INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) Part IV: 1980—81
PERIODICAL EXAMINATIONS
Design of Experiments

Date: 22.9.80 Maximum Marks: 100 Time: 3 hours

Note: Question No.5 is compulsory. Of the rest, answer any four. All the questions carry equal marks.

1.a) Formulate precisely the weighing problem involving a chemical balance without bias.

b) Define and characterize an optimum weighing design in this context.

c) Show that an optimum weighing design exists only if the number of weighing operations is an integral multiple of 4.

2. Define a Hadamard matrix (H_N) of order N. Give an example of H_8. Explain the role of such matrices in constructing exactly or asymptotically optimum weighing designs for a chemical balance with or without bias.

3.a) "In a spring balance weighing design, there is always loss of precision whenever the bias component is present; however, the presence of bias also helps in achieving orthogonal estimates of the individual weights in certain cases" - justify the statement.

b) Indicate a method of constructing a spring balance (with bias) weighing design for s^2 objects in exactly (s^2+2s+1) weighing operations so as to provide orthogonal estimation of the individual weights [You may use the fact that the BIBD ($b = s^2+s$, $v = s^2$, $r = s+1$ $k = s$, $\lambda = 1$) exists].

4.a) Show that in a symmetrical BIBD ($b = v$, $r = k$, λ), $|B_i \cap B_j| = \lambda \quad \forall\ 1 \leq i \neq j \leq b$ in usual notations. Conversely, suppose in a BIBD (b, v, r, k, λ), $|B_i \cap B_j| = \lambda$ $\forall\ 1 \leq i \neq j \leq b$. Can you conclude that the BIBD must be symmetrical?

b) For a BIBD (b, v, r, k, λ), work out the expression for $|NN'|$ where $N(v \times b)$ is the incidence matrix. Hence, or otherwise, show that a necessary condition for a symmetrical BIBD to exist is that $(r-\lambda)$ must be a perfect square when v is even.

c) Construct the symmetrical BIBD ($b = v = 21$, $r = k = 5$, $\lambda = 1$).

5. What is meant by 'Dual' of a block design? What are Linked Block (LB) designs? How would you analyze an LB design? Work out expressions for the different Sum of Squares (S.S) in the ANOVA table for an LB design with parameters b^*, v^*, r^* and k^*.

6. Submit practical note-book.

1980-81 : 421

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) Part IV: 1980-81

PERIODICAL EXAMINATIONS

Measure Theory

29.9.80 Maximum Marks. 100 Time: 3 hours

Note: The paper carries 110 marks. Maximum you can score is 100

You may assume the Caratheodory extension Theorem and the relevant formulae; but whenever you use them state them explicitely.

If $\underset{=}{G}$ is a collection of subsets, $\sigma(\underset{=}{G})$ stands for the σ-field generated by $\underset{=}{G}$.

1.a) What is a complete measure space and the completion of a measure space.

b) Let X be the real line, $\underset{=}{A}$ be the σ-field of countable subsets of X and their complements and μ be the set function on $\underset{=}{A}$ defined by $\mu(A) = 0$ if A is countable and $\mu(A) = 1$ if A is co-countable. Show that μ is a measure. Find the completion of $(X, \underset{=}{A}, \mu)$. [4+6]=[10]

2. Explain clearly and explicitely how each distribution function on the real line corresponds to a probability measure on the Borel σ-field of the real line. Show that this correspondence is one-to-one and onto all the probability measures on the Borel σ-field of the real line. [15]

3. Let $\underset{=}{G}_1 = \{(a,b): -\infty < a < b < \infty\}$

$\underset{=}{G}_2 = \{[a,b]: -\infty < a < b < \infty\}$

$\underset{=}{G}_3 = \{(a,b]: -\infty < a < b < \infty$ and a and b are rational$\}$.

Show that $\sigma(\underset{=}{G}_1) = \sigma(\underset{=}{G}_2) = \sigma(\underset{=}{G}_3)$. [10]

4. If $\underset{=}{A}$ is a field of sets and A is a set which is not in $\underset{=}{A}$ show that the field generated by $\underset{=}{A}$ and A is the collection of all sets of the type $(B \cap A) \cup (C \cap A^c)$ where B and C belong to $\underset{=}{A}$. [10]

5. Let μ be a bounded finitely additive measure on a field $\underset{=}{A}$. Let $\underset{=}{C}$ be a compact class contained in $\underset{=}{A}$ such that $\underset{=}{C}$ approximates $\underset{=}{A}$ with respect to μ. Then show that μ is a measure on $\underset{=}{A}$. If $\bar{\mu}$ is a measure on $\sigma(\underset{=}{A})$ such that $\bar{\mu}$ is an extension of μ then show that $\underset{=}{C}_\delta$ approximates $\sigma(\underset{=}{A})$ with respect to $\bar{\mu}$, where $\underset{=}{C}_\delta$ is the collection of all countable intersection of sets from $\underset{=}{C}$. [20]

6. In any bounded measure space if $\{A_n\}_{n \geq 1}$ is a sequence of sets such that $\sum_{n=1}^{\infty} \mu(A_n \Delta A_{n+1}) < \infty$ then show that $\mu(\overline{\lim} A_n) = \mu(\underline{\lim} A_n)$. [20]

p.t.o.

7. Let μ be a bounded measure on the σ-field $\sigma(\underline{\underline{A}})$ generated by a field $\underline{\underline{A}}$. Show that for every $A \in \sigma(\underline{\underline{A}})$ and $\epsilon > 0$ there is a set $B \in \underline{\underline{A}}$ such that $\mu(B \triangle A) < \epsilon$. [15]

8. Let $(X, \underline{\underline{A}}, \mu)$ be a measure space. Let $Y \subset X$ be a subset.

a) Show that $Y \cap \underline{\underline{A}}$ defined by $\{Y \cap A : A \in \underline{\underline{A}}\}$ is a σ-field.

b) Is $\mu^*(Y) = 1$ show that τ given by

$\tau(Y \cap A) = \mu(A)$ is well defined and is a measure.

c) If τ given by $\tau(Y \cap A) = \mu(A)$ is well defined show that $\mu^*(Y) = 1$. [10]

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) Part IV : 1980-81
PERIODICAL EXAMINATIONS
Inference

Date 6.10.80 Maximum Marks: 100 Time: 3 hours

Note: Answer all questions. Marks are indicated in brackets. The class notes are allowed to use during the examination.

1. Let $x_1, \ldots, x_n$ be a random sample from $N(\theta,1)$. We want to estimate θ under the loss function $L(\hat{\theta},\theta) = |\hat{\theta} - \theta|$. If θ has a prior distribution $N(0, \sigma^2)$, find the Bayes estimate for θ. [16]

2. Let $x_1, \ldots, x_n$ be a random sample from $N(\theta,1)$, We wat to estimate θ under the loss function $L(\theta,\theta) = (\theta - \theta)^2$ find a minimax estimate of θ. [15]

3. Let $\Theta = (0,\infty)$, $\mathcal{Q} = [0,\infty)$ and x be a poisson distribution

$$P(x=k) = e^{-\theta} \frac{\theta^k}{k} \qquad k = 0,1,\ldots.$$

we want to estimate θ.

a) If the loss function is $L(\theta,a) = \frac{(\theta - a)^2}{\theta}$, find a minimax rule.

b) If the loss function is $L(\theta,a) = (\theta - a)^2$, does there exist a minimax rule [15]

4. Let x be a random variable with density

$$f_\theta(x) = \theta e^{-\theta x} \qquad x>0,\ \theta>0.$$

We want to estimate θ based on a single observation under the loss function $L(\theta,\theta) = \theta(\theta - \theta)^2$.
Find the Bayes estimate for θ. [19]

5. Two contestants statistician and nature simultaneously put up either one or two fingers. The statistician wins if the sum of the digits showing is odd and the nature wins if the sum of the digits showing is even. The loss table is given by

H \ θ	1	2
1	-2	3
2	3	-4

The statistician is allowed to ask the nature how many fingers he intends to put up and that nature must answer truthfully with probability $\frac{3}{4}$. The statistician therefore observes a random variable x(the answer nature gives) taking values 1 or 2.

$$P(x=1 \mid \theta = 1) = \frac{3}{4}, \quad P(x=2 \mid \theta = 1) = \frac{1}{4}$$
$$P(x=1 \mid \theta = 2) = \frac{1}{4}, \quad P(x=2 \mid \theta = 2) = \frac{3}{4}$$

a) Characterize all non-randomized rules.

b) Compute the risk table for all these non-randomized rules

c) Draw the risk set.

d) Find the minimax rule.

[40]

1980-81: 461/582

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) Part-IV/M.Stat. (Previous): 1980-81

PERIODICAL EXAMINATIONS

SQC and OR

Date: 27.10.80 Maximum Marks: 100 Time: 3 hours

Note: Question 1 is compulsory. Attempt any three other questions from the remaining.

1. Explain the following:

(a) Quality of a product (b) Statistical control (c) Rational subgrouping (d) Screening (e) Type A and Type B, Oc curve (f) Average Run Length of a control chart (g) Control limits versus confidence limits (h) Optimal basic feasible solution to a linear programming problem.

(5 x 8) = [40]

2. A sugar mill sells sugar cubes in packets of net weight equal to 450 gms. Dialy five packets are being sampled from the production line and weighed in the laboratory. The averages and ranges of the net weight in gms. for twenty five days are as follows:

Day	Average	Range	Day	Average	Range	Day	Average	Range
1	437.2	19.0	11	450.5	15.2	21	459.0	18.8
2	446.6	21.2	12	454.2	20.8	22	447.5	19.0
3	449.6	32.8	13	452.9	13.2	23	456.2	18.0
4	459.5	13.0	14	454.4	16.0	24	450.5	24.2
5	446.6	24.7	15	448.6	15.2	25	453.9	22.0
6	445.2	9.4	16	454.2	16.8			
7	456.8	20.2	17	451.8	28.2			
8	454.9	21.2	18	457.9	18.0			
9	458.0	30.2	19	449.0	24.8			
10	455.0	25.8	20	453.5	17.8			

(a) Draw an $\bar{X}$ - R chart and examine whether the process is under control.

(b) Obtain estimate of μ and standard deviation under statistical control and hence compute the proportion of underweight packets being produced by the mill.

(14+6) = [20]

3.(a) The following data give the results of inspection of enamel plates of a standard size for spots.

Plate Number	No. of spots	Plate Number	No. of spots
1	8	7	10
2	7	8	5
3	9	9	20
4	11	10	24
5	12	11	35
6	8	12	10

Contd....2

Contd....

Plate Number	No. of spots	Plate Number	No. of spots
13	10	19	23
14	18	20	11
15	10	21	13
16	10	22	16
17	18	23	14
18	19	24	13

Draw neatly a suitable control chart and examine if the process is under statistical control.

(b) Given lot size N, AOQL = p_L and process average is $\bar{p}$, prove that for a single sampling plan (n,c)

$$p_L = y \left(\frac{1}{n} - \frac{1}{N} \right)$$

where $y = e^{-x} \frac{x^{c+2}}{c!}$ under usual notation.

(12+8) = [20]

4. For the sampling plan N = 3000, n = 100 and c = 1

(a) Draw the O.c. curve and read the value of IQL.

(b) Draw AOQ curve and determine AOQL approximately.

(10+10) = [20]

5.(a) Formulate the following as a linear programming problem:

Consider a company that has one production line upon which it produces a single homogeneous commodity. We suppose that the commodity sells for a fixed unit price, that the cost for regular-time production, overtime production, and storage are known and vary between time periods; that the rate of production per unit time is known; and that an accurate sales forecasts in the form of demand during each of a number of successive time periods is known. It is desired to formulate a production schedule that will meet the sales forecast and minimise the combined costs of production and storage.

- t = number of time periods
- η_i = number of units of finished product to be sold during ith time period
- s_0 = initial stock
- m_i = maximum number of units that can be produced each time period on regular time
- n_i = maximum number of units that can be produced each time period on overtime
- a_i = cost of storage of one unit of product during time period i
- c_i = cost of production of one unit on regular time during time period i
- d_i = cost of production of one unit on overtime during time period i
- x_i = regular - time production during ith time period
- y_i = overtime production during ith time period

Contd...Q.5(a)

For correctness, it will be assumed that each time period is one month in length and that the stock for each month is taken on the last day of the month. This is equivalent to adding the month's production to stock and withdrawing the month's sales from stock at the end of each month.

(b) For the linear programming problem: maximise $Z = CX$ subject to $AX = b$, $X \geq 0$ if a set $k \leq m$ of vectors $a_1, a_2 \cdots a_k$ can be found which are linearly independent and such that $a_1x_1 + a_2x_2 + \cdots + a_kx_k = b$, $x_j \geq 0$, $j = 1, 3 \ldots n$ where A is $m \times n$ matrix and other symbols have the usual meaning, then show that $X = (x_1, x_2, \cdots x_k, 0, \ldots 0)$ is an extreme point of the convex set of feasible solution.

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) Part IV: 1980-81
Complex Analysis

1980-81: 411

PERIODICAL EXAMINATION

Date: 10.11.80 Maximum Marks: 100 Time: 3 hours

Note: The paper carries 120 marks. You may answer as many questions as you like. The maximum you can score is 100. If you use any theorems proved in class, state them clearly and precisely.

1. Let f be a realvalued function on an open set Ω in R^n.

a) Define partial derivatives of f.

b) Let $\Omega = \{ \underline{x} : \|\underline{x}\| < r \}$ where r is a positive real number. Prove that if the partial derivatives are identically zero on Ω, then f is a constant.

c) Is the result true if Ω is a general open set in R^n ? Give a proof if it is true, a counterexample otherwise.

d) Let $\Omega = \{ \underline{x} : \|\underline{x}\| < r \}$ and let f satisfy

$$|f(\underline{x}) - f(\underline{y})| \leq K |\underline{x} - \underline{y}|^{3/2}, \text{ for all } \underline{x}, \underline{y} \in \Omega ,$$

where K is a constant. Prove that f is a constant.

(3+6+4+7) = [20]

2.(a) Let f be a complexvalued function on an open set Ω of the complex plane. Define the complex derivative at a point in Ω. Ignoring the complex structure, f can be regarded as a function on the open subset Ω of R^2 into R^2. Define the total derivative at a point in Ω. What is the relationship between these two derivatives ? Does the existence of one imply the existence of the other ? When do you say that f is holomorphic on Ω ?

(b) Give an example each of i) an open set in $\mathbb{C}$ which is connected but not convex, ii) an open set in $\mathbb{C}$ which is convex but not connected, iii) an open set in $\mathbb{C}$ which is neither convex nor connected.

(c) State the Cauchy - Riemann equations. State sufficient conditions in terms of partial derivatives under which f is holomorphic on Ω.

(d) Let f be a nonconstant holomorphic function on a region Ω. Show that the function g defined by $g(z) = f(\bar{z})$, $z \in \Omega$ is not a holomorphic function on Ω. Show that, however, the function h defined by $h(z) = \overline{f(\bar{z})}$ is a holomorphic function.

(12+6+8+10) = [36]

p.t.o.

3.(a) Define the exponential function on $\mathbb{C}$ and show that, restricted to the real line, this function is nonnegative, strictly increasing and onto $(0, \infty)$.

(b) Find the sum of the series $\sum_{n=1}^{\infty} n\, z^n$ for $|z| < 1$.

(c) Find the radius of convergence of the power series

$$\sum_{n=0}^{\infty} 2^n z^{n!} .$$

(d) Let $z_n \to z$. Prove that

$$\lim_{n \to \infty} \left(1 + \frac{z_n}{n}\right)^n = e^z \quad (z_n, z \in \mathbb{C}).$$

(8+7+7+8) = [30]

4.(a) Define a smooth path and a piecewise smooth path in $\mathbb{C}$. Give an example of a piecewise smooth path which is not smooth. Define the integral of a function over a piecewise smooth path.

(b) Let γ be the path describing a semicircle with centre 0 and radius r, starting at r. Compute $\int_\gamma \bar{z}\, dz$.

(c) Let γ be the positively oriented circle with centre 0 and radius 1. Compute $\int_\gamma \frac{z^2 + e^z}{z}\, dz$.

(d) Let γ_1 and γ_2 be two piecewise smooth paths defined on $[a, b]$ into $\Omega = \{z : |z| < r\}$. Let f be a holomorphic function on Ω. Show that if $\gamma_1(a) = \gamma_2(a)$ and $\gamma_1(b) = \gamma_2(b)$, then

$$\int_{\gamma_1} f(w)\, dw = \int_{\gamma_2} f(w)\, dw .$$

(10+8+8+8) = [34]

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) Part-IV/M.Stat. (Previous): 1980-81
SQC and OR

PERIODICAL EXAMINATIONS (Supplementary)

Date: 24.11.80 Maximum Marks: 100 Time: 3 hours

Note: Question 1 is compulsory. Attempt any three other questions from the remaining.

1. Explain the following:

(a) Control limits versus confidence limits, (b) Operating characteristic (OC) function of an $\bar{X}$ chart, (c) Shewhart's control chart, (d) 100%. inspection versus sampling inspection, (e) Percentage inspection and its demerits, (f) Dodge and Romig's sampling plans, and (g) Operating characteristic (OC) function of double sampling plan by attributes and (h) Extreme point of convex set of feasible solution and optimal feasible solution.

(5 x 8) = [40]

2. The specification on certain characteristic is 45 ± 2. Daily five items are selected at random from the production process and are measured. The averages and ranges for 24 samples each of size 5 are given below:

Day	Average	Range	Day	Average	Range	Day	Average	Range
1	43.7	1.9	9	45.0	1.5	17	45.1	2.8
2	44.6	2.1	10	45.4	2.1	18	45.7	1.8
3	44.9	3.2	11	45.2	1.3	19	44.9	2.5
4	45.9	1.3	12	45.4	1.6	20	45.3	1.8
5	44.6	2.5	13	44.8	1.5	21	45.9	1.9
6	44.5	0.9	14	45.4	1.7	22	44.7	1.9
7	45.6	2.0	15	45.1	2.8	23	45.6	1.8
8	45.4	2.1	16	45.7	1.8	24	45.0	2.4

(a) Draw and $\bar{X}$ - R chart and examine whether the process is under control.

(b) Obtain estimate of μ and standard deviation under statistical control and compute the percentage of defective items being produced.

(14+6) = [20]

3.(a) The following table gives the number of defective pieces due to run out in sample of 150, taken from each day's production.

Sample No.	Number of defective	Sample No.	No. of defectives
1	8	11	3
2	3	12	3
3	6	13	4
4	3	14	5
5	5	15	6
6	10	16	7
7	6	17	7
8	3	18	1
9	6	19	1
10	0	20	

Examine if the process is under control.

p.t.o

Contd..... Q.No.3

(b) The drawing specification for order diameter of a part is given as 89.450 gms as minimum and 89.970 mm. as maximum. Data on past performance gave the homogenised $\bar{R}$ for sample of size 4 as .0094 mm. The contribution of the part to the profit may be taken as Rs.50.00. However, the performance at this stage was considered to be unsatisfactory because of non-conformance to the specification. An out of specification item is rejected or reworked according as its diameter lies below 89.450 mm. or above 89.970 mm. respectively. The cost of rejection and rework per item can be taken as Rs.20/- and Rs.80/- respectively. Determine the average at which the process may be set to maximise the profit.

(12+8) = [20]

4. For the sampling plan N = 5000, n = 200 and c = 2

(a) Draw the O.C. curve and read the value of incoming lot quality for probability of acceptance 0.90.

(b) Draw AOQ curve and hence or otherwise verify the formula

$$p_L = y\left(\frac{1}{n} - \frac{1}{N}\right)$$

where $y = e^{-x}\frac{x^{c+2}}{c!}$ under usual notation.

[Hint. y = 1.371 for c = 2] (10+10) = [20]

5.(a) Let C be the set of all feasible solutions to a linear programming problem. Show that (i) C is a convex set and (ii) the maximum (minimum) value of the objective function of the linear programming problem occurs at one of the extreme point of C.

(b) A shop has three cargo holds, forward, aft. and center. The limitations are:

Hold	Cargo capacity by weight (tons)	Cargo capacity by volume (c.ft.)
forward	2000	100,000
center	3000	135,000
aft	1500	30,000

The following types of cargo is available for loading into the ship.

Cargo type	Total amount available (tons.)	Volume (c.ft.)	Profit per ton (Rs.)
A	6000	60	6
B	4000	50	8
C	2000	25	5

In order to preserve the trim of the ship, the weight in each hold must be proportional to the capacity in tons. How should the cargo be distributed so as to maximise the profit ?

(3+5+12) = [20]

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) Part-IV : 1980-81

Multivariate Analysis

PERIODICAL EXAMINATION

1980-81 : 451

Date: 17.11.80 Maximum Marks: 100 Time: 3 hours

Note: The paper carries 120 marks. You may attempt any part of any question. The maximum you can score is 100.

1.(a) Derive the density of a nonsingular multivariate normal variable starting from the definition (in terms of linear combinations). [15]

(b) Show that the sample mean vector and sample dispersion matrix from a multivariate normal distribution are independently distributed. [10]

2.(a) For the nonsingular multivariate normal distribution of $\underset{1\times p}{U'} = (\underset{1\times p_1}{U^{(1)'}}, \underset{1\times(p-p_1)}{U^{(2)'}})$ derive the conditional distribution of $U^{(2)}$ given $U^{(1)}$. [10]

(b) Based on a random sample of n observations on U, find maximum likelihood estimators of the parameters of such a conditional distribution. [15]

(c) Write down the model, sampling distribution results and the analysis of variance table for testing hypothesis of multiple regression parameters to be zero in the univariate case. Do the same for the case of multivariate regression with suitable arguments. [15]

3. Using only the definition of Wishart distribution and the properties of idempotent matrices (i.e., not using univariate results) show that if $\mathcal{U}$ is a nxp matrix of n independent $N_p(\cdot\ , \Sigma)$

(a) $\mathcal{U}' A \mathcal{U} \sim W_p$ iff A is idempotent. [5]

(b) $\mathcal{U}' A_1 \mathcal{U}$ and $\mathcal{U}' A_2 \mathcal{U}$ are independent Wishart iff A_1, A_2 are idempotent and $A_1 A_2 = 0$. [10]

4.(a) Defining Hotelling's T^2 statistic derive its distribution. [5]

p.t.o.

Contd..... Q.No.4

(b) Suppose $X_1, X_2, \ldots, X_p$ are measurements on the leftside of an organism and $X_{p+1}, X_{p+2}, \ldots, X_{2p}$ are measurements of the same quantities on the rightside. Assuming multivariate normality of the population distribution of the 2p-dimensional variable, formulate the problem and explain a method with formulae, of testing the left-right symmetry of the organism.

[10]

5.(a) In a psychiatric clinic, data on 670 patients were collected before and after a certain treatment with the help of tests. The following data represent statistics based on the differences (after - before). The variables are:

w : withdrawal retardation

h : hostility - suspiciousness

d : anxious depression

Mean vector : 0.0776   -0.4868   -0.0855

Inverse of Estimated Dispersion Matrix (based on 668 d.f.)

0.28113	0.03450	0.12522
	0.30783	-0.07606
		0.33330

Examine if the treatment is effective. [10]

(b) In an experiment with 2^2 factorial treatments ((1), a, b, ab) with 50 3-dimensional observations under each treatment the following data were obtained. Setup the Analysis of Dispersion Table. Test if interaction is present.

Treatment Total Vectors

(1)	5.396	2.770	4.260
a	5.006	3.428	1.462
b	5.121	3.042	2.563
ab	5.001	3.325	2.156

Pooled Dispersion Matrix on 196 d.f.

0.1953	0.0922	0.0996
	0.1211	0.0472
		0.1255

[15]

INDIAN STATISTICAL INSTITUTE

1980-81 : 472/583

B.Stat.(Hons.) Part-IV and M.Stat. Previous Year: 1980-81

Computer Programming

MID-YEAR EXAMINATION

Date: 8.12.80 Maximum Marks: 100 Time: 2 hours

Note: Answer all the questions.

1. Write a programme in FORTRAN to find and print the maximum, minimum, second maximum and second minimum of a given set of N values. There are N+1 data cards. The first data card contains the value of N punched in Cols. 1-3 in the mode I3. Each one of the later cards contains one value punched in Cols. 5-11 in the mode F 7.3.

[25]

2. The following instructions were given for the Ith sweep-out in the solution of N linear equations in N unknowns AX = Y, when Ith diagonal element after (I-1) sweep-outs is not zero. Examine whether the given set of instructions achieve the objective. If they do not, modify the instructions suitably to achieve the objective.

```
      DØ1J = I, N
      A(I, J) = A(I, J)/A(I, I)
1     Y(I) = Y(I)/A(I, I)
      DØ2K = I, N
      DØ2J = I, N
      A(K, J) = A(K, J) - A(K, I) * A(I, J)
2     Y(K) = Y(K) - A(K, I) * Y(I)
```

[7]

3. Given below is a programme in FORTRAN. Write down what will be the computer output when the machine comes to a halt.

```
      DIMENSION  A(4, 4)
      DØ1I  =  1, 4
      DØ1J  =  1, 4
1     A(I, J) = (-1)**(I-J)*(I-J)** 2 *(4-I)* I ** 2
      R = 0
      DØ2J = 1, 4
2     R = R + ABS (A(1, J))
      RM = R
      L  = 1
      DØ3I = 2, 4
      R = 0
      DØ4J = 1, 4
4     R = R + ABS (A(I, J))
      IF (RM-R) 5, 3, 3.
```

p.t.o.

Contd..... Q.3

```
5     RM = R
      L  = I
3     CONTINUE
      PRINT 10, L, RM
10    FORMAT (11X, I3 , F 10.2)
      STOP
```

[10]

4. Write a short note on function and subroutine subprogrammes. [8]

ASSIGNMENTS [50]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) Part-IV: 1980-81

1980-81 : 422

Measure Theory

MID-YEAR EXAMINATION

Date: 12.12.80 | Maximum Marks: 100 | Time: 3 hours

Note: The paper carries 117 marks. The maximum you can score is 100. All the functions are defined on appropriate measure spaces $(X, \underline{A}, \mu)$. All the measures are finite. $\sigma(\underline{C})$ stands for the σ-field generated by $\underline{C}$. Write your answers clearly and legibly.

1.(a) Define $\int f\,d\mu$ for any nonnegative measurable function f.

(b) Prove the most general form of the monotone convergence Theorem (Give complete details).

(3+15) = [18]

2.(a) State and Prove the Dominated convergence Theorem (Give complete details).

(b) If $\{f_n\}_{n\geq 1}$ is a sequence of measurable functions such that $|f_n(x)| \leq \frac{1}{2^n}$ for all x. then show that $\int \sum_{n=1}^{\infty} f_n\,d\mu = \sum_{n=1}^{\infty} \int f_n\,d\mu$.

(c) Is it in general true that $\int f_n\,d\mu \to \int f\,d\mu$ for any sequence of measurable functions $\{f_n\}_{n\geq 1}$ converging to f ? [Hint: Try the functions $nI_{(0,\,1/n)}$].

(15+5+5) = [25]

3. If $(X_1, \underline{A}_1, \mu_1)$ and $(X_2, \underline{A}_2, \mu_2)$ are two measure spaces prove the existence of a measure μ on $\underline{A}_1 \otimes \underline{A}_2$ such that $\mu(A_1 \times A_2) = \mu_1(A_1).\mu_2(A_2)$ for $A_1 \in \underline{A}_1$ and $A_2 \in \underline{A}_2$. Show that such a measure is unique.

[15]

4.(a) State Fubini's Theorem.

(b) Show that the set $A = \{(x, y) \in [0, 1] \times [0, 1] : x+y \geq 1\}$ belongs to $\underline{B} \otimes \underline{B}$ where $\underline{B}$ is the Borel σ-field of $[0, 1]$. Find $\lambda \otimes \lambda(A)$ where λ is the Lebesgue measure on $\underline{B}$. Evaluate all the integrals explicitely.

(3+15) = [18]

p.t.o.

5. Describe clearly the correspondence between the probabilities on the Borel σ-field of the real line and the distribution functions on the real line. Show that this correspondence is one-one. [10]

6. Let μ be a real valued function defined on the power set $\underline{P}(X)$ of a set X satisfying the properties

(i) $0 \leq \mu(A) \leq 1$, $\mu(\emptyset) = 0$ and $\mu(X) = 1$

(ii) $\mu(A \cup B) + \mu(A \cap B) \leq \mu(A) + \mu(B)$

(iii) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$

(iv) $A_n \uparrow A \Rightarrow \mu(A_n) \uparrow \mu(A)$

Then show that $\underline{D} = \{ A : \mu(A) + \mu(A^c) = 1 \}$ is a σ-field and that μ restricted to $\underline{D}$ is a probability.

[15]

7.(a) If $\underline{G}_1$ is a collection of subsets of a set X_1 and $\underline{G}_2$ is a collection of subsets of a set X_2 show that

$$\sigma(\underline{G}_1) \otimes \sigma(\underline{G}_2) = \sigma(\{ A \times B : A \in \underline{G}_1, B \in \underline{G}_2 \}).$$

(b) If f is a quasiintegrable function such that $\int_E f \, d\mu = 0$ for all $E \in \underline{A}$ then show that $f = 0$ a.e. $[\mu]$.

(c) Prove that $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$ for any quasi-integrable function f.

(d) If f is integrable show that $\mu(\{ x : |f(x)| \geq 1 \}) < \infty$.

(4+4+4+4) = [16]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) Part-IV: 1980-81

1980-81: 452

Multivariate Analysis

MID-YEAR EXAMINATION

Date: 15.12.80 Maximum Marks: 100 Time: 4 hours

Note: The paper carries 120 marks. You may attempt any part of any question. Any score over 100 will be taken as 100.

1.(a) Show that if $X_1, X_2, \ldots, X_n$ are i.i.d. $N_1(0, 1)$, then the joint density of $X_1, X_2, \ldots, X_n$ is constant on the sphere $\sum_{i=1}^{n} x_i^2 = C$. Show that this is a characteristic property of $N_1(0, 1)$ among random variables with differentiable densities.

[Hint: Consider the partial derivates of the joint density on $\sum_{i=1}^{n} x_i^2 = C$.] [5]

(b) Show more generally that if $\underset{p \times n}{X'}$ is a random sample of size n from a p-variate differentiable density $f(x_1, x_2, \ldots, x_p)$ then f is N_p iff the joint density of X' depends only on the matrix $X'X$. [7]

2. Let $U \sim N_p(\mu, \Sigma)$ where $\sigma_{ii} = 1$, $\sigma_{ij} = \rho$ for $i \neq j$, $\Sigma = ((\sigma_{ij}))$. Find on the basis of a random sample of size n from U, maximum likelihood estimators of μ, ρ. [15]

3. Let $\bar{U}$ be the sample mean and S be the sample corrected sum of squares and products matrix from a random sample of size n from $N_p(\mu, \Sigma)$, with $|\Sigma| \neq 0$. Then let

$$T^2 = n(n-1)(\bar{U} - \mu)' S^{-1}(\bar{U} - \mu).$$

Show that $T^2 = n(n-1) \max_{\underset{p\times 1}{L}} \frac{[L'(\bar{U} - \mu)]^2}{L' SL}$.

Give an interpretation of this result in terms of test criteria for $H_0 : \mu = \mu_0$. [10]

p.t.o.

4. Write down the multivariate linear model. Assuming results on univariate theory, Wishart distribution and estimation of the multivariate parameters in the linear model (but quoting clearly and precisely), develop procedures for testing a multivariate linear hypothesis. [30]

5.(a) Formulating the problem clearly derive the first principal component from a dispersion matrix. [6]

(b) Formulating the problem clearly derive the first canonical variables of the partitioned dispersion matrix

$$\begin{bmatrix} \underset{p \times p}{\Sigma_{11}} & \Sigma_{12} \\ \Sigma_{21} & \underset{q \times q}{\Sigma_{22}} \end{bmatrix}.$$ [9]

6.(a) Let $U = \begin{bmatrix} \underset{k \times 1}{X} \\ \underset{(p-k) \times 1}{Y} \end{bmatrix}$. Show that the first k canonical variables of X with respect to U are the first k principal components of X. [7]

(b) Find the principal components and their variances of $R = ((\rho_{ij}))$ with $\rho_{ij} = \rho$, for $i \neq j$ and $\rho_{ii} = 1$, for all i. [8]

7. At an agricultural experimental station, 52 samples of soil were collected of which 25 contained the organism Azotobacter and 27 did not. Three characteristics of the soil were studied:

H_1 : pH, X_2 : amount of readily available phosphate

X_3 : total nitrogen content.

Data are summarised below.

	With Azotobacter		
	X_1	X_2	X_3
Total	184.9	2196	911
	Without Azotobacter		
Total	160.6	963	592

Corrected sum of squares and products matrix based on 50 d.f.

20.9570	421.9173	148.2403
	88,897.3600	9,135.0933
		78.0385

Test it the two groups are different in the mean vector and if so find the linear discriminant function between the two groups. (State all your assumptions clearly.) [23]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) Part-IV: 1980-81

1980-81 : 442

Design of Experiments

MID-YEAR EXAMINATION

Date: 18.12.80 Maximum Marks: 100 Time: 3 hours

Note: Question No.6 is compulsory. Of the rest, answer any four questions. All the questions carry equal marks.

1.(a) What are lattice designs ? Indicate a method of construction of an m-ple lattice design involving v treatments. Show the actual layout for $v = m^2 = 9$.

(b) Work out the usual analysis (under a fixed-effects additive model) of an m-ple lattice design and give the expression for the average variance.

2. In the context of the analysis of a connected block design, recall that, in the usual notations,

$$E(\underline{Q}) = C\underline{\tau}, \quad D(\underline{Q}) = \sigma^2 C$$

where C is the C-matrix of rank (v-1).

Consider the following canonical reduction of C:

$$C = \sum_{i=1}^{v-1} \theta_i \, \underline{\xi}_i \underline{\xi}_i' \quad \text{with } \theta_i > 0, \ \underline{\xi}_i \perp \underline{\xi}_j,$$

$$\underline{\xi}_i \perp (1, 1, \ldots, 1)' \quad \text{etc.}$$

Define $\underline{\eta}_i = \underline{\xi}_i' \underline{\tau}$ $(i = 1, 2, \ldots, v-1)$.

(a) Show that all the η_i's are estimable

(b) Determine $D(\hat{\underline{\eta}})$ where $\underline{\eta} = (\eta_1, \eta_2, \ldots, \eta_{v-1})'$ and, hence, calculate $\bar{V} = \frac{1}{v-1} \Sigma V(\hat{\eta}_i)$ and show that $\bar{V} \propto (\Sigma \theta_i^{-1})$.

(c) For given b, v and k (in usual notations), assume that all the designs involved are binary. Show that $\bar{V}$ in (b) attains a minimum for the BIBD(b, v, r, k, λ), whenever it exists.

[Hint: Use the H.M. - A.M. inequality]

3.(a) In relation to a 2^n-factorial experiment, define main effect and interactions of orders 1, 2, ... Explain the role of Hadamard matrices in getting the algebraic expressions for the various factorial effects and interactions in terms of the yields arising out of a 2^n - factorial experiment laid out in r complete blocks. Explain Yates' technique in this context.

p.t.o.

Contd......Q.No.3

(b) Deduce that in the above experiment, each main effect and interaction has variance $(\sigma^2 / r.2^{n-2})$ where σ^2 is the per plot intrablock variance.

4.(a) In a $(2^5, 2^2)$ experiment, the Key block is known to contain the following treatment combinations (in usual notations):

1, ac, de, acde, abd, abe, bcd, bce.

Find out the confounded interactions and show the complete layout.

(b) Suppose we want to construct a $(2^5, 2^2)$ confounded experiment in minimum number of replications so as to ensure balance over all the 3 - and 4 - factor interactions (without confounding any main effect or 2-factor interactions). In addition to the replication in (a) above, what others are needed ? Show only the Key blocks of the others.

5.(a) Construct a confounded 3^3 design in 9 blocks of 3 plots each in which 4 of the d.f. are carried by the pencils P(1, 1, 1) and P(1, 1, 2). What are the other d.f. confounded ?

(b) Describe the missing plot technique and bring out Fisher's rule. Work out the ANCVA for an RBD with one missing observation.

6. Submit Practical Note Book.

1980-81 : 412

INDIAN STATISTICAL INSTITUTE
B.Stat (Hons.) Part IV: 1980-81
Complex Analysis
MID-YEAR EXAMINATION

Date 20.12.80 | Maximum Marks : 100 | Time: 3 hours

Note: This paper carries 112 marks. Answer as many questions as you can. The maximum you can score is 100. If you use any theorems proved in class, state them clearly and precisely.

1. a) Define the radius of convergence of a power series. Compute the radius of convergence of the power series $\sum_{n=1}^{\infty} \frac{n}{2^n} z^n$

   b) Show that the function $f(z) = \bar{z}$ does not have a complex derivative at any point in the complex plane.

   c) Define the index of a piecewise smooth closed path at a point. Let γ be the path on $[0, 4\pi]$ defined by

$$\gamma(t) = e^{it} \quad \text{if } 0 \le t < 2\pi$$
$$= e^{i(4\pi - t)} \quad \text{if } 2\pi \le t \le 4\pi.$$

   What is the index of γ at $z = 0$? [5+5+5 = 15]

2. a) State and prove Cauchy's theorem for convex regions.

   b) Let Ω be a region in the complex plane and γ a piecewise smooth closed path in Ω. Let f be a holomorphic function on Ω. Is it always true that $\int_\gamma f(z)\,dz = 0$? If the answer is yes, give a proof, otherwise give an example. [10+7 = 17]

3. a) Let f be an entire function taking values in $\mathbb{C} - D_a(\delta)$, where $D_a(\delta) = \{z : |z-a| < \delta\}$. Then show that f has to be a constant.

   b) Conclude, using a), that the range of a nonconstant entire function is dense in the plane. [8+4 = 12]

4. Evaluate the following integrals.

a) $\int_{\gamma} \operatorname{Re} z \, dz$, where γ is the positively oriented circle with centre 0 and radius 1.

b) $\int_{\gamma} \frac{dz}{z^2+1/4}$, where γ is as in a).

c) $\int_{\gamma} e^z z^{-n} dz$, where γ is as in a). [5+5+5 = 15]

5. a) How many holomorphic functions are there on $D = \{z : |z| < 1\}$ such that $f(-\frac{1}{n}) = \frac{1}{n^3}$, $n \geq 2$?

b) How many holomorphic functions are there on D taking values in $\bar{D} = \{z : |z| \leq 1\}$ such that $f(0) = i$? [5+7 = 12]

6. a) Define the order of a zero of a holomorphic function. What are the orders of the zeros of $(z+3)(z^2-4)^3$?

b) Define removable singularity, pole, order of a pole and essential singularity. Identify (giving reasons) the nature of the isolated singularities for the following functions :

i) $\frac{1}{z(z^2+4)^2}$ ii) e^{-1/z^2} iii) $\frac{e^z-1}{z}$ [6+15 = 21]

7. a) Define the residue at a pole. State the residue theorem for convex regions.

b) Evaluate the following integrals.

i) $\int_{-\infty}^{\infty} \frac{x^2-x+2}{x^4+10x^2+9} dx$ ii) $\int_{-\infty}^{\infty} \frac{e^{itx}}{1+x^2} dx$, where t is a real number. [6+14 = 20]

INDIAN STATISTICAL INSTITUTE

M. Stat. Previous Year and
B.Stat.(Hons.) Part-IV: 1980-81

SQC and OR

MID-YEAR EXAMINATIONS

Year No.
1980-81 : 462/584

Date: 24.12.80 Maximum Marks : 100 Time: 3 hours

Note: Attempt only four questions. The marks allotted for each question are given in the brackets.

1. (a) Given a convex set C and a point X_0 not belonging to C, show that there exists a hyperplane $\alpha x = \beta$ passing through X such that C is on one side of the hyperplane.

   (b) If a linear programming problem can be solved then its optimal solution has at the most m positive variables where m denotes the rank of the coefficient matrix A for the problem; maximise = C X subject to A X = b, X ≥ 0, the symbols have the usual meaning.

(8+12) = [20]

2. (a) Define a supporting hyperplane. Show that a closed convex set which is bounded from below has an extreme point in every supporting hyperplane.

   (b) Solve by simplex method the following L.P.P.

$$\text{Max.} = 10x_1 + x_2 + 2x_3$$

$$\text{Subject to} \quad x_1 + x_2 - 2x_3 \leq 10$$
$$x_1 + x_2 + x_3 \leq 20$$
$$x_1, x_2, x_3 \geq 0$$

(2+6+12) = [20]

3. (a) What are artificial Variables ?

   (b) Solve the following linear programming problem by using charne's M - method

$$\text{Minimise} = 3x_1+4x_2+x_3+6x_4$$

$$\text{subject to} \quad 5x_1-2x_2+x_3-3x_4 \geq 2$$
$$6x_1+x_2-5x_3-3x_4 \geq 5$$
$$-x_1+4x_2+3x_3+7x_4 \geq 6$$
$$\text{all } x_j \geq 0$$

(4+16) = [20]

4. (a) Explain briefly the Revised simplex Form I to solve the linear programming problem.

   (b) Show that if Primal linear programming problem has an optional solution then its dual problem also has an optional solution.

4. (c) Solve the following problem by duality consideration

$$\text{Minimise } Z = 3x_1+2x_2+x_3+4x_4$$

$$\text{subject to } \quad 2x_1+4x_2+5x_3+x_4 \geq 10$$

$$3x_1-x_2+7x_3-2x_4 \geq 2$$

$$5x_1+2x_2+x_3+6x_4 \geq 15$$

$$x_1, x_2, x_3 \text{ and } x_4 \geq 0$$

(3+3+14) = [20]

5. (a) For an item the unit cost of storage is Re 1 per month, ordering cost is .. 25 per order and demand is 200 units per month.

Find the optimum quantity to be ordered and hence determine the total cost of ordering and storage per month.

(b) Probability that exactly d units of a spare part are required is given by P_d. The unit cost of spare is c money unit, loss incurred due to shut down of machine arising out of shortage per unit is U money units and salvage value of an unused spare is V money units. Work out an optimal inventory policy minimising the average total loss, stating your assumptions, if any.

(c) An equipment is purchased for N money units The running cost and resale value of the equipment at the end of ith year are r_i and s_i respectively. Obtain a condition for optimal replacement period from given information. (5+10+5) = [20]

Sessionals and educational visit report .... [20]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) Part-IV: 1980-81

1980-81: 432

Inference

MID-YEAR EXAMINATION

Date: 27.12.80 Maximum Marks: 100 Time: 3 hours

Note: Open note examination. Points of each question is in the margin.

1. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. Poisson with parameter λ i.e.

$$P(X = K) = e^{-\lambda} \cdot \frac{\lambda^K}{K!}, \quad K = 0, 1, 2, \ldots$$

Find the U.M.V. unbiased estimate of $P(X = 0) = e^{-\lambda}$. [20]

2. Give an example of a family of distribution, which is boundedly complete but not complete. [15]

3. Let $X_1, \ldots, X_n$ be i.i.d. $N(\theta, 1)$. We want to estimate θ under the loss function

$$L(a, \theta) = |a - \theta|$$

Show that $\bar{X}_n$ is an admissible estimate of θ. [20]

4. Let X be a Poisson r.v. with parameter θ. We want to estimate θ under the loss function

$$L(a, \theta) = \frac{(a - \theta)^2}{\theta}$$

By using Cramer-Rao inequality method show that $\delta(X) = X$ is an admissible estimate of θ. [20]

5. Let $X_1, \ldots, X_n$ be a sample from uniform distribution $\mathcal{U}(0, \theta)$ i.e.

$$f(x) = \begin{cases} \frac{1}{\theta} & 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$$

Find the best invariant estimate of θ if the loss function is

$$L(\theta, a) = \left| \frac{a}{\theta} - 1 \right|$$ [10]

p.t.o.

6. Let $X_1, \ldots, X_n$ be a sample from a gamma distribution $\gamma(\alpha, \beta)$ with α known and β unknown i.e.

$$f(x) = \begin{cases} \beta^{-\alpha} (\Gamma(\alpha))^{-1} e^{-\frac{x}{\beta}} x^{\alpha-1} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Show that best unbiased estimate of β is also the best invariant estimate of β when the loss function is

$$L(\beta, a) = (a/\beta) - 1 - \log(a/\beta).$$

[15]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV year: 1980-81

1980-81: 433

Advanced Linear Estimation and Inference
PERIODICAL EXAMINATION

Date: 2.3.81 | Maximum Marks: 100 | Time: 3 hours

Note: All questions have equal weight.

1. Let $X_1, \ldots, X_n$ be a sample from the uniform distribution $u(\theta, 2\theta)$, where $\Theta = (0, \infty) = \mathcal{A}$ and $L(\theta, a) = \frac{(\theta-a)^2}{\theta^2}$.
   Show that the best invariant decision rule is

$$d(X) = \frac{(n+2)[(V/2)^{-(n+1)} - u^{-(n+1)}]}{(n+1)[(V/2)^{-(n+2)} - u^{-(n+2)}]}$$

   Where $u = \min(X_1, \ldots, X_n)$, $V = \max(X_1, \ldots, X_n)$.

2. Let X and Y be random variables with joint density

$$f_{X,Y}(x,y \mid \lambda,\mu) = \lambda\mu\, e^{-\lambda x-\mu y}\, I_{(0,\infty)}(x)\; I_{(0,\infty)}(y)$$

   Find a UMP unbiased test of size α for testing $H_0: \lambda = \mu$ vs $H_1: \lambda \neq \mu$.

3. Let $X_1, \ldots, X_n$ be a sample from $u(\theta_1, \theta_2)$, i.e., uniform distribution on the interval (θ_1, θ_2). We want to test

$$H_0: \theta_2 \geq 0 \quad \text{vs} \quad H_1: \theta_2 < 0.$$

   Find a UMP unbiased test of size α.

4. Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be independent samples from $N(\mu, \sigma^2)$ and $N(\eta, \sigma^2)$. Find a UMP unbiased test for testing $H_0: \mu = \eta$ vs $H_1: \mu \neq \eta$ at level of significance α.

5. Let $X_1, X_2, \ldots, X_n$ be a sample from

$$f(x \mid \theta) = [2(1 + \cosh(x_i - \theta))]^{-1}$$

   Find the locally best test for

$$H_0: \theta = 0 \quad \text{vs} \quad H_1: \theta > 0$$

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV year: 1980-81
Probability Theory and its Applications

PERIODICAL EXAMINATION

1980-81 + 423

Date: 9.3.81 Maximum Marks: 100 Time: 3 hours

Note: Answer as much as you can. The maximum you can score is 100.

1.(a) Let X be a random variable with $E(e^{aX}) < \infty$, $(a > 0)$. Show that

$$P\left\{X \geq \varepsilon\right\} \leq \frac{E(e^{aX})}{e^{a\varepsilon}}$$ [10]

(b) Let $A(r) = [E|X|^r]^{1/r}$, $r > 0$, for a random variable X. Prove that

$$A(r) \leq A(s) \quad \text{if } 0 < r < s$$ [10]

2.(a) Show that the sequence of random variables X_n converges to a random variable X a.e. if and only if for every $\varepsilon > 0$

$$\lim_{N \to \infty} P\left\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon \text{ for some in } n \geq N\right\} = 0$$ [10]

(b) Define convergence in probability. Show that a sequence of random variables X_n converges in probability to a random variable X if and only if $\lim_{n \to \infty} E\left[\frac{|X_n - X|}{1 + |X_n - X|}\right] = 0$. [10]

3. X_n, $n \geq 1$ are independent random variables with $P\{X_n = 1\} = p_n$ and $P\{X_n = 0\} = 1 - p_n$

(a) Show that $X_n \to 0$ in probability if and only if $\lim_{n \to \infty} p_n = 0$. [6]

(b) Show that $X_n \to 0$ a.e. if and only if $\sum_1^\infty p_n < \infty$. [14]

(c) If $p_n = p$ for all n, show that

$$\left[\frac{X_1 + X_2 + \dots + X_n}{n} - p\right] \to 0 \text{ in probability.}$$ [10]

p.t.o.

.(a) Let Y be a non-negative random variable. Prove that

$$\sum_{n=1}^{\infty} P\left\{Y \geq n\right\} \leq E(Y)$$

[Hint: $\displaystyle\sum_{n=1}^{\infty} P\left\{Y \geq n\right\} = \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} P\left\{k \leq Y < k+1\right\}$

$\displaystyle = \sum_{k=1}^{\infty}\sum_{n=1}^{k} P\left\{k \leq Y < k+1\right\} = \sum_{k=1}^{\infty} k\, P\left\{k \leq Y < k+1\right\}$]

[10]

(b) $X_1, X_2, \ldots$ is a sequence of <u>identically</u> distributed random variables with $E\,|X_1| < \infty$. Using (a), show that

$$P\left\{\omega : |X_n(\omega)| \geq n \text{ for infinitely many } n\right\} = 0.$$

[20]

5. Decide whether the following events are in the tail σ-field with respect to the sequence $\left\{X_n\right\}$ of random variables. Give reasons for your answer.

(a) $\left\{\omega : X_n(\omega) \text{ converges to } 5\right\}$

(b) $\left\{\omega : \dfrac{X_1(\omega) + X_2^2(\omega) + X_3^3(\omega) + \cdots\cdots + X_n^n(\omega)}{n} \text{ converges}\right\}$

(c) $\left\{\omega : X_1^2(\omega) + X_2^2(\omega) + \cdots\cdots + X_n^2(\omega) \text{ converges to } 2\right\}$

(d) $\left\{\omega : X_n(\omega) > n^2 \text{ for infinitely many } n\right\}$

(5 x 4) = [20]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV year

and

M.Stat. Previous Year : 1980 - 81

1980-81: 473/572

Biostatistics

PERIODICAL EXAMINATIONS

Date: 23.3.81 Maximum Marks: 100 Time: 3 hours

Note: The question paper carries 112 marks. You can answer any part of any question. Maximum marks you can score is 100.

1. What is random mating ? State and prove Hardy-Weinberg equilibrium law for the case of a character controlled by a single autosomal locus.

   Describe the corresponding progress of a population under random mating for a sex-linked gene.

   [20]

2. For the case of two linked loci (autosomal characters), indicate the progress of the population under random mating and show that in the long run the population tends to one in which all genes are combined strictly at random according to their frequencies.

   [20]

3. Consider the case of a single Mendelian character in which the genotypic array in generation 0 is

$$p_0^2\, AA + 2p_0 q_0\, Aa + q_0^2\, aa, \quad p_0 + q_0 = 1$$

   and suppose that only (1-k) of the recessives survive to reproduce in each generation, but otherwise there is no selection, $0 < k < 1$ and the mating is random.

   Prove from the beginning the relation

$$u_{n+1} = \frac{u_n(1+u_n)}{1+u_n-k},$$

   where $p_n = \dfrac{u_n}{1+u_n}$ = frequency of gene A in the n-th generation.

   Hence, show that if k is small,

$$\Delta u_n = u_{n+1} - u_n = \frac{k\, u_n}{1+u_n}$$

   and approximately

$$kn = u_n - u_0 + \log_e\left(\frac{u_n}{u_0}\right)$$

   [20]

p.t.o.

4. Suppose F_2 data are available on two autosomal characters (You can assume complete dominance for both the characters). State, giving sufficient reasons, how you will test for the presence of linkage. Assuming recombination fraction (p) for the male and female same and both in the coupling phase in F_1 generation, find the maximum likelihood estimator of p and its variance. Obtain also the expression for i_p = per unit of item information contained in the F_2 data regarding the recombination fraction p. [20]

5. Write short notes on:

   (i) Mendel's law of segregation

   (ii) Mendel's law of independent assontment

   (iii) Mitotis and meiosis

[12]

6. (i) Suppose a recessive trait occurs in 1 in 1000 of a random mating population. How many generations of complete selection against the recessive individuals would be necessary to reduce the proportion to 1 in 100,000 ?

   (ii) Suppose the mating is random and we start with the population

$$\frac{1}{2} S_1 S_2 + \frac{1}{3} S_1 S_3 + \frac{1}{6} S_2 S_3 ,$$

where S_1, S_2 and S_3 are self-sterility genes. What are the proportions of $S_1 S_2$ individuals in generation 5, 10, 100, 1000 ?

(10 + 10) = [20]

1980-81: 413/521

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV year
and
M.Stat. Previous Year: 1980-81
Modern Algebra

PERIODICAL EXAMINATION

Date: 30.3.81 Maximum Marks: 100 Time: 3 hours

Note: All questions carry equal marks. Answer any 5 questions.

1. Prove that any group of order 13 is abelian.

2. Let G be a group. Suppose there is an integer k such that for any a, b in G, $(ab)^k = a^k b^k$, $(ab)^{k+1} = a^{k+1} b^{k+1}$ and $(ab)^{k+2} = a^{k+2} b^{k+2}$. Prove that G is abelian.

3. Let M, N be normal subgroups of a group G with $M \cap N = (e)$, where e is the identity element of G. Show that for m in M and n in N, mn = nm.

4. Let F be a field and K a nonempty subset of F. Show that K is a subfield of F if, and only if, for any a, b in K, a-b is in K and, if $b \neq 0$, ab^{-1} is in K.

5. Let F be a field and F[x] the ring of polynomials over F. Prove that any ideal in F[x] is a principal ideal. Show that an ideal (p(x)) in F[x] is maximal if, and only if, p(x) is irreducible.

6. Let D be an integral domain. Let M, N be two ideals in D with $M \cap N = (0)$. Show that M = (0) or N = (0).

7. Let m, n ... relatively prime positive integers. Let D be an integral domain and a, b elements of D with $a^m = b^m$ and $a^n = b^n$. Show that a = b.

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV Year: 1980-81

1980-81: 443

Advanced Sample Surveys

PERIODICAL EXAMINATION

Date: 6.4.81 Maximum Marks: 100 Time: 3 hours

Note: Answer questions numbered 1 and 2 and any one from the rest. Marks allotted to them are indicated in parentheses.

Practical records carrying 20 marks are to be submitted at the Examination Hall.

1. Answer any two of the following:

(i) In sampling a finite population of size N by PPSWR method in n draws obtain formulae for expectation and variance of effective sample-size in terms of N, n and normed size-measures P_i's.

(5+5) = [10]

(ii) In two-stage sampling from a finite population show that you may unbiasedly estimate the variance of an unbiased estimate of the population total from a single sample even if the second stage units are chosen by a systematic sampling method provided the first stage units are chosen by PPSWR method.

[10]

(iii) Show how Lahiri's scheme of sampling a unit from a finite population really ensures selection with probabilities proportional to size-measures of units. Also demonstrate the corresponding result for the extended scheme due to Lahiri-Midzuno.

(7+3) = [10]

(iv) Show that the usual ratio-estimator becomes unbiased for a finite population mean if Midzuno sampling scheme is employed. How will you unbiasedly estimate the variance of the ratio-estimate in this case ?

(3+7) = [10]

2. The following table gives information on 10 households in a street in Calcutta about their composition and ownership of T.V. sets.

Serial number of household	Household size	Whether a T.V. set is possessed ('1' if 'yes' ; '0' if 'no')
1	7	1
2	3	1
3	5	0
4	4	1
5	2	0
6	3	1
7	8	0
8	3	1
9	5	1
10	5	0

p.t.o.

Contd..... Q.No.2

Choose a sample of households in two draws without replacement with selection-probabilities proportional to their sizes.

From your sample obtain three different unbiased estimates for the proportion of households possessing a T.V. set. What are the errors of your estimates ? Obtain an unbiased estimate of the variance of one of your estimates.

(8+3 x 5+1+6) = [30]

3. Explain fully why Murthy estimator is more efficient than Des Raj estimator and the latter is more efficient than Hansen-Hurwitz estimator when same normed size-measures are used for each and each is based on samples taken with the same number $n(\geq 2)$ of draws from a given finite population.

(15+15) = [30]

4. Show that in the class of all unbiased estimators for a finite population total there does not exist one with a uniformly smallest variance unless one has a census.

Show that in the class of homogeneous linear unbiased estimators (HLUE) for a finite population total there does not exist one with a uniformly smallest variance, unless a sampling design belongs to an exceptional class. Characterize this exceptional class. Show that for such an exceptional design a unique uniformly minimum variance estimator exists in the HLUE class.

Obtain an admissible estimator for a finite population total in the HLUE class, for a general class of sampling designs.

(8+8+3+5+6) = [30]

5. Explain how you may unbiasedly estimate a finite population total on taking a sample in several stages choosing the first stage units (fsu's) with varying probabilities without replacement, the sample-size being fixed in advance. Obtain formulae for its variance and unbiased variance-estimator.

Suppose you are to estimate a finite population mean of second-stage-unit (ssu's) values from a two-stage sample when each fsu contains the same number of ssu's. If your plan is to use the sample mean on taking an SRSWOR of fsu's and independent SRSWOR's of ssu's from each selected fsu, then considering the variance function and a simple cost function discuss how you may reasonably decide on the choice of the sample sizes at the two stages of sampling.

(4+8+8+10) = [30]

1980-81 : 3E1/481/571

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year (Elective-5: Economics)
and
B.Stat. (Hons.) IV Year and M.Stat. Previous Year: 1980-81
(Econometrics)

PERIODICAL EXAMINATIONS

Date: 13.4.81 Maximum Marks: 100 Time: 3 hours

Group - A

(Answer any two questions)

1.(a) Define the Lorenz Curve and Lorenz ratio of an income distribution.

(b) Find the equation of the Lorenz curve for an exactly Paretean income distribution over the income range (c, ∞), where c is subsistence income $(c > 0)$.

(c) In (b) above, what would be the equation of the Lorenz curve of the truncated income distribution over (c', ∞) where $c' > c$?

(d) Suppose, you are given some income data and you plot a graph showing $\log T_x$ against $\log N_x$, where N_x is the number of earners earning x or more, and T_x is the total income of these N_x persons. What would be the equation of the graph if the income distribution is Paretean ?

(6+10+7+7) = [30]

2.(a) Find the expressions for mean and variance of lognormal distribution $\Lambda(\mu, \sigma^2)$.

(b) Suppose income $X \sim \Lambda(\mu, \sigma^2)$ and consumer expenditure $C = \alpha X^{\beta}$ exactly, where α, β are positive constants. What can you say about the size distribution of C ? What are its mean and median ?

(c) Give a broad account of the method of quantiles for estimation of parameters of a two-parameter lognormal distribution. Mention, in particular, the choice of quantiles which gives estimates with the highest asymptotic efficiency.

(6+10+14) = [30]

3. Write short notes on any two of the following:

(a) The moment distribution property of the lognormal distribution and its uses.

(b) General properties of a Lorenz curve.

(c) Measurement of income inequality.

(2 x 15) = [30]

p.t.o.

Group - B

4. Examine the size distribution of income given below and fit a Paretean distribution over the appropriate range. (Compute expected frequencies and show the fitted Pareto line, but you need not test the goodness of fit)

Income (Rs.)	No. of earners
10001 - 20000	6286
20001 - 30000	1404
30001 - 50000	696
50001 - 75000	218
75001 -100000	82
100001 -200000	74
200001 -	25

[30]

5. Practical Record. ... [10]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV Year: 1980-81

Sequential Analysis and Nonparametric Methods

PERIODICAL EXAMINATION

Date: 20.4.81 Maximum Marks: 100 Time: 3 hours

1. Let $f(x, \theta)$ be the frequency function of a random variable x with parameter θ.

   (i) State what is meant by a sequential procedure of testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$. How does it differ from a fixed sample testing procedure ?

   (ii) Define Wald's SPRT of strength (α, β) for the above problem. Determine the approximate values of the constants involved.

   (iii) If (α', β') be the strength of SPRT with approximate values of constants as defined in (ii), show that

$$\alpha' + \beta' \leq \alpha + \beta$$

2.(a) Let $\{x_i\}$ be i.i.d. random variables having a frequency function $f(x, \theta)$. Let for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$, SPRT is constructed and a terminal decision is taken at the n-th stage. Let

$$Z_n = \sum_{i=1}^{n} \log \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)}$$

Then show that

$$E\,(e^{Z_n} \mid H_0 \text{ accepted}, \theta_0) = \frac{L(\theta_1)}{L(\theta_0)}$$

$$E\,(e^{Z_n} \mid H_0 \text{ rejected}, \theta_1) = \frac{1 - L(\theta_1)}{1 - L(\theta_0)}$$

Assume SPRT in this case terminates with probability 1.

(b) Let $\{x_i\}$ be i.i.d. random variables having c.d.f. $F(x, \theta)$, $\theta \in (H)$ and consider $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$. A sequential test is defined as follows.

At the n-th stage, accept H_0 if $x_n \in R^0$ and reject H_0 if $x_n \in R^1$ where R^0 and R^1 are disjoint sets of real line.

Then show that for all $\theta \in (H)$

$$L(\theta) = \gamma_0(\theta)[\gamma_0(\theta) + \gamma_1(\theta)]^{-1}$$

$$E(n, \theta) = [\gamma_0(\theta) + \gamma_1(\theta)]^{-1}$$

Where $\gamma_0(\theta) = \int_{R^0} dF(x, \theta)$, $\gamma_1(\theta) = \int_{R^1} dF(x, \theta)$.

3.(a) For a SPRT of strength (α, β) for testing $H_0 : \Theta = \Theta_0$ against $H_1 : \Theta = \Theta_1$, define the O.C. and A.S.N. function, the underlying cdf being $F(x, \Theta)$.

(b) Find an approximate expression of O.C. function of the above SPRT under suitable assumptions to be stated by you.

(c) Let the underlying density of the above problem is $N(\Theta, \sigma^2)$, and we want to test $H_0 : \sigma = \sigma_0$ against $H_1 : \sigma = \sigma_1$ Θ being known. Find the expression for O.C. in this case. Also deduce few standard points of O.C. in the present set up.

4.(a) Let $Z_1, Z_2, \ldots$ be identically distributed random variables. Then under certain conditions show that

$$E\left(\sum_{i=1}^{N} Z_i\right) = E(N)\,E(Z_1)$$

State the conditions.

(b) Let Z be a random variable with moment generating function

$$\phi(t) = E(e^{tz})$$

State a set of sufficient conditions so that the equation

$$\phi(t) = 1$$

has non zero solutions. Discuss different cases. Give proofs whenever necessary.

5. Describe Stein's two stage procedure for obtaining a fixed width confidence interval for μ in $N(\mu, \sigma^2)$ with confidence coeffecient at least $1-\alpha$, σ^2 being unknown and α is preassigned. State and prove relevant results.

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV Year : 1980-81

1980-81: 424

Probability Theory and its Applications

ANNUAL EXAMINATION

16.5.81 Maximum Marks: 100 Time: 3 hours

Note: Answer <u>any five</u> questions.

1.(a) Define the distribution function of a random variable. [3]

(b) Prove that a distribution function can have at most countably many discontinuity points. [5]

(c) Prove that if $F_n(x)$, $n \geq 1$, $F(x)$ are distribution functions on R such that $\lim_{n \to \infty} F_n(r) = F(r)$ for every rational number r, then F_n converges weakly to F. [12]

2. $X_1, X_2, \ldots$ are independent random variables with distribution functions $F_1, F_2, \ldots$ respectively. Show that $P\left\{\sup_{n \geq 1} X_n < \infty\right\}$ is 0 or 1 and is in fact 1 if and only if $\sum_{1}^{\infty} (1 - F_n(x)) < \infty$ for some x. (5+15) = [20]

3.(a) Define convergence in distribution for a sequence random variables. Let X_n, $n \geq 1$, X be random variables defined on the same probability space. Show that if $X_n \xrightarrow{P} X$, then X_n converges in distribution to X. (3+7) = [10]

(b) Show that if $X_n \xrightarrow{P} X$, then there is a subsequence $\{X_{n_k}\}$ of $\{X_n\}$ such that $X_{n_k} \to X$ almost everywhere as $k \to \infty$. [10]

4.(a) Suppose that $\{X_n, n \geq 1\}$ is a sequence of independent random variables with $E(X_n) = 0$ for all n. If $\sum_{1}^{\infty} \frac{E(X_n^2)}{n^2} < \infty$, then show that $\frac{1}{n} \sum_{1}^{n} X_k$ converges to 0 almost surely. [15]

(b) $X_1, X_2, \ldots$ are independent and identically distributed random variables with $0 < E(X_1) < \infty$. Show that $\sum_{1}^{\infty} X_n = \infty$ with probability one. [5]

p.t.o.

5. μ is a probability measure on R with characteristic function $\varphi(t)$. By using arguments similar to those used in the proof of the inversion formula, show that

$$\mu\{a\} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{-ita} \varphi(t)\, dt .$$

[20]

6.(a) Prove that if a random variable X has $E|X| < \infty$, the characteristic function $\varphi(t)$ of X is differentiable and

$$\varphi'(t) = E(iX \cdot e^{itX}).$$

[5]

(b) μ_n, $n \geq 1$, is a sequence of probability measures on R with characteristic functions $\varphi_n(t)$, such that $\varphi_n(t)$ converges for every t to a limit function $\varphi(t)$, where φ is continuous at $t = 0$. Show that the sequence $\{\mu_n, n \geq 1\}$ is tight. (Hint: Observe that $\varphi(0) = \lim_{n \to \infty} \varphi_n(0) = 1$.)

[15]

INDIAN STATISTICAL INSTITUTE 1980-81: 414/526

B.Stat. (Hons.) IV / M.Stat. Previous Year: 1980-81

Modern Algebra

ANNUAL EXAMINATION

19.5.81 Maximum Marks: 100 Time: 3 hours

Note: Answer as many as you like. Maximum marks you can score is 100.

1. Let G be a group with identity element e. If G and (e) are the only subgroups of G, show that G is finite. [15]

2. Let p, q be distinct prime numbers. If a, b are elements of a group G such that order of a is p and order of b is q, show that order of ab is pq. [15]

3. Let G be a group of prime order. Show that there is an automorphism of G onto itself which is not the identity map. [20]

4. Prove that the polynomial $1 + x + x^3 + x^4$ is not irreducible over any field F, where 1 stands for the multiplicative identity of F. [15]

5. Let E be an extension of a field F and ϕ an automorphism of E onto itself leaving every element of F fixed. Let f(x) be a polynomial over F having a root a in E. Show that $\phi(a)$ is also a root of f(x). [10]

6. Let, for any prime p, $J_p = \{0, 1, \ldots, p-1\}$ with the operations addition and multiplication mod. p. Let F be a field. Show that F either has a subfield isomorphic to J_p for some p or a subfield isomorphic to the field of rational numbers (with the usual operations). [20]

7. Show that there exists an infinite field having finite characteristic. [20]

8. Let K be an extension of a field F. Let H be the set of all elements of K which are algebraic over F. Show that H is a subfield of K. [20]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV Year: 1980-81

1980-81: 454

Sequential Analysis and Nonparametric Methods

ANNUAL EXAMINATION

21.5.81 Maximum Marks: 100 Time: 3 hours

Note: All questions due to be answered. Questions carry equal marks.

1. Let $X_{(1)} \leq \cdots \leq X_{(n)}$ be the order statistics based on a random sample from a continuous population having pdf $f(x)$ and cdf $F(x)$.

(a) Find the density of $X_{(i)}$.

(b) If $f(x)$ is symmetric around 0, show that

$$E(X_{(i)}) = -E(X_{(n-i+1)}), \quad i = 1, \ldots, n,$$

provided necessary expectations exist. What will be the corresponding statement if the underlying distribution is symmetric about $\mu \neq 0$?

(c) Define ith cover $c_i = F(X_{(i)}) - F(X_{(i-1)})$, $i = 1, \ldots, n+1$ $X_{(0)} = -\infty$, $X_{(n+1)} = \infty$. Find the distribution of c_i and hence or otherwise show that $E(c_i) = {}^1/n+1$, $i = 1, \ldots, n+1$.

2.(a) Describe the Pearsonian Ch. square test and the Kolnogorov D_n test for goodness of fit. Discuss their relative merits and demerits.

(b) Assuming the parent cdf continuous show that D_n is a compldtely distribution free test statistic.

(c) Show how D_n can be used to provide a confidence internal of given confidence coefficient for the parent cdf.

3.(a) Distinguish between sign test and Wilcoxon signed rank test. State their relative merits and demerits.

(b) Show that the null distribution of Wilcoxon signed rank statistic T^+ is symmetric about its mean.

(c) Calculate the first two moment of T^+ under the null hypothesis.

p.t.o.

4.(a) Define a linear rank statistic based on two independent samples from two continuous populations. Assuming the null hypothesis, find its mean and variance.

(b) What do you mean by the symmetry of the null distribution of a linear rank statistic about a constant μ ? Give a set of sufficient condition for a linear rank statistic to be symmetrically distributed around its mean μ .

(c) Define the Wilcoxon statistic for testing the identity of two populations against location alternatives. Find a relation between this statistic and the Mann-Whitney U statistic.

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV Year: 1980-81

1980-81: 434

Advanced Linear Estimation and Inference

ANNUAL EXAMINATION

23.5.81 | Maximum Marks: 100 | Time: 3 hours

Note: Answer any five. All questions carries equal weight. Open note test.

1. Consider the game where $\Theta = \{\theta_1, \theta_2\}$, $\mathcal{A} = \{a_1, a_2\}$ and the loss is given by

$L(\theta, a)$

	a_1	a_2
θ_1	-3	2
θ_2	2	-5

A randomized strategy $\delta \in \mathcal{A}^*$ may be represented as a number q, $0 \leq q \leq 1$, with the understanding that a_1 is taken with probability q and a_2 with probability $1-q$. Draw the risk set for the game $(\Theta, \mathcal{A}^{k}, L)$. Find the minimax rule. Indicate the Bayes rules. Is the minimax rule a Bayes rule ?

2. Let $\Theta = (0, \infty)$, $\mathcal{A} = [0, \infty)$, let X be a Poisson random variable with parameter θ. Let the loss function be $L(\theta, a) = (\theta - a)^2/\theta$. Find a Minimax estimate for θ.

3. Let $X_1, \ldots, X_n$ be independent each being $N(0, \theta)$. We want to estimate θ under the loss function $[\delta(X) - \theta]^2/\theta^2$. Find an admissible minimax estimate of θ.

4. Find the best invariant estimate of θ based on a sample of size n from a uniform distribution

$$\mathcal{U}(\theta - \tfrac{1}{2}, \ \theta + \tfrac{1}{2})$$

using the loss function $L(\theta, a) = (a - \theta)^2$. Is this estimate unbiased ? If this estimate is unbiased, is it an UMVU estimate ?

5. Let X and Y be random variables with joint density

$$f_{X,Y}(x, y \mid \lambda, \mu) = \begin{cases} \lambda\mu \, e^{-\lambda x - \mu y}, & x > 0, \ y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

p.t.o.

Contd..... Q.No.5

We want to test $H_0 : \mu = \lambda$ vs $H_1 : \mu \neq \lambda$. Find the group of transformations under which problem is invariant. Find a UMP invariant size α test.

6. Let $X_1, X_2, \ldots, X_n$ be a random sample from

$$f(x \mid \theta) = \begin{cases} \dfrac{1}{1-\theta} & \theta \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the M.L.E. of θ. Find its asymptotic distribution as $n \to \infty$.

1980-81: 3E2/462/573

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year (Elective-5: Economics)/B.Stat.(Hons.) IV Year
and
M.Stat. Previous Year: 1980-81

Econometrics

SEMESTRAL II and ANNUAL EXAMINATIONS

27.5.81 Maximum Marks: 100 Time: 3 hours

Note: Answer any three questions from Group A and all questions from Group B.

Group - A

1.(a) Define the Lorenz ratio (LR) in terms of the Lorenz curve (LC) and show how LR is related to the Gini mean difference.

(b) Obtain the equation of the LC for a lognormal distribution.

(10+10) = [20]

2.(a) Define Engel elasticity of demand for a consumer item. Briefly explain some uses of information on Engel elasticities for various items of the household budget.

(b) How can one verify Engel's law given budget data for a sample of households ? Describe briefly.

(12+8) = [20]

3. How would you examine whether economies of scale in household consumption are significant or not, for each of the items in the household budget, given budget data for a sample of households ?

How do such economies of scale arise ? Are they equally important for all items ? How would you simplify the Engel relationship for an item if economies of scale are known to be negligible for that item ?

(10+10) = [20]

4. Discuss briefly any two of the undernoted problems in the context of estimation of demand functions from time series data:

(a) multicollinearity (b) aggregation (c) identification.

(10+10) = [20]

5. Write short notes on any two of the following:

(a) $\bar{R}^2$, the adjusted coefficient of multiple determination.

(b) Measurement of variables in the Cobb-Douglas production function.

(c) Examining returns to Scale through Cobb-Douglas production function.

(2 x 10) = [20]

p.t.o.

Group - B

6. The following data relate to rural areas of Punjab and are based on the 28th round of the NSS (October '73 - June '74). A few households with per capita monthly expenditure below Rs 28 have been left out.

monthly per capita expenditure (Rs.)	average expenditure per person per month (Rs.) on	
	cereals	all items
x	y	x
28 - 34	10.34	31.22
34 - 43	11.70	34.48
43 - 55	12.88	49.10
55 - 75	14.53	64.46
75 - 100	16.98	86.29
100 - 150	18.22	117.72
150 - 200	22.13	166.02
200 -	19.86	253.65

Assuming that the Engel curve for cereals has the semilog form, estimate the Engel elasticity for cereals at x = Rs.28, Rs.55 and Rs.200.

[30]

7. Practical Record. [10]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) IV Year: 1980-81

1980-81: 444

Advanced Sample Surveys

ANNUAL EXAMINATION

30.5.81 Maximum Marks: 100 Time: 3 hours

Note: Answer any three questions each carrying 25 marks.

Assignment records to be submitted at the examination hall at the start of the examinations will carry 25 marks.

1. The table below gives the data on last month's expenditures on food for 8 households of various compositions. Draw a PPSWOR sample of four households. Use the sample to unbiasedly estimate the total expenditure on food last month in these 8 households. Also obtain an unbiased estimate of the variance of your estimate.

Serial No. of household	Household size	Last month's expenditure on food (in 100 Rupees)
1	7	17.9
2	2	8.6
3	4	14.3
4	5	21.6
5	6	14.0
6	1	3.1
7	2	5.4
8	5	14.9

2.(a) Obtain the Yates - Grundy form of the variance of Horvitz - Thompson estimator for a finite population total based on a fixed sample-size design.

(b) Find an unbiased estimator for the variance of Horvits - Thompson estimator for a finite population total based on the Poisson scheme of sampling. Indicate how the variance in this case may be unbiasedly estimated even from a sample of effective size one.

3.(a) Describe the circular pps-systematic method of sampling a finite population with inclusion-probabilities proportional to given normed size - measures.

(b) Using the data based on a PPSWR sample chosen in n draws from a population of N units show how you may unbiasedly estimate the gain in precision of PPSWR method over a comparable SRSWOR method of sampling in n draws, the purpose being to estimate the population mean in either case.

p.t.o.

4.(a) Explain what you mean by a self-weighting design and indicate its use.

If fsu's are selected by stratified PPSWR method and selected fsu's are sub-sampled independently by SPSWOR method illustrate how you may choose a self-weighting design.

(b) A population of size N consists of L (≥ 2) well-defined strata. A large SPSWOR of size n' is chosen from it with n'_h's, the numbers of units in it falling in the various strata, at least quater then 2 for every stratum h (= 1, ..., 2). From this first phase sample a stratified simple random (WOR) second phase sample is chosen with sample-sizes $n_h = \nu_h n'_h$'s (with ν_h pre-assigned, $0 < \nu_h < 1, \forall h$) from the units observed to fall in the respective strata out of the chosen first phase sample. Show how you may use the data to unbiasedly estimate the population mean if the strata-sizes N_h's (h = 1, ..., L) be initially unknown. Obtain a neat expression for the variance of your estimate.

5. The data below gives the composition of 10 households along with ages of the inmates.

Serial no. of household	Ages of household inmates (l.b.d.)
1	52, 45, 13, 10
2	35, 27. 3
3	68, 57, 34, 27, 24, 21, 17
4	41, 36, 63
5	33, 29, 1
6	77, 38, 35, 2
7	27, 25, 0
8	43, 39, 6, 0
9	65, 57, 37, 31, 4, 1
10	56, 24, 21, 17, 11

Draw an SRSWOR of 4 households and from each selected household independently take SRSWOR's of 2 inmates and record their ages only. From such sample-data obtain an appropriate estimate of the average age of all the members of these 10 households. Also give an estimate of the variance of your estimate.

1980-81: 474/574

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) IV Year and M.Stat. Previous Year: 1980-81

Bio-statistics

ANNUAL EXAMINATIONS

1.6.81 Maximum Marks: 100 Time: 3 hours

Note: The paper carries 114 marks. You can answer any part of any question. Maximum marks you can score is 100.

1. Suppose we have a pair of alleles A, a with A completely dominant to a and a random sample of N dominant individuals. We have n progenies by selfing from each individual of the sample and find that D have no recessive progeny and H have one recessive progeny or more. The problem is to estimate the proportion of homozygous dominants in the population. Set up the maximum likelihood equation, and derive the estimator and its variance.

[20]

2. Describe the Proband method and sib method for estimating the segregating parameter θ in a human population. Indicate a method for obtaining maximum likelihood estimator of θ under ascertainment through affected children.

[22]

3. Define the terms 'inbreeding coefficient' and 'coefficient of parentage'. If F_n denotes the common inbreeding coefficient in generation n, obtain an expression for F_n in the following two cases:

   (i) full sibbing in each generation,

   (ii) perfect random mating, with the population size constant in each generation.

4. Suppose we have two dominant factors A, a and B, b and obtain an F_2 from a coupling double heterozygote cross. Assume the recombination fraction in male gametogenesis is p_1 and in female gametogenesis is p_2. Obtain the maximum likelihood estimator of $P = (1-p_1)(1-p_2)$ and its variance. Assuming in addition $p_1 = p_2 = p$, find an estimator of p and its variance.

[20]

5. The data in the accompanying table represent the sample distributions of the A-B-O blood groups among controls, and among stomal and duodemal ulcer patients in a certain population, in a given period.

p.t.o.

Contd..... Q.No.5

Blood Group / Sample	O	A	B	AB	Total
Control	4578	4219	890	313	10,000
Stomal ulcer	181	96	18	5	300
Duodenal ulcer	298	214	39	13	564

(i) Is there any evidence that stomal ulcer is associated with O blood group ?

(ii) Is there any evidence that duodenal ulcer is associated with O blood group ?

(iii) Is there any evidence that susceptibility to stomal ulcer is lower in individuals of the A blood group than in those of the B or AB blood group ?

[30]