{X}$ are independent, each having the same variance σ^2 and $X_p = \sum_{i=1}^{p-1} X_i$. Show that

$$\mathscr{D}(\underset{\sim}{X}) = \sigma^2 \begin{bmatrix} I_{p-1} & \vdots & \underset{\sim}{1}_{p-1} \\ \cdots & \cdots & \cdots \\ \underset{\sim}{1}'_{p-1} & \vdots & (p-1) \end{bmatrix}$$

where I_r is identity matrix of order r and $\underset{\sim}{1}_r$ denotes a column vector of unity of order r. Is $\mathscr{D}(\underset{\sim}{X})$ non-singular ? If not why not ? (3+1+1)

(c) Let the joint density of $\underset{\sim}{X}$ be given by

$$f(\underset{\sim}{x}) = (\text{const}) \exp\left\{-\tfrac{1}{2}(\underset{\sim}{x} - \underset{\sim}{b})'A(\underset{\sim}{x} - \underset{\sim}{b})\right\}$$

where A is positive definite and $\underset{\sim}{b}$ is a fixed vector. Show that

(i) $\mathscr{E}(\underset{\sim}{X}) = \underset{\sim}{b}$ and (ii) $\mathscr{D}(\underset{\sim}{X}) = A^{-1}$. (14)

2.(a) Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$. Consider a nonsingular linear transformation $\underset{\sim}{Y} = B\underset{\sim}{X}$, (B is n.s.). Show that $\underset{\sim}{Y} \sim N_p(B\underset{\sim}{\mu}, B\Sigma B')$. (6)

(b) Decompose $\underset{\sim}{X}$ into subvectors $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$, $\underset{\sim}{X}^{(1)}$ having q components and $\underset{\sim}{X}^{(2)}$ having (p-q) components. Decompose $\underset{\sim}{\mu}$ and Σ correspondingly. Prove in the case $\Sigma_{12} = 0$, that $\underset{\sim}{X}^{(i)} \sim N(\underset{\sim}{\mu}^{(i)}, \Sigma_{ii})$, i = 1,2. (6)

p.t.o.

2.(c) Using (b) or otherwise show that even if $\Sigma_{12} \neq 0$

$\underset{\sim}{X}^{(1)} \sim N(\underset{\sim}{\mu}^{(1)}, \Sigma_{11})$. (5)

(d) Show that the conditional distribution of $\underset{\sim}{X}^{(1)} | \underset{\sim}{X}^{(2)} = \underset{\sim}{x}^{(2)}$ is $N_q(\underset{\sim}{\nu}_1, \Sigma_{11.2})$ where

$$\underset{\sim}{\nu}_1 = \underset{\sim}{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{x}^{(2)} - \underset{\sim}{\mu}^{(2)})$$

and $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. (5)

(e) Use (d) to show in the bivariate case that the conditional distribution of X_1 given $X_2 = x_2$ is

$N(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1-\rho^2))$ where

$\mu_i = E(X_i)$, $\sigma_i^2 = \text{var}(X_i)$, and $\rho = \text{cor}(X_1, X_2)$. (3)

3.(a) Define the partial correlation between X_i, X_j given $X_{r+1}, \ldots, X_p$; $i,j = 1,2, \ldots, r$. Prove the following formulae

(i) $$\rho_{12.3} = \frac{\rho_{12} - \rho_{13}\,\rho_{23}}{\sqrt{(1-\rho_{13}^2)}\sqrt{(1-\rho_{23}^2)}}$$

(ii) $$\rho_{ij.r+1,\ldots,p} = \frac{\rho_{ij.r+2,\ldots p} - \rho_{i,r+1.r+2\ldots p}\,\rho_{j,r+1.r+2\ldots p}}{\sqrt{(1-\rho_{i,r+1.r+2\ldots p}^2)}\sqrt{(1-\rho_{j,r+1.r+2\ldots p}^2)}}$$

(2+4+6)

(b) Define the multiple correlation coefficient of X_1 on $X_2, \ldots, X_p$. Show that if $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, then the multiple correlation coefficient defined above is the same as $\text{cor}(X_1, \hat{X}_1)$ where

$$\hat{X}_1 = \mathcal{E}(X_1 | \underset{\sim}{X}^{(2)}) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{X}^{(2)} - \underset{\sim}{\mu}^{(2)}).$$

(4+9)

4.(a) Suppose $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ i.i.d. $N_p(\underset{\sim}{\mu}, \Sigma)$. Show that

$\bar{\underset{\sim}{X}} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S = \frac{1}{(n-1)}\sum_{i=1}^{n}(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})(\underset{\sim}{X}_i - \bar{\underset{\sim}{X}})'$ are unbiased for $\underset{\sim}{\mu}$ and Σ respectively. You may quote the univariate results you are using. (4)

contd....

4.(b) Prove that $\underset{\sim}{\bar{X}}$ and $\frac{(n-1)}{n}$ S are joint MLE's of $\underset{\sim}{\mu}$ and Σ respectively. (12)

(c) Find the MLE of Σ when it is known that $\mu = \mu_0$ (given). (9)

5.(a) Let $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ be independent random vectors such that $\underset{\sim}{X}_i \sim N_p\ (\underset{\sim}{\mu}_i, \Sigma)$. Then for orthogonal $C = (C_{ij})$. Prove that $\underset{\sim}{Y}_i = \sum_j C_{ij} \underset{\sim}{X}_j \sim N(\underset{\sim}{\nu}_i, \Sigma)$, $\underset{\sim}{\nu}_i = \sum_j C_{ij} \underset{\sim}{\mu}_j$, $\underset{\sim}{Y}_1, \ldots \underset{\sim}{Y}_n$ are indep. Also show that

$$\sum_i \underset{\sim}{Y}_i \underset{\sim}{Y}_i' = \sum_i \underset{\sim}{X}_i \underset{\sim}{X}_i'.$$ (10+5)

(b) Using the above theorem prove that

(i) $\underset{\sim}{\bar{X}} \sim N_p\ (\underset{\sim}{\mu}, \frac{1}{n} \Sigma)$

(ii) $\underset{\sim}{\bar{X}}$ and S are independent. (5+5)

1985-86 | 381

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year:1985-86

NONPARAMETRIC AND SEQUENTIAL METHODS
PERIODICAL EXAMINATION

Date : 24.2.86. Maximum Marks : 75 Time: 3 Hours

Note : Show all your work. Attempt questions 1 and 2 and any two from questions 3,4 and 5.

1. Stating explicitly the definitions and the results you need to assume, prove :
   (a) If $X_1, \ldots, X_n$ is a random sample from a population with a distribution symmetric about μ, then an odd translation statistic is symmetrically distributed about μ.
   (b) If two random variables V and W satisfy
   $$(V - \mu, W) \stackrel{d}{=} (\mu - V, W)$$
   for some constant μ, then, if it exists, Cov (V,W) is zero
   Give an example, each for (a) and (b), to illustrate the corresponding result. (6+10+4)

2. Describe the Wilcoxon-Signed rank test. Derive the asymptotic null distribution of the corresponding test statistic. (15)

3. State and prove the one-sample U-statistic theorem. (20)

4. Explain the notion of a LMP rank test. Deduce the test statistic for the LMP rank test corresponding to the two-sample location problem. (20)

5. (a) Briefly describe the one and two sample non-parametric tests for the location parameter(s). Compute the expectations and variances, under H_0, for three of these test statistics.
   (b) Describe and discuss a non-parametric test to detect whether two independent samples have been drawn from the same population. (10+10)

1985-86 | 332(b)

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1985-86

SEMESTRAL - I BACKPAPER EXAMINATION

Statistical Inference

Date: 31.12.1985 Maximum Marks: 100 Time: 3 hrs.

Note: Answer all questions. Question 8 has an alternative.

1. Show that, for $0 < p < 1$, sample pth fractile is consistent estimator of population pth fractile when population pth fractile is uniquely defined. [10]

2. Prove that an MLE is a function of sufficient statistic. [10]

3. Show that when original data are replaced by sufficient statistics, Fisher information remains unchanged. [6]

4. Give a consistent estimator of Fisher Information based on n iid Bernoulli observations. [5]

5. State Rao-Blackwell theorem. [3]

6. $X_1 = 2$, $X_2 = 1$, $X_3 = 2$, $X_4 = 1$, $X_5 = 1$ are iid observations on the number of deaths in one car accidents with at least one death. Assume that the observations are from a Poisson distribution with parameter λ truncated at zero.

   Compute UMVUE of λ. [25]

7. Let the distribution of X belong to one parameter exponential family with parameter θ. Show that, for any test φ, θ_1 and θ_2 with $\theta_1 < \theta_2$, there exists a test φ^* such that

   (a) $E_{\theta_i} \varphi^*(X) = E_{\theta_i} \varphi(X)$, $i = 1, 2$.

   Also show that for any such test φ^* of (a) above,

   (b) $E_{\theta} \varphi^*(X) \leq E_{\theta} \varphi(X)$, $\theta_1 < \theta < \theta_2$.

   (18+6) = [24]

p.t.o.

8. Either

Show that the usual t-test is UMP unbiased for testing

$$H_0 : \left| \mu - \mu_0 \right| \leq \delta_0 \text{ against } H_1 : \left| \mu - \mu_0 \right| > \delta_0 ,$$

where μ_0 and δ_0 are given.

Or

The following gives the weights of the anterior muscles of both hind legs of 16 normal rabbits. Construct a UMP unbiased size $\alpha = 0.05$ test to test the null hypothesis that the two legs differ atmost by 0.05 gms in respect of the anterior muscle weight.

Muscle weights (grams)

Rabbit number	Leg left	Leg right	Rabbit number	Leg left	Leg right
1	5.0	4.9	9	5.3	5.2
2	4.8	5.0	10	5.3	5.5
3	4.3	4.3	11	5.3	5.5
4	5.1	5.3	12	5.9	5.9
5	4.1	4.1	13	6.5	6.8
6	4.0	4.0	14	6.3	6.3
7	7.1	6.9	15	6.6	6.6
8	5.9	6.3	16	6.2	6.3

[17]

1985-86	322(b)

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1985-86

SEMESTRAL - I BACKPAPER EXAMINATION

Stochastic Processes - 2

Date: 31.12.1985 Maximum Marks: 100 Time: 3 hours

Note: Maximum Marks you can score is 100.

1.(a) What is meant by a stopping time for a sequence of random variables ?

(b) Suppose $(X_n)_{n\geq 1}$ are i.i.d. with finite mean μ. Let N be a stopping time for this sequence. Show that

$$E\left(\sum_{1}^{N} X_i\right) = \mu\, E(N).$$

(c) If (N_t) is a renewal process, then show that

$$\frac{E(N_t)}{t} \longrightarrow \frac{1}{\mu} \text{ as } t \longrightarrow \infty.$$

(5+7+13) = [25]

2.(a) What is a Nonhomogeneous Poisson Process ?

(b) Show that the output process of an $M/G/\infty$ Queue is a nonhomogeneous Poisson Process.

(10+15) = [25]

3.(a) Consider a discrete time Markov Chain with countable state space. When do you say that an initial distribution is stationary ?

(b) Consider a stationary Markov Chain. When do you say that it is reversible ?

(c) Consider an irreducible stationary Markov Chain. Make the following statement precise and then prove it. The chain is reversible iff for any state i the probability of any path from i to i is same as the probability of the reversed path.

(5+5+10) = [20]

p.t.o.

4. Consider a continuous time Markov Chain with two states 0, 1. Waiting times at the states 0 and 1 be assumed exponential with parameters λ and μ respectively.

   (a) Write down the Kolmogorov's forward equations for $p_{ij}(t)$.

   (b) Solve the differential equations and explicitly calculate $p_{ij}(t)$ $i, j = 0, 1$.

(10+10) = [20]

5. Let (X_t) be a standard Brownian motion. Let $a > 0$. Let T be the hitting time of a for the Brownian motion.

   (a) Define T mathematically.

   (b) Evaluate $P(T \leq t)$.

   (c) Show that T is finite almost surely.

   (d) Show that $E(T) = \infty$.

(3+7+3+7) = [20]

1985-86 312(b)

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1985-86

DIFFERENCE AND DIFFERENTIAL EQUATIONS
SEMESTRAL-I BACKPAPER EXAMINATION

Date : 30.12.85. Maximum Marks : 100 Time: 3 Hours

Note : Answer as many questions as you can. Total marks of all questions is 110. You can score a maximum of 100.

1.(a) State and prove the nature of the Euler-Lagrange equations satisfied by the extremal curves for the functional

$$v[y_1(x), y_2(x)] = \int_{x_0}^{x_1} F(x, y_1, y_2, y_1', y_2')\, dx$$

You must explain your class of admissible curves.(You may assume the fundamental lemma of the calculus of variations but you must explain its statement and application). [12]

(b) Find the extremal curves for the functionals.

(i) $v[y(x)] = \int_{x_0}^{x_1} \left(x \frac{dy}{dx} + y \right) dx$ [6]

(ii) $v[y(x)] = \int_{x_0}^{x_1} \frac{\sqrt{1+y'^2}}{y}\, dx$

You must find the extremals passing through (x_0, y_0) and (x_1, y_1). [6]

2. State and prove a theorem about real analytic solution to the initial value problem

$$\frac{d^2y}{dx^2} + P(x,y)\frac{dy}{dx} + Q(x,y) = 0$$

with $y(0) = y'(0) = 0$

where P and Q are real analytic functions of (x,y) in a rectangle $|x| \leq a$, $|y| \leq b$. You must use the majorant method. [18]

p.t.o.

3.(a) Prove that the Laplace transform of $\frac{1}{2}(\sin x - x\cos x)$ is $\frac{1}{(p^2+1)^2}$. Hence solve

$$\frac{d^4y}{dx^4} + 2\frac{d^2y}{dx^2} + y = 0$$

given $y = 0,\ y' = 1,\ y'' = 0,\ y''' = 0$ at $x = 0$. [1:

(b) Find the general solutions of

(i) $\frac{d^4x}{dt^4} - 2\frac{d^3x}{dt^3} - 3\frac{d^2x}{dt^2} + 8\frac{dx}{dt} - 4x = 0.$ [5:

(ii) $y'' + 4y = \tan 2x.$ [8:

4. A spherical raindrop of radius a falls from rest through a vertical height h. Throughout motion it accumulates condensed vapour at the rate of k grammes per square centimeter per sec. Show that at the end of its fall its radius is

$$k\sqrt{\frac{2h}{g}}\cdot\left[1+\sqrt{1+\frac{ga^2}{2hk^2}}\right].$$

(Only gravity acts. No air resistance). [12

5. Prove that

$$M(x,y)\,dx + N(x,y)dy$$

is an exact differential if and only if $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Here M and N are C^1 functions on a disc $\{(x,y) : x^2+y^2 < R^2\}$. [1:

6.(a) Find the general solution of

$$t_{k+2} - 5t_{k+1} - 6t_k = 3^k$$ [6.

(b) State and prove the Banach contraction principle for fixed points in complete metric spaces. [5.

1985-86	352(b)

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1985-86

SEMESTRAL - I BACKPAPER EXAMINATION

Elective - 4: Physical and Earth Sciences

Date: 30.12.1985 Maximum Marks: 100 Time: 3 hrs.

Note: Attempt any five questions. Maximum score for each full question is 20. Maximum score for each subdivision in a full question is indicated along the right margin. Answers should be brief and to the point. Draw sketches wherever necessary.

1.(a) Give the main features of explosive volcanism. Name two volcanic rocks and indicate the structures associated with them. (3+4) = [7]

(b) Describe the following textures:

hypidiomorphic and porphyritic. (3x2) = [6]

(c) Arrange the following rocks in order of increasing S_1O_2 content.

Andesite, Dunite, Granite, Granodiorite, Gabbro.

Give the essential mineral composition of one of the above rocks. (5+2) = [7]

2.(a) Name three basic controlling factors of metamorphic changes and elaborate on them. (3+6) = [9]

(b) What are the metamorphic equivalents of shale, limestone and basalt ? [3]

(c) What is a metamorphic grade ? Arrange the following index minerals in order of increasing metamorphic grade —— biotite, chlorite, garnet, sillimanite. (2+2) = [4]

(d) Name one typical metamorphic rock from Eastern Ghats of India and indicate the condition of its formation. (1+3) = [4]

p.t.o.

3.(a) Distinguish between brittle and plastic deformation. [3]

(b) Draw a folded layer labelling the following features:

antiform, synform, inflection point, hinge, limb. [5]

(c) Give the main features of shear folding. [5]

(d) Define the following terms associated with an inclined fault plane :

strike, dip, netslip. [6]

(e) Which of the following lies above the fault plane:

hanging wall, footwall ? [1]

4.(a) What sort of tectonic province does the Indian peninsula represent ? Give the main features of such a province.

(1+4) = [5]

(b) What is an unstable region in the earth's crust ? Name three types of unstable regions and give their characteristics.

(2+3+6) = [11]

(c) What sort of lava does well up along the Mid-Atlantic Ridge and how does the age of the lava flow charge as one proceeds towards the African coast ? (2+2) = [4]

5.(a) Give the geologic evidences in favour of Continental Drift. Name the continental masses which constituted the Gondwanaland.

(5+2) = [7]

(b) What is a magnetic anomaly on the ocean-floor ? Indicate how the magnetic anomaly pattern on the ocean floor can be used to suggest Sea-floor Spreading.

(2+4) = [6]

(c) What is a subduction zone ? Give the geophysical features associated with such a zone.

(3+4) = [7]

6.(a) How would you graphically represent data on grain size measurements and cross-bedding azimuth ?

(4+4) = [8]

Contd..... 3/-

Contd..... Q.No.6

(b) What is ϕ - scale ? Give the ϕ values associated with following grades:

pebbles, very fine sand.

(2+2) = [4]

Hint:

Cobbles	Fine sand
— 64 mm —	— 0.125 mm —
Pebbles	Very fine sand
— 4 mm —	— 0.0625 mm —
Granules	Silt.

(c) In an area successive beds are exposed in parallel stream sections running transverse to bed strike. Two fossils H and T occur together in 14 of the observed stream sections. In 10 of these H continues in beds occurring above those in which both H and T are found. In the rest of the stream sections T continues in beds occurring above those in which both H and T are found. How would you proceed to test the null hypothesis that H and T have equal range of stratigraphic occurrence in the area.

[8]

Hint: Coin tossing experiment. One-tail critical values of the test statistic R (binomial distribution) is as follows:

No. of trials	Size of Critical region	
N	a = 0.05	a = 0.01
14	R = 3	R = 2

THE GEOLOGIC COLUMN

			Age in million years
CENOZOIC	QUATERNARY	Holocene	
		Pleistocene	1.5-3.5
	TERTIARY	Pilocene	7
		Miocene	26
		Oligocene	37-38
		Eocene	53-54
		Paleocene	64-65
MESOZOIC	CRETACEOUS		136
	JURASSIC		180
	TRIASSIC		225
PALAEOZOIC	PERMIAN		280
	CARBONIFEROUS		345
	DEVONIAN		395
	SILURIAN		500
	CAMBRIAN		570
PRE CAMBRIAN			

Contd..... 2/-

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

SEMESTRAL - I BACKPAPER EXAMINATION

Sample Surveys

Date: 27.12.1985 Maximum Marks: 100 Time: 4 hrs.

1.(a) Define a 'sampling design'. What do you understand by the terms 'inclusion probability of a unit' and 'joint inclusion probability of a pair of units' for a sampling design.

(2+2+3) = [7]

(b) For a probability proportional to size with replacement sampling design of n draws write down π_i, the probability of inclusion of a unit U_i and π_{ij}, the joint inclusion probability of a pair of units (U_i, U_j).

(2+3) = [5]

2.(a) For stratified simple random sampling (without replacement) to estimate the population mean, write down the optimum allocation of a fixed total sample size n to the strata, explaining clearly your notations. How does one implement this allocation in practice ?

(2+3) = [5]

(b) In the above situation, suppose that the actual (a) allocation in practice turns out to be n_i^a for the ith stratum while the optimum (O) allocation is n_i^o. Obtain an expression for the relative loss of efficiency measured by

$$\frac{\text{Var.}_a(\hat{\bar{Y}}_{st}) - \text{Var.}_o(\hat{\bar{Y}}_{st})}{\text{Var.}_o(\hat{\bar{Y}}_{st})}$$

where the notation is self explanatory and the stratum sizes are assumed large. Further, derive a **quick** upper bound to the above expression in terms of θ, the relative deviation of sample allocations given by $\theta = \max_i \left|\frac{n_i^o - n_i^a}{n_i^a}\right|$.

(7+3) = [10]

Contd..... 2/-

Contd..... Q.No.2

(c) Write down the 'combined and separate regression estimators' in stratified random sampling. Which of these do you recommend ? Give reasons.

[6]

3.(a) Define the term 'Intra Cluster Correlation Coefficient'. For a population of 14 clusters each of size 6, find the lower and upper bounds for the intra cluster correlation coefficient among the elements of the cluster.

(3+3) = [6]

(b) A population consists of N clusters of varying sizes M_i, $i = 1, 2, \ldots, N$. Suppose that n clusters are selected with probabilities proportional to the cluster size and with replacement. Write down an unbiased estimator for the population mean $\overline{Y} = \sum_{i=1}^{N} M_i \overline{Y}_i / \sum_{i=1}^{N} M_i$ and an unbiased estimator of its variance, where $\overline{Y}_i$ is the ith cluster mean, $i = 1, 2, \ldots, N$. On the other hand, if the estimator $t = \sum_{i=1}^{n} M_i \overline{Y}_i / \sum_{i=1}^{n} M_i$ is used, suggest a sampling scheme that makes t unbiased for $\overline{Y}$.

(3+4+4) = [11]

4. A sample survey was conducted to estimate the total household expenditure in an urban area. The design adopted was a stratified two-stage one with census enumeration blocks as first stage units and households within them as second stage units. From each stratum 4 blocks were selected with probability proportional to population and with replacement and 4 households were selected from each selected block with equal probability and without replacement. The data on household expenditure for the sample households together with information on selection probabilities are given below :

Contd..... 3/-

Contd..... Q.No.4

Stratum	Sample block	Inverse of probability of selection	Total no. of households	Weekly household expenditure of sample households			
				1	2	3	4
I	1	67.68	189	110	281	120	114
	2	338.12	40	80	60	122	125
	3	101.50	135	122	210	171	105
	4	69.03	160	244	115	312	128
II	1	113.34	73	345	359	160	117
	2	441.00	26	97	179	144	85
	3	31.50	240	100	115	50	172
	4	661.57	14	102	40	126	148
III	1	15.80	287	122	176	108	140
	2	21.00	257	125	110	134	215
	3	48.89	68	300	115	67	110
	4	26.73	218	263	75	142	54

Total no. of households in Stratum I, II, III are 12848, 8422, 6354 respectively.

(a) Obtain an unbiased estimate of the total weekly household expenditure. [10]

(b) Obtain an unbiased estimate of the sampling variance of the above estimate. [15]

(c) Compare the efficiency of the above design with that of unistage simple random sampling with replacement of households in each stratum. [12]

5. For estimating the total Y of current population in a region, two subsamples of 6 villages each are selected (circular systematically) from each stratum with independent random starts. Using the data given in the table, obtain a ratio estimate for Y taking the previous census population (x) as the auxiliary information. Also obtain an unbiased estimate of the sampling error of your above estimate.

Total number of villages (N) and sample totals of x and y

Stratum Number	No. of villages N	sub-sample 1		sub-sample 2	
		x	y	x	y
1	2044	3722	3935	3456	3641
2	1304	3625	4033	4171	4649
3	1265	2769	3050	3746	4043

(Total of x for the region = 3,155,680). (8+5) = [13]

[1985-86] [352]

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year : 1985-86

ELECTIVE-4 : PHYSICAL AND EARTH SCIENCES

SEMESTRAL-I EXAMINATION

Date : 28.11.85. Maximum Marks : 100 Time: 3 Hours

Note : Attempt any five questions. Maximum score for each full question is 20. Maximum score for each subdivision in a full question is indicated in the right margin. Answers should be brief and to the point. Draw sketches wherever necessary.

1. What is the most common oxide in the chemical make-up of crustal rocks ? Name three bases of chemical classification of igneous rocks. [2+3]

How would you distinguish a plutonic rock from its volcanic equivalent ? Name two families of intermediate plutonic rocks and show how the broad distinction between them and further subdivision in each of them can be effected with the help of mineral compositions. [3+2+8]

What is the volcanic equovalent of diorite ? Name two textures expected in this volcanic rock and indicate its common geologic setting. [1+2+2]

2. Separate the processes from the products in the following set; arrange them in two columns matching each product against the relevant process.

Ash beds, Coesite, Compression of layered plastic material, Contact aureole, Differentiation, Fast cooling, Fine-grained igneous rocks, Folds, Igneous rocks of different mineral composition but common parentage, Incongruent melting, Magmatism, Migmatites, Partial melting, Plutonic rocks, Reaction rims, Sandstone, Sedimentation, Shock metamorphism, Thermal metamorphism, Volcanism. [10+10]

3.(a) According to one theory Gondwanal and broke up in Cretaceous. Which of the following are necessary corollaries. [5]

(i) Oldest sediments of the Atlantic Ocean are Cretaceous.

(ii) Oldest lavas of the Atlantic floor are Cretaceous.

(iii) Oldest sediments of the Tethys are Cretaceous.

p.t.o.

3.(a)(iv) Pre-Cretaceous sediments of Peninsular India are intra cratonic.

(v) Jurassic fauna of India and Africa are likely to be similar.

(vi) Post-Cretaceous polar-wandering curves for Africa and Australia are similar.

(vii) Closing of the Tethys began in Tertiary times.

(b) Indicate the geophysical record on the ocean floor which le to the hypothesis of Sea-floor Spreading. [4]

(c) What is a convergent plate boundary ? Describe the tectonic features associated with such a boundary. [2+6]

4.(a) What is plastic deformation ? What sort of structure would you expect in case of plastic heterogeneous deformation ? [2+1]

(b) Draw neat sketches to show synformal anticline and antiformal syncline; label hinge on the first and axial plane on the second. [2+2+1+1]

(c) What is flexure-slip folding ? Indicate the geometric form of fold so produced. [2+1]

(d) Draw a sketch to illustrate horst and graben structure. Where in Africa and Europe do extensive graben structure occur ? [4+3]

5. Write notes on :

(a) Main tectonic provinces of the Indian subcontinent;

(b) Airy model of isostatic adjustment;

(c) Earth's core;

(d) Mid-Oceanic Ridges. [4x5 = 20]

6.(a) Name two geologic parameters, one scalar and the other vector, on which quantitative measurements can be made. [2]

(b) Describe how you would measure grain size in unconsolidated sands. [5]

(c) What type of distribution is generally observed in the grain size analysis of clastic sedimentary rocks ? [3]

(d) Discuss the applicability of skewness, sorting and percentile measures in the discrimination of sediments of different environments. [10]

[1985-86|35.(P)]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1985-86

ELECTIVE-4 : BIOLOGICAL SCIENCES - PRACTICAL
SEMESTRAL-I EXAMINATION

Date : 28.11.85. Maximum Marks : 100 Time : 2 Hours

1. Determine the ABO blood groups of the given blood samples by slide technique against

(i) known anti sera
(ii) known red cells

[1985-86 | 252]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) II Year : 1985-86

ELECTIVE-2 : ECONOMICS
SEMESTRAL-1 EXAMINATION

Date : 28.11.85. Maximum Marks: 100 Time: 3 Hours.

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

1. Consider an individual living for two periods with given money income I_1 and I_2 in the two periods. There is a set of indifference curves representing his preference over C_1 and C_2 where C_1 and C_2 are consumption expenditures (in terms of money) in the two periods. Assume that the individual is free to borrow or lend at a given market rate of interest r.

   (i) Write down his budget constraint and represent it in a diagram. What does the slope of the budget line represent? [5]

   (ii) Would the individual be worse off if the option of borrowing or lending is taken away from him ? Why ? [5]

   (iii) Suppose the rate of interest rises. How would it affect the welfare level of the individual ? [7]

   (iv) What would the budget line look like if (a) the borrowing rate of interest is different from the lending rate of interest ? (b) the individual has the option of lending but not borrowing ? [8]

2. Consider a price taking firm producing a single output q with two inputs x_1 and x_2. The production function is given by

   $q = Ax_1^{\alpha} x_2^{1-\alpha}$ where $A > 0$ and $0 < \alpha < 1$ are given constants.

   The price of the final output is given by p and the prices of the two inputs are given by w_1 and w_2.

   (i) Defining returns to scale, show that the production function exhibits constant returns to scale. [4]

   (ii) Find out the elasticity of substitution between the two inputs. [4]

p.t.o.

2.(iii) Can you determine the profit maximizing output level of the firm ? Give reasons for your answer. [10]

(iv) Does your answer to (iii) change if the firm is not a price taker in the commodity market, i.e. if instead it faces an inverse demand curve for its product of the form $p = f(q)$? [7]

3. Suppose that a single-output producer sells his output in two separate markets. In the first market he is a price taker and in the second market he acts as a monopolist. Discuss the profit maximizing behaviour of the producer. In which of the markets will the producer charge a higher price ? Why ? How is the equilibrium of the producer affected if suddenly it becomes possible for everybody to transport goods costlessly from one market to the other ? [25]

4. Consider an economy producing two goods (X_1 and X_2) with two factors of production labour (L) and land (V). To produce one unit of good 1 it is necessary to employ 1 unit of L and 1 unit of V. To produce one unit of good 2 it is necessary to employ 1 unit of L and 4 units of V. The total amount of L and V available in the economy are 10 units and 20 units respectively. Assume perfect competition and free mobility of factors. Finally, the relative demand for the two outputs is given by $X_1 / X_2 = \alpha \frac{P_2}{P_1}$ where $\alpha > 0$ is a constant and P_1, P_2 are commodity prices.

(i) Represent diagramatically the production possibility frontier of the economy. [5]

(ii) Find out the range of values of α for which both L and V are fully employed. [15]

(iii) If $\alpha = 1$, determine the factor prices.
(You can assume one of the goods to be the numeraire) [5]

5. Discuss the role of income and substitution effects in determining the stability of a two-good, two-person exchange economy. If income effects cancel out in the aggregate, would you expect the equilibrium to be stable ? Briefly indicate why the analysis gets complicated if there are more than two goods in the economy. [25]

1985-86 | 354

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

ELECTIVE-4 : BIOLOGICAL SCIENCES - THEORY
SEMESTRAL-I EXAMINATION

Date : 28.11.85. Maximum Marks : 100 Time: 3 Hours

Note : Answer 6 questions, at least 2 from each group.

GROUP 1

1. Write a note on DNA as the genetic material. Describe the double helix and fundamental principles of the genetic code.

2. How can we determine the mode of inheritance of a dominant, rare, fully penetrant diallelic trait ?

3. Compare and contrast the characteristics of sex-linked and sex-limited traits. Illustrate your answer with suitable examples.

4. Write short notes on any three of the following :
   (i) Down's syndrome,
   (ii) Holandric inheritance,
   (iii) Co-dominant alleles,
   (iv) Penetrance.

GROUP 2

5. (i) Write the genotypes against each phenotype of the A_1A_2BO blood group system.
   (ii) What is the principle of blood grouping ?
   (iii) What kinds and proportions of blood group phenotypes would you expect among children of the following parents :
   (a) both parents of type M,
   (b) both parents of type N,
   (c) both parents of type MN, and
   (d) one parent of type M and the other of type N.

p.t.o.

6.(i) What is Rhesus blood group ?

(ii) Describe what you know about the haemolytic disease of the new born.

7.(i) What are the common abnormal haemoglobins present in Indian populations ? Briefly discuss their distribution in the Indian subcontinent.

(ii) By what technique most of the abnormal haemoglobins can be detected ? Give a brief description of the technique.

(iii) Which are the most frequent abnormal haemoglobins present in Africa and Southeast Asia ?

(iv) In case of haemophilia, if a carrier woman marries a normal man, what types of offspring would you expect and in what proportions ?

8. Write short notes on any <u>three</u> of the following :

(i) Polymorphism,
(ii) Genetic drift,
(iii) Backcross,
(iv) Hybrid vigour.

1985-86 | 332

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1985-86

STATISTICAL INFERENCE - THEORY

SEMESTRAL-I EXAMINATION

Date : 25.11.85. Maximum Marks : 50 Time : 3 Hours

Note : Answer question 1 and any five from the rest.

1. Discuss the following : complete sufficient statistic, minimal sufficient statistic, Fisher information, near MLE, score in method of scoring, Type I and Type II errors, size and level of a test, randomized and non-randomized tests, UMP and UMP unbiased tests, test having Neyman Structure. $(10 \times \frac{1}{2}$

2.(a) If T is minimal sufficient for θ in Ω_1, a subset of Ω, and sufficient for θ in Ω, then show that T is minimal sufficient for θ in Ω.

(b) Hence find a minimal sufficient statistic for θ based on n iid observations from Uniform $(\theta - \frac{1}{2}, \theta + \frac{1}{2})$ distribution. (4+5 = 9)

3.(a) Compute Chapman-Robbins lower bound for the variance of an estimator of θ based on X with $P_\theta[X=0]=1-\theta$, $P_\theta[X=1] = \theta$, $0 < \theta < 1$.

(b) Compute Cramer-Rao lower bound for the variance of an estimator for estimating σ^2 based on n iid observation from $N(0,\sigma^2)$. (5+4 = 9)

4.(a) Derive UMVUE of the variance of UMVUE of θ based on n iid observations from

$$P_\theta[X=x]= a(x)\,\theta^x / \left(\sum_{i=0}^{\infty} a(i)\theta^i\right),\ \theta > 0,\ a(x) > 0,\ x=0,1,2,\ \ldots$$

(b) Find UMVUE of the power function of a size α UMP unbiased test for $H_0: \lambda \leq \mu$ against $H_1: \lambda > \mu$ based on X, which is Poisson (λ), and Y, which is Poisson (μ). (6+3

p.t.o.

5. Based on n iid observations from $N(\mu, \sigma^2)$, find
   (a) an ancillary statistic
   (b) MLE of μ/σ
   (c) UMVUE of μ/σ.

(1+4+4 = 9)

6. (a) State and prove Neyman-Pearson fundamental lemma.
   (b) State generalised Neyman-Pearson lemma. (2+5+2 = 9)

7. Based on n iid observations from $N(\mu, 1)$
   (a) Construct the MP test of size α for $H_0 : \mu = \mu_0$ against $H_1 : \mu = \mu_1 > \mu_0$, μ_0 and μ_1 are given.
   (b) Show that the above test is UMP size α for $H_0 : \mu \leq \mu_0$ against $H_1 : \mu > \mu_0$. (5+4 = 9

8. (a) When do you say that a rv X has MLR in x ?
   (b) Let X be uniform in $(\theta, \theta+1)$. Does X have MLR in x ?
   (c) When X has MLR in x, show that the test ϕ given below non-decreasing power function.

$$\phi(x) = \begin{cases} 1 & \text{if } x > x_0,\ x_0 \text{ given} \\ \gamma & \text{if } x = x_0 \\ 0 & \text{if } x < x_0 \end{cases}$$

(1+1+7 = 9)

9. Suppose that the distribution of X is in one-parameter exponential family with parameter θ. Then, given any test ϕ, θ_1 and θ_2 with $\theta_1 < \theta_2$, show that there exists a two-sided test ϕ^* for which

$$E_{\theta_i}\phi(X) = E_{\theta_i}\phi^*(X), \ i = 1,2.$$

10. Based on n iid observations from $N(\mu, \sigma^2)$, deduce an explicit form for a UMP unbiased test of size α for $H_0 : \mu \leq \mu_0$ against $H_1 : \mu > \mu_0$, μ_0 given. (9)

[1985-86 [332(P)]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

STATISTICAL INFERENCE - PRACTICAL
SEMESTRAL-I EXAMINATION

Date : 25.11.85.   Maximum Marks : 50   Time: 3 Hours

Note : State your working formulae clearly and show computational steps. There is no need to derive working formulae and no credit will be given for derivation. Accuracy of computation is important. Final answer should be clearly displayed.

Answer any <u>three</u> of the first five questions.

1.(a) Probability function of a rv X is given below.

$\theta \backslash x$	-5	-4	-3	-2	-1	0	1	2	3	4	
θ_0	.40	.01	.14	.17	.11	.01	.06	.02	.05	.02	
θ_1	.28	.01	.05	.39	.01	.02	.03	.03	.01	.12	.05

Obtain the most powerful test of size $\alpha = 0.2$ for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ and compute the probability of type II error of the test.

(b) A telephone sub-office has 20 employees. On a given day only 8 employees were present in the sub-office, of which 5 employees were female.

Construct a UMP test of size $\alpha = 0.1$ to test the null hypothesis that the sub-office has at most 10 female employ

(7+1+7 =

2. Let X and Y be independent Poisson random variables with λ and μ respectively. Compute the power, at $\lambda = 1$ and $\mu =$ of a UMP unbiased size $\alpha = .2$ test for $H_0 : \mu \leq \lambda$ against $H_1 : \mu > \lambda$.

3. Based on an observation from the density

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1}, & 0 < x < 1,\ \theta > 0 \\ 0, & \text{elsewhere} \end{cases}$$

construct a UMP unbiased size $\alpha = 0.05$ test for $H_0: \theta = 1$ against $H_1: \theta \neq 1$. (15)

4. The breaking strength of a particular type of string is $N(\mu, \sigma^2 = 4.41)$. Breaking strengths (in lbs.) of 15 randomly cut pieces of the string are :

15.9, 17.9, 21.0, 20.5, 16.8, 19.1, 17.5, 18.4,
16.9, 18.1, 19.7, 17.6, 21.8, 19.3, and 21.9.

Construct a UMP unbiased test of size $\alpha = 0.05$ for testi $H_0: \mu = 18$ lbs. against $H_1: \mu \neq 18$ lbs. and construct the associated confidence interval for μ. (12+3)

5.(a) Based on 25 observations from $N(\mu, \sigma^2)$, following we calculated :

$$\sum_{i=1}^{25} X_i = 140, \quad \sum_{i=1}^{25} X_i^2 = 900.$$

Find MLE of μ and σ^2 when (i) $-\infty < \mu < \infty$, $\sigma > 0$, and (ii) $-\infty < \mu < \infty$, $0 < \sigma \leq 4$.

(b) Following 9 independent observations are from Poisson with expectation λ :

0, 0, 2, 1, 0, 2, 1, 0, 1.

Compute UMVUE of $e^{-\lambda}$. (1+3+11=15)

6. Practical records. (5)

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

DIFFERENCE AND DIFFERENTIAL EQUATIONS
SEMESTRAL-I EXAMINATION

Date: 22.11.85. Maximum Marks : 100 Time: 3 Hours.

Note : Answer as many questions as you can. The total marks of all questions is 110 but you can score a maximum of 100.

1.(a) State a theorem which guarantees the existence and uniqueness of a real-analytic solution to

$$\frac{dy}{dx} = x+e^y, \quad y=c \text{ when } x=0.$$

(You must show how your theorem applies to this problem). [5]

(b) By actually substituting $y = c + c_1x + c_2x^2 + \ldots$ in the differential equation above find the values of c_1, c_2 and c_3 in terms of c. [10]

2. Let $F(x,y,\lambda)$ be a real-analytic function of all the three variables in the region $|x| \le a$, $|y| \le b$, $|\lambda| \le c$.

Consider the initial value problem

$$\frac{dy}{dx} = F(x,y,\lambda), \quad y = 0 \text{ when } x = 0$$

for each value of the parameter λ in $|\lambda| \le c$. Prove carefully, using the majorant principle that the formal power-series solution

$$y = \sum_{p,q=0}^{\infty} a_{pq}\, x^p \lambda^q$$

must converge in some rectangle around origin in (x,λ)-plane and will give the solution to the initial value problem for each λ near zero.
(You must find a 'dominating' system which you can explicitly solve, and then apply the majorant method.) [15]

3. Find the Euler-Lagrange type of differential equation that is satisfied by an admissible function $y = y(x)$ which extremises the functional

$$v[y(x)] = \int_{x_0}^{x_1} F(x,y,y',y'', y\ )\, dx.$$

p.t.o.

Q.3. contd....

Here F is a C^∞ function of its four arguments. Admissible functions are C^4 functions on $[x_0, x_1]$ with assigned boundary values as shown :

$$y(x_0) = y_0 \quad , \quad y(x_1) = y_1$$
$$y'(x_0) = y_0' \quad , \quad y'(x_1) = y_1' .$$

What is the order of the differential equation you find for $y(x)$?

(You need not prove the fundamental lemma of the calculus of variations but you must carefully state and explain its application to your proof.) [15]

4.(a) A particle moves from (x_0, y_0) to (x_1, y_1) along a C^2 path so that its speed $\frac{ds}{dt}$ is always proportional to the ordinate y. Find the paths which extremise the time taken to execute the motion. [10]

(b) Find the two-parameter families of extremals for the functionals

(i) $v[y(x)] = \int_{x_0}^{x_1} \frac{1 + y^2}{(y')^2} dx$

(ii) $v[x(t)] = \int_{t_0}^{t_1} \frac{1}{t^3} \left(\frac{dx}{dt} \right)^2 dt$ [6+5]

5.(a) Using Laplace transforms solve

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 4y = xe^{-2x}$$

given y = -1 and y′= 0 at x = 0. [3]

(b) Find the general solutions of the difference equations

(i) $x_{k+2} + 3x_{k+1} + 5x_k = 7^k$

(ii) $y_{k+2} + 2y_{k+1} - 3y_k = c$ (a constant). [5+6]

6.(a) A raindrop has its radius increasing at a uniform rate by accumulation of moisture. If it is given an initial horizontal velocity v_0 prove that (under gravity) it describes a hyperbola with a vertical asymptote. Assume its initial radius was r_0. Find the position of the centre of the raindrop as a function of time. (Neglect air resistance).

(b) Solve by variation of parameters

$$\frac{d^2x}{dt^2} + x = \frac{1}{\sin t} .$$ [8]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

STOCHASTIC PROCESSES-2
SEMESTRAL-I EXAMINATION

Date : 18.11.85. Maximum Marks : 100 Time: 3 Hours

Note : Maximum marks you can score is 100.

1. Let (X_t) be the standard Brownian motion.

(a) Show that $\frac{X_n}{n} \longrightarrow 0$ almost surely as $n \to \infty$

(b) Fix $0 < s < t$. Write the joint density of X_s, X_t explicitly. (Here explicitly means you have to evaluate explicitly the quadratic form involved).

(c) Define a process Y_t as follows :

$$Y_0 = 0, \quad Y_t = t \cdot X_{1/t} \quad \text{for } t > 0.$$

Show that this process is again Gaussian with mean zero and evaluate the covariance kernel.

[5+5+10]

2. Consider a pure death process on $\{0,1,2, \ldots\}$ with death rates μ_n $n \geq 1$ and $\mu_0 = 0$.

(a) Write down the forward equations.

(b) Solve for $p_{ij}(t)$ in terms of $p_{i,\, j+1}(t)$.

(c) Find $p_{ii}(t)$.

(d) Find $p_{i,\, i-1}(t)$.

(e) If $\mu_n = n\mu$ show that

$$p_{ij}(t) = \binom{i}{j} (e^{-\mu t})^j (1 - e^{-\mu t})^{i-j} \quad 0 \leq j \leq i.$$

[5+7+5+5+8]

p.t.o.

-

3. Consider a renewal process (N_t) driven by F. As usual assume that F is not concentrated at the origin. Fix $x < t$. Show with usual notation

$$P(X_{N(t)+1} \leq x) = \int_{t-x}^{t} [F(x) - F(t-y)]\, dm(y).$$ [10]

4.(a) Explain what is a delayed renewal process (N_t) with initial distribution G and interarrival distribution F.

(b) Show that N_t is finite almost surely.

(c) Show that $N_t \to \infty$ almost surely.

(d) Show that $\frac{N_t}{t} \longrightarrow \frac{1}{\mu}$ where $\mu = \int_0^\infty x dF(x)$.

As usual F and G are not concentrated at 0. [5+5+5+10]

5. Consider two independent Poisson processes (X_t) and (Y_t) with intensities λ and μ respectively. Let T_5 and T_6 be the times of occurrences of the 5^{th} and 6^{th} events of the (X_t) process. Put $N = Y_{T_6} - Y_{T_5}$. Thus N is the number of events of the Y process during the interval $(T_5, T_6]$. Calculate the distribution of N. [15]

6. State Blackwell's renewal theorem and Feller's renewal theorem. You must explain the terms involved in the statements. [10]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

SAMPLE SURVEYS
SEMESTRAL-I EXAMINATION

Date : 20.11.85. Maximum Marks : 100 Time : 4 Hours.

Note : Answer all questions.

1.(a) Define a 'sampling design'. From a population of size N, one unit is drawn with probability of selection proportional to its size measure x. The rest of (n-1) units in the sample are selected from the remaining (N-1) units of the population by simple random sampling without replacement. Write down p(s), the probability of obtaining the sample for this design. Hence show that the ratio estimator $\hat{R} = \sum_1^n y_i / \sum_1^n x_i$ is unbiased for $R = \sum_1^N Y_i / \sum_1^N X_i$, the Ratio of population totals of the two variates y and x, for this design. [2+3+4] = [9]

(b) For the above sampling design obtain the probability of inclusion of the ith unit. [3]

2.(a) For stratified simple random sampling (without replacement) to estimate the population mean, write down the optimum allocation of a fixed total sample size n to the strata, explaining clearly your notations. How does one implement this allocation in practice ? [2+3] = [5]

(b) In the above situation, suppose that the actual (a) allocation in practice turns out to be n_i^a for the i th stratum while the optimum (O) allocation is n_i^o. Obtain an expression for the relative loss of efficiency measured by

$$\frac{Var_a(\hat{\bar{Y}}_{st}) - Var_O(\hat{\bar{Y}}_{st})}{Var_O(\hat{\bar{Y}}_{st})}$$

where the symbols have the usual meaning and the stratum sizes are assumed large. Further, derive a quick upper

p.t.o.

2.(b) contd....

bound to the above expression in terms of θ, the relative deviation of sample allocations given by

$$\theta = \left| \frac{n_i^o - n_i^a}{n_i^a} \right|$$

[7+3] =[10]

(c) Indicate the situations when you use the 'combined and separate ratio estimators' in stratified random sampling, justifying your answer with necessary formulae and assumptions. [5]

3.(a) Define the term 'Intra class correlation coefficient'. A population consists of 12 clusters each of size 5. Find bounds for the intra cluster correlation coefficient among the elements of the cluster. [3+3]=[6]

(b) A population consists of N clusters of varying sizes M_i, $i = 1,2, .., N$. Suppose that n clusters are selected at random and without replacement. Let Y_{ij} be the value taken by the study variate y on the jth unit of the ith cluster, $j = 1,2,..., M_i$; $i = 1,2, .., N$. Let $\bar{Y}_i = \sum_{j=1}^{M_i} Y_{ij} / M_i$ be the ith cluster mean. Write down the conventional unbiased estimator for the mean $\bar{Y} = \sum_{i=1}^{N} M_i \bar{Y}_i / \sum_{i=1}^{N} M_i$. Suggest a method of estimating its sampling error. Also comment on any other suitable estimator in this case. [3+5+3] =[11]

4. A sample survey was conducted to estimate the total yield of paddy in a district. A stratified two-stage sampling design was adopted with villages as first stage units and plots in them as second stage units. From each stratum 4 villages were selected with probability proportional to area and with replacement and 4 plots were selected from each selected village with equal probability and without replacement. The data on yield for the sample plots together with information on selection probabilities are given in the table below :

contd....

Q.No.4 contd....

Stratum	Sample village	Inverse of probability of selection	Total no. of plots.	Yield of sample plots 1	2	3	4
I	1	440.21	28	104	182	148	87
	2	660.43	14	108	64	132	156
	3	31.50	240	100	115	50	172
	4	113.38	76	346	350	157	119
II	1	21.00	256	124	111	135	216
	2	16.80	288	123	177	106	138
	3	24.76	222	264	78	144	55
	4	49.99	69	300	114	68	111
III	1	67.68	169	110	281	120	114
	2	339.14	42	80	61	118	124
	3	100.00	134	121	212	174	106
	4	68.07	161	243	115	314	129

Using the above data

(i) Obtain an unbiased estimate of the total yield of paddy in the district.

(ii) Obtain an unbiased estimate of the variance of the above estimate.

(iii) What are the possible sources of non-sampling errors in the above study ? [10+15+8]=[33]

5. An experienced teacher makes a guess of the scores x_i in an examination for each of the 200 students of his class based on past performance. He obtains an average score of $\bar{X} = 58$. For a simple random sample (without replacement) of 10 students the following results are obtained after the examination :

	Student 1	2	3	4	5	6	7	8	9	10
Actual score y_i	61	42	50	58	67	45	39	57	71	[illegible]
Guessed score x_i	59	47	52	60	67	48	44	58	76	[illegible]

(i) Find the regression estimate $\hat{\bar{Y}}$ of the average score $\bar{Y}$ for the class and estimate its sampling error. [5+3]=[illegible]

(ii) Comment on the teacher's ability of predicting the scores of students. If, as an alternative, an estimate $\hat{\bar{Y}}' = \bar{X} + (\bar{y} - \bar{x})$ is used, obtain the gain in precision, if any of $\hat{\bar{Y}}'$ over $\hat{\bar{Y}}$. [1+6]=[7]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1985-86

STATISTICAL INFERENCE
SUPPLEMENTARY PERIODICAL EXAMINATION

Date : 18.10.85. Maximum Marks : 100 Time: 3 Hours

Note : Answer all questions clearly justifying your answers.

1. Parameter space consists of 3 points θ_1, θ_2, and θ_3. Following table gives the probability function $f(x,\theta)$ of a random variable X.

θ \ x	x_1	x_2	x_3	x_4
θ_1	0.2	0.3	0.1	0.4
θ_2	0.5	0.1	0.2	0.2
θ_3	0.3	0.0	0.4	0.3

Find MLE of θ. [10]

2.(a) Obtain Fisher Information (matrix), in n iid observations from $N(\mu, \sigma^2)$, about μ and σ.

(b) Hence show that the sample mean based on n iid observations from $N(\mu,1)$ is UMVUE of μ. [7+5=12]

3. Let, for $0 \leq p \leq 1$,

$$P[X = -1] = p,$$
$$P[X = x] = (1-p)^2 p^x, \quad x = 0,1,2, \ldots..$$

(a) Find the class of unbiased estimators of zero.

(b) Hence show that the statistic T given by

$$T(0) = 1,$$
$$T(x) = 0, \quad x \neq 0$$

is UMVUE of $(1 - p)^2$.

[8+5=13]

4. (a) State Factorization theorem.

(b) Show that an MLE is always a function of a sufficient statistic.

[3+2 =5]

p.t.o.

5.(a) State Rao-Blackwell theorem.

(b) Starting with the estimator $2X_1$, find UMVUE of θ based on a sample $X_1, X_2, \ldots, X_n$ from $U(0, \theta)$.

[3+7 = 10]

6. Observed cell frequencies and cell probabilities are given below for a coupling intercross, where θ is the recombination fraction of two linked genes.

Class	Frequency	Probability
1	125	$\frac{1}{4}(3 - 2\theta + \theta^2)$
2	18	$\frac{1}{4}(2\theta - \theta^2)$
3	20	$\frac{1}{4}(2\theta - \theta^2)$
4	34	$\frac{1}{4}(1 - 2\theta + \theta^2)$

Find MLE of θ.

[45]

7. Practical records. [5]

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year: 1985-86

ELECTIVE 4 : PHYSICAL AND EARTH SCIENCES

PERIODICAL EXAMINATION

Date : 4.9.85.   Maximum Marks : 100   Time: 3 hours

Note : Attempt Question No.1 and any <u>five</u> from the rest.

1. Fill up the blanks (any 13). Only write down one of the four choices for each blank.

(i) The Tidal Hypothesis was proposed by ____ ____ (Moulton-Chamberlin/Kant-Laplace/Jeans-Jeffreys/Ringwood-Alfven).

(ii) Quartz, the most abundant mineral available, has a chemical composition _______(Fe_3O_4/SiO_2/$MgSiO_3$/$CaSO_4$).

(iii) The overall density of the earth is _____(4.5/5.5/6.5/7.5).

(iv) When a sea invades the land, it is called ________. (regression/transgression/onlap/offlap).

(v) Fossils obtained over a long distance of an exposure of a thick Sedimentary rock formed from the meandering activity of river do not necessarily mean that they are of the ________ (same/different/preserved/graded) age.

(vi) The _________ (eyes/muscles/blood-cells/teeth) of a vertebrate animal are most likely to be preserved.

(vii) The ________ (colour/thickness/ripple mark/chemical composition) helps to determine the top and the bottom surfaces of a bed.

(viii) The Lower Jurassic (190 my) dinosaur bones of the Indian Statistical Institute have been obtained from _______. (mixed/continental/turbidite/marine) rocks of the Godavari Valley.

(ix) A rock composed of 40% by volume of pebbles is called ________ (mudstone/phyllite/marble/conglomerate).

(x) Petroleum is a/an _______ (rocky/mineral/organic/inorganic) substance.

(xi) Sedimentary structures include _______ (grain-size/sorting/facies/stromatolite).

(xii) A bedding in which grain-size decreases upward is called a _______ (laminated/cross- /trough/graded) bed.

(xiii) A horizontal sequence of beds lying over a tilted sequence of beds has a/an _______ (unconformable/tectonic/sedimentary/metamorphic) contact.

p.t.o.

1.(xiv) In a fossiliferous horizontal sequence of beds conformably lying over another fossiliferous horizontal sequence of beds, the evolutionary sequence shown by the organisms suggests a big time gap between the two sequences; hence, the two sequences are said to have a/an ______ (tectonic/laminated/tilted/unconformable) relationship.

(xv) ______ (Mud/Feldspar/Basalt/Granite) is a coarse-grained light-coloured plutonic igneous rock.

2. Who is said to be the pioneer in suggesting that the solar system had a cold beginning ? Who modified his hypothesis and how did the modifier explain the origin of the solar system ? [2+2+12]

3. Can a mineral be a crystal as well as a crystalline substance ?
How does a mineral form ? Name three rock-forming minerals along with their respective chemical compositions. Give an example of an amorphous mineral and state its usefulness. [$6+4+4\frac{1}{2}+1\frac{1}{2}$]

OR

What is understood by the Moh's Scale of Hardness ? Describe how you would proceed to determine the hardness of Magnetite (Fe_2O_3) which has a hardness ranging from 5.5 to 6.5. [6+10]

4. What are the uses of the P- and S- wave study in an earthquake ? What is the Low Velocity Zone ? Does it occur above the Mohorovicic discontinuity ?

Describe in short the physical characteristics of the earth's core. [6+4+2+4]

5. In a deep mine, it may be observed that with every 30 meter descent, there is an increase in temperature through 1°C. What is then the temperature of the earth's core ?
What is sial and sima ? What is the reason for the higher heat flow in the sial ?
Which hypothesis, hot earth or cold earth, supports the heat flow of the earth's interior ? Why ? [5+6+5]

6. What is the Geological Time Scale ? What is meant by 'Precambrian' time ? What is the approximate age of India's Gondwana coals ? When did the first birds appear in the earth and what is their name ? [6+6 + 2+ 2]

contd....

7. What is the importance of free Oxygen in the atmosphere ? When did it first form ? What could be the reason for its origin ? [2+4+10]

OR

"In understanding the evolution of hydrosphere, salinity is taken as a factor". - Elucidate.

What is the logic given when scientists say that a considerable portion of all the water in oceans has appeared since late Mesozoic (144 my - 66 my) ? [8+8]

8. What is a 'basin' and what kind of sedimentary processes go on in a basin ?
What is meant by "sorting" of sediments ? In what way the fabric of a sedimentary deposit help in understanding the sedimentary processes ? [4+4+4+4]

OR

What is a cross-bedding and how does it form in a fluviatile environment ? How many kinds of cross-beddings are there ? In what ways cross-bedding study is useful ? [2+6+4+4]

9. What are the different kinds of continental environment ? Describe a kind of river environment and its sedimentary facies.

What are the criteria that you would be looking for in a sedimentary gravel deposit to prove that it is of glacial origin ? What is the name of the deposit ? [6+8+2]

10. Describe how the hard parts of an organism can be altered while it is undergoing fossilization. Give an example to show how fossils can be useful. [10+6]

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : [illegible]

ELECTIVE-4 : BIOLOGICAL SCIENCES
PERIODICAL EXAMINATION

Date : 4.9.85. Maximum Marks : 100 Time: 3 Hours.

Note : Attempt any five questions. All questions carry equal marks.

1. Write a short history of Genetics from Mendel to DNA double helix.

2. Write what you know about transcription and translation, genetic code, introns and exons.

3. Describe Mendel's laws of inheritance.

4. List the criteria of inheritance due to a single, completely dominant, rare, autosomal gene. Illustrate your answer with the help of a pedigree.

5. Distinguish between sex - linked, sex limited and sex influenced modes of inheritance.

6. Write short notes on any <u>three</u> of the following :
   1. Down Syndrome.
   2. Penetrance.
   3. Holandric inheritance.
   4. Haemophilia.
   5. Co - dominant inheritance.

7. Compare and contrast between multiple allelic and polygenic inheritance.

8. Write short notes on any <u>three</u> of the following :
   1. Carrier
   2. Homozygous
   3. Turner Syndrome
   4. Autosomal recessive inheritance.

1985-86 | 331

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1985-86

STATISTICAL INFERENCE - THEORY

· PERIODICAL EXAMINATION

Date : 2.9.85. Maximum Marks : 50 Time: 3 Hours.

Note : Answer any five questions.

1.(a) Let ξ_p (uniquely defined) be the p-th fractile of a distribution and $Z_{p,n}$ be the sample p-th fractile based on n iid observations from the distribution, $0 < p < 1$. Show that $Z_{p,n}$ converges to ξ_p in probability as $n \to \infty$.

(b) Let $X_1, X_2, \ldots, X_n$ be iid uniform distribution in $[0,\theta]$, $\theta > 0$. Find the asymptotic distribution of $n(\theta - T)/\theta$, as $n \to \infty$, where T is the maximum of $X_1, X_2, \ldots, X_n$. [5+5 = 10]

2. Based on n iid observations from $N(\mu, \sigma^2)$, find

(a) MLE of μ and σ^2,

(b) a statistic (may be vector valued) which is sufficient for the family of distributions, and

(c) find UMVUE of μ and σ^2. [6+2+2=10]

3. Let $f(x; \theta)$ be the probability (density) function of X, $T \equiv T(X)$ be a statistic, and $I(\theta; Y)$ be the Fisher Information about θ in Y. Show that, for all θ,

(a) under some regularity conditions (state them)

$$I(\theta; X) = - E_\theta \left[\frac{\partial^2 \ln f(X; \theta)}{\partial \theta^2} \right],$$

(b) $I(\theta; X) \geq I(\theta; T)$, and

(c) $I(\theta; X) = I(\theta; T)$ iff T is sufficient.

[4+3+3=10]

4.(a) State and derive Chapman Robbins lower bound for variance of an estimator.

(b) Compute the lower bound for estimating θ based on X such that $P_\theta[X = 1] = \theta$, $P_\theta[X = 0] = 1 - \theta$, $0 < \theta < 1$.

[5+5 =10]

p.t.o.

5.(a) State and prove Cramer-Rao inequality.

(b) Hence show that the sample mean of n iid observations from $N(\mu,1)$ is UMVUE of μ. [6+4 = 10]

6.(a) State and prove Rao-Blackwell theorem.

(b) Based on n iid observations from Poisson distribution with parameter λ, find UMVUE of $e^{-\lambda}$ starting with T such that

$$T = \begin{cases} 1 & \text{if } X_1 = 0 \\ 0 & \text{else.} \end{cases}$$

[5+5=10]

7.(a) If X_1 and X_2 are iid, show that

$$V_0 [X_1^2 + X_1X_2] \geq V_0 [\tfrac{1}{2} (X_1 + X_2)^2].$$

(b) $X_1, X_2, \ldots, X_n$ are iid Bernoulli with probability of success p. Find UMVUE of $p(1-p)$. [5+5=10]

1985-86 | 331

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year : 1985-86

STATISTICAL INFERENCE – PRACTICAL
PERIODICAL EXAMINATION

Date : 2.9.85. Maximum Marks : 50 Time: 3 Hours.

Note : State your working formulae clearly and show computational steps. There is no need to derive the working formulae and no credit will be given for derivation. Accuracy of computation is important and the final answer should be clearly displayed.

Answer all questions.

1. Following is a sample of size 10 from $N(\mu, \sigma^2)$:

   14, 12, 7, 11, 27, 9, 13, 12, 13, and 8.

   Obtain (i) MLE, and (ii) UMVUE of μ and σ^2 when

   (a) $-\infty < \mu < \infty$, $\sigma > 0$, and

   (b) $13 < \mu < \infty$, $\sigma > 0$. [5+7 = 12]

2. Following is a sample of size 10 from uniform $[0,\theta]$, $\theta > 0$:

   11.4, 11.2, 10.7, 12.7, 11.1, 10.9, 11.3, 11.2, 11.3, and 10.8.

   Estimate θ by the method of

   (a) moments,

   (b) maximum likelihood, and

   (c) minimum variance criterion. [3+2+3 = 8]

3. Observed cell frequencies and cell probabilities are given below for a coupling intercross, where θ is the recombination fraction of two linked genes.

Class	Observed frequency	Probability
1	125	$\frac{1}{4}(3 - 2\theta + \theta^2)$
2	18	$\frac{1}{4}(2\theta - \theta^2)$
3	20	$\frac{1}{4}(2\theta - \theta^2)$
4	34	$\frac{1}{4}(1 - 2\theta + \theta^2)$

   Find MLE of θ. [10]

4. Practical records. [5]

[1985-86]

INDIAN STATISTICAL INSTITUTE
B.Stat.,(Hons.) III Year : 1985-86

DIFFERENCE AND DIFFERENTIAL EQUATIONS
PERIODICAL EXAMINATION

Date : 30.8.85. Maximum Marks : 100 Time: 3 Hours.

Note : Answer as many questions as you can. Complete and clear answers will count more than partial answers to many questions. Workings and proofs must be shown. The total marks of all questions is 110: You can get a maximum of 100.

1.(a) Find the second order differential equation whose general solution is

$$y = A \sin^{-1} x + B$$

A and B being arbitrary constants. (5)

(b) Obtain a differential equation representing all tangent lines to the parabola $y^2 = 4x$. What are all the solutions of your equation? (10)

2.(a) Find a second order partial differential equation

(involving at most $\frac{\partial^2 z}{\partial x^2}$, $\frac{\partial^2 z}{\partial y^2}$, $\frac{\partial^2 z}{\partial x \partial y}$, $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$) which

is satisfied by any $z = z(x,y)$ of the form

$$z = xf(y + 2x) + g(y + 2x),$$

where f and g are two arbitrary smooth functions of one variable. (8)

(b) Prove that if the family of integral curves of the linear differential equation

$$\frac{dy}{dx} + P(x)\cdot y = Q(x)$$

is cut by the vertical line $x = c$, then the tangents to the solution curves at the points of intersection all pass through a fixed point. Find the coordinates of this point of concurrence.(in terms of P,Q and c). (10)

3.(a) Consider $\frac{dy}{dx} = f(x,y)$; f continuous on $a \leq x \leq b$ and $-\infty < y < \infty$, and f is Lipschitz in the y-variable on this strip. Let (x_0, y_0) be a point of the strip. Define

$$y_0(x) = y_0 \text{ on } [a,b],$$

and $y_n(x) = y_0 + \int_{x_0}^{x} f(t,y_{n-1}(t))\,dt$, $n = 1,2,\ldots$ on [a,b].

p.t.o.

3.(a) contd....

Prove that $y_n(x)$ converges uniformly on all of $[a,b]$ to a continuous limit function $y(x)$. (You are not asked to show that the limit function solves the initial value problem). (10)

(b) If $f(x,y) = xy^2$ then does the hypothesis of part (a) apply ? Prove yes or no. State carefully a theorem which will guarantee some existence and uniqueness statement for a solution of $y' = xy^2$ passing through any given (x_0, y_0) in $\mathbb{R}^2$. (8)

4. Solve the equations :

(i) $y^2 \frac{dy}{dx} - y^3 \tan x = \sin x$

(ii) If $(x^2 + y^2)^\alpha$ is an integrating factor for

$$(x+y)\,dx - (x - y)\,dy = 0$$

find α and solve the equation. (6)

(iii) Using Laplace transforms solve

$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = x$$

given $y(0) = 2$ and $y'(0) = -1$. (8)

(iv) Using the method of variation of parameters solve

$$y'' + 2y' + y = e^{-x} \log x .$$

(You must find the complementary function and then use variation to obtain a particular integral). (8)

5.(a) Let $y(x)$ be some solution of the equation

$$x\frac{d^2y}{dx^2} + \frac{dy}{dx} + xy = 0.$$

You are told that the Laplace transform of $y(x)$ exists. Find the form of this Laplace transform. (10)

(b) Solve :

(i) $(y^4 - 2x^3y)\,dx + (x^4 - 2y^3x)\,dy = 0.$ (6)

(ii) $\frac{dy}{dx} = f(ax + by + c)$. (5)

(c) Give an example of a continuous function $f(x,y)$ on a rectangle in $\mathbb{R}^2$ such that $\frac{dy}{dx} = f(x,y)$ has two distinct solutions passing through some interior point of your rectangle. (4)

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1985-86

STOCHASTIC PROCESSES - 2

PERIODICAL EXAMINATION

Date : 28.8.85. Maximum Marks : 100 Time: 3 Hours

Note : Each question carries 20 marks. Maximum you can score is 100.

1. Consider a Poisson process (N_t) with intensity λ. Fix $0 < s < t$ and an integer $n \geq 1$. Show that the conditional distribution of N_s given $N_t = n$ is Binomial $(n, \frac{s}{t})$. What is the joint distribution of (N_1, N_2, N_3) ?

2. The number of accidents in Calcutta upto time t be denoted by C_t and the corresponding number in Bombay be B_t. We assume that (C_t) is $PP(\lambda_1)$ and (B_t) is $PP(\lambda_2)$ and that the two processes are independent.

   (a) Let T be the time of the first accident in Bombay. What is the distribution of T ?

   (b) Let Z be the number of accidents in Calcutta upto (and including) the time of the first accident in Bombay. In other words : $Z = C_T$. Show that

$$P\ (Z=k) = \frac{\lambda_2}{\lambda_1+\lambda_2} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \quad k = 0,1,2, \ldots.$$

3. Let (N_t) be $PP(\lambda)$. Let X_0 be a random variable independent of the process and $P(X_0 = 1) = P(X_0 = -1) = \frac{1}{2}$. Put $X_t = X_0.(-1)^{N_t}$.

   (a) Show that for every t, the distribution of X_t is same as that of X_0.

   (b) Fix $s < t$, show that the joint distribution of (X_s, X_t) is same as the joint distribution of

$$(X_{s+2}, X_{t+2}).$$

p.t.o.

4. Consider a Poisson process (N_t) with intensity 1. An event occuring at time u has chance e^{-u} of being recorded - independent of other events. Let R_t be the number of recorded events upto (and including) time t. Show that (R_t) is a nonhomogeneous Poisson Process with intensity function e^{-t}. That is, show

(a) $R_0 = 0$

(b) $R_{t+s} - R_t \sim \text{Poisson}\left(\int_t^{t+s} e^{-u}\, du\right)$

(c) (R_t) process has independent increments.

5. Consider a renewal process (N_t) driven by F. As usual F is assumed to be not concentrated at 0. Fix t. Show that for $x<t$

$$P(X_{N(t)+1} \leq x) = \int_{t-x}^{t} [F(x) - F(t-y)]\, dm(y)$$

Recall $m = \sum_1^{\infty} F_n$, and F_n is the c.d.f of $X_1 + \ldots + X_n$ where $X_1, \ldots X_n$ are i.i.d with c.d.f F.

6. With the usual notation of renewal theory, state which of the following statements are true and which are false. Give justifications.

(a) $N(t) < n$ if and only if $S_n > t$.

(b) $N(t) \leq n$ if and only if $S_n \geq t$.

(c) $N(t) > n$ if and only if $S_n < t$.

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1985-86

SAMPLE SURVEYS
PERIODICAL EXAMINATION

Date: 26.9.85. Maximum Marks : 100 Time : $3\frac{1}{2}$ Hours.

1.(a) Define the terms 'Sampling Design' and 'Sampling Scheme'. What do you understand by the 'inclusion probability of a unit, π_i' and 'joint inclusion probability of a pair of units, π_{ij}'. Calculate π_i and π_{ij} for a Simple Random Sampling With Replacement design of n draws.
(3+3+2+2+2+3)=(15)

(b) Let the population size be 3 and the sample size be 2 and let $s_1 = \{U_1, U_2\}$, $s_2 = \{U_1, U_3\}$ and $s_3 = \{U_2, U_3\}$. Under the Simple Random Sampling design let $p(s_i) = 1/3$ for $i = 1,2,3$. Define the estimator t by

$$t = \begin{cases} t_1 = (y_1+y_2)/2 \text{ if } s_1 \text{ occurs} \\ t_2 = (y_1/2) + (2y_3/3) \text{ if } s_2 \text{ occurs} \\ t_3 = (y_2/2) + (y_3/3) \text{ if } s_3 \text{ occurs.} \end{cases}$$

Show that t is unbiased for $\bar{Y}$ and that there exist populations (Y_1, Y_2, Y_3) for which $V(t) \not> V(\bar{y})$, where $\bar{y}$ is the conventional sample mean. What does this example show ?
(3+5+2)=(10)

(c) Of 105 office-going commuters sampled using a SRSWOR design from a population of 1241 commuters, 11 have expressed that they did not prefer a 5-day week. Estimate the proportion of commuters in the population that do not prefer a 5-day week and also obtain an approximate 95% confidence interval for the proportion. (3+7) = (10)

2.(a) Explain how you would draw a sample of size n from a population of N units with probability of selection P_i, proportional to a given size measure X_i ($i = 1,2,.., N$) with replacement using Lahiri's method of selection. Show that this procedure indeed gives the selection probability for the ith unit equal to $P_i = X_i / \sum_{i=1}^{N} X_i$. (5+5)=(10)

p.t.o.

2.(b) A sample of 6 factories is drawn from a population of 74 factories with probability of selection of a factory proportional to the size x (no. of workers) with replacement and the number of absentees is observed :

Sampled factory	x size (in '000s)	y no. of absentees
1	21	105
2	101	524
3	14	73
4	6	31
5	41	200
6	12	64

It is also known that the total size of all the 74 factories is $X = \sum_{i=1}^{74} X_i = 2{,}949{,}000$.

(i) Estimate the <u>average</u> number of absentees in the factory.

(ii) Calculate an unbiased estimate of the sampling error of your estimate in (i) above.

(iii) Estimate the gain in efficiency of using a PPSWR design compared to a simple random sampling with replacement design. (6+12+12)=(30)

3.(a) What are the similarities and differences between 'Linear Systematic Sampling' and 'Circular Systematic Sampling'? (3)

(b) When the values of the y - characteristic are known to be of the form $Y_i = \alpha + \beta i$ and when the population size is a multiple of the sample size, would you prefer Systematic Sampling to Simple Random Sampling ? Give reasons. (6)

(c) On the basis of a single systematic sample, explain how you would obtain an estimate for the variance of the estimated population mean. Comment on this variance estimate. (4+2)=(6)

(d) 3 systematic samples of size 4 each were drawn from a population of size 28 independently. The data is given below :

	y - values
sample 1 :	104, 203, 178, 165
sample 2 :	206, 109, 167, 154
sample 3 :	416, 809, 689, 647

Obtain an unbiased estimate of the population mean and an unbiased estimate of its sampling error. Also comment on the non-sampling error, if any. (2+4+4)=(10)