INDIAN STATISTICAL INSTITUTE

M.Stat.(M-stream) I Year : 1987-88

DESIGN OF EXPERIMENTS I

Semestral-II Backpaper Examination

Date : 1.7.88. Maximum Marks : 100 Time : 3 Hours

Note : Answer any five questions.

1.(a) Explain in brief the role of randomisation and replication in Design of Experiments, giving suitable examples.

(b) What is a uniformity trial ? What is its role in the context of agricultural field trials ? [12+8] = [20]

2. If N(v) represents the maximum number of mutually orthogonal latin squares of order v, show that

(i) $N(v) = v-1$, if v is a prime number or a prime power

(ii) $N(v) \le v-1$, otherwise. [20]

3. Construct a balanced 2^5 designs in 5 replicates, each replicate being laid out in 4 blocks of size 8 each. All main effects and two factor interactions should remain unconfounded in the design. Explain the method neatly and write down the principal block for each of the replicates. [20]

4.(a) A 2^4 design is to be laid out in a 4x4 square. Apparently three factorial effects are to be confounded with rows and three with columns. Construct a design in which no main effects are confounded and exactly two 2-factor interactions are confounded.

(b) Construct a $\frac{1}{8}$th replicate of a 2^7 design in which no main effect has as an alias another main effect or a two factor interaction. [10+10]=[20]

5.(a) What is a BIB Design ?

(b) Prove that in a BIB Design with $v = nk$ i.e. $b = nr$ (where n is a positive integer), $b \ge v+r-1$.

(c) What is a resolvable BIB Design ? What is an affine resolvable BIB Design ? Give an example of an affine resolvable BIB Design.

(d) Show that for an affine resolvable BIB Design v/k^2.

[2+5+5+8]=[20]

6. Write short notes on :

(a) Missing plot technique

(b) Replicated Latin Square Design

(c) Split plot design.

(d) Size and shape of plots and blocks. [5x4] = [20]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

ANALYSIS II
Semestral-II Backpaper Examination

Date : 29.6.88. Maximum Marks : 100 Time : 3 Hours

Note : Answer all questions.

1.(a) Show that any n-dimensional normed linear space is linearly homeomorphic to $\mathbb{C}^n$.

Hence show that any finite-dimensional normed linear space is a Banach space.

Also show that any two norms on a finite-dimensional normed linear space are equivalent.

(b) Show that an infinite dimensional Banach space cannot have a countable basis. [12+8]

2.(a) Prove that a normed linear space X is separable iff there exists a countable subset of X whose span is dense in X.

Hence show that if every orthonormal set in a Hilbert space H is countable, then H is separable.

(b) Deduce the following from the Closed Graph Theorem :

Any one-to-one bounded linear transformation of a Banach space X onto a Banach space Y is a homeomorphism. [12+8]

3.(a) Let X and Y be two normed linear spaces and T a linear transformation of X onto Y. Then prove that T^{-1} exists and is bounded iff there exists m > 0 such that $\|Tx\| \geq m\|x\|$, $\forall x \in X$.

(b) Let H be a Hilbert space. Show that if $x_n \to x$ and $y_n \to y$ in H, then $(x_n, y_n) \to (x, y)$.

(c) Let M be a linear subspace of a Hilbert space H. Show that M is closed iff $M = M^{\perp\perp}$. [7+6+7]

4.(a) Let $\{e_i : i \in I$ be an orthonormal set in a Hilbert space H. Show that for any $x \in H$ the set $S = \{e_i : (x, e_i) \neq 0\}$ is countable.

(b) Let H be a Hilbert space. Show that the adjoint operation $T \mapsto T^*$ on $\mathcal{B}(H)$ is one-to-one, onto and norm preserving.

p.t.o.

4.(c) An operator T on H is said to be normal iff $T.T^* = T^*T$.

Show that T is normal iff $\|T^*x\| = \|Tx\|$, $\forall x \in H$.

[7+7+6]

5.(a) For any non-empty set S, let $\ell_2(S)$ be the set of all complex functions f on S such that $\{s \in S : f(s) \neq 0\}$ is countable and $\Sigma|f(s)|^2 < \infty$. With pointwise addition and scalar multiplication and inner product defined by

$$(f,g) = \Sigma f(s)\overline{g(s)},$$

$\ell_2(S)$ is a Hilbert space.

Now, suppose $S = \{e_i : i \in I\}$ is a complete orthonormal set in a Hilbert space H. For each $x \in H$, define a function φ_x on S by

$$\varphi_x(e_i) = (x, e_i).$$

Show that $\varphi_x \in \ell_2(S)$. Also show that the map $x \longrightarrow \varphi_x$ is an isometric isomorphism of H onto $\ell_2(S)$.

(b) For any two vectors x and y in a Hilbert space H show that $|(x,y)| \leq \|x\| . \|y\|$.

(c) If M and N are closed linear subspaces of a Hilbert space H, then prove that M+N is closed.

[10+5+5]

1987-88|512(b)

INDIAN STATISTICAL INSTITUTE

M.Stat.(M-stream) I Year : 1987-88

PROBABILITY II

SEMESTRAL-II BACKPAPER EXAMINATION

Date : 27.6. 88. Maximum Marks : 100 Time : 3 Hours

Note : Marks for each question are indicated in [ ]. Answer all questions.

1. Let λ stand for Lebesgue measure on the unit interval. Fix $\varepsilon > 0$. Show that given any Borel subset B of the unit interval, there is an open set U such that $B \subseteq U$ and $\lambda(U - B) < \varepsilon$. [15]

2. Let $\{f_n\}$, $\{g_n\}$, $\{h_n\}$ be three sequences of integrable functions on a measure space $(\Omega, \mathcal{A}, \mu)$

   such that $f_n \leq h_n \leq g_n$ a.e.(μ)

   and that $f_n \to f$ a.e.(μ)

   $h_n \to h$ a.e.(μ)

   $g_n \to g$ a.e.(μ)

   Assume that $\int f_n \, d\mu \to \int f d\mu$ and $\int g_n d\mu \to \int g d\mu$.

   Show that $\int h_n d\mu \to \int h d\mu$ [12]

3. (a) State the Borel Cantelli lemma. [5]

   (b) Consider an independent sequence of tosses with a fair coin and ℓ_n denote the run length of heads starting from the n^{th} trial.

   Show that $P(\{w : \limsup_n \frac{\ell_n(w)}{\log_2 n} \geq 1\}) = 1$. [20]

4. State and prove the Central Limit Theorem for a sequence of mutually independent random variables.
   (You may use without proof all the facts on characteristic functions and weak convergence that you need). [30]

5. State and prove Kolmogorov's maximal inequality. [12]

6. Construct random variables $\{X_n\}_{n\geq 1}$ and X such that $X_n \Rightarrow X$ but X_n does not converge to X in probability. [6]

1987-88 | 512

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

PROBABILITY II
Semestral II Examination

Date : 13.5.88. Maximum Marks : 100 Time : $3\frac{1}{2}$ Hours

Note : Marks for each question are indicated in [ ].
Answer all questions.

1. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $\{f_n\}_{n\geq 1}$ a sequence of $\mathcal{A}$-measurable functions on Ω. Show that if $\{f_n\}$ is uniformly integrable then $\{f_n\}$ is uniformly absolutely continuous. [10]

2.(a) State the Borel-Cantelli Lemma. [5]

(b) Consider a sequence of tosses with a fair coin. Let ℓ_n stand for the run length of heads starting from the n^{th} trial, i.e., for a sample point w,

$\ell_n(w) = k$ iff trials numbered n to n+k-1 result in Heads and the $(n+k)^{th}$ trial results in Tail.

Show that

$$P\left(\left\{w : \limsup_n \frac{\ell_n(w)}{\log_2 n} \leq 1\right\}\right) = 1.$$ [15]

3.(a) Define the weak convergence of a sequence of distribution functions. [5]

(b) Construct random variables $\{X_n\}_{n\geq 1}$ and X such that $X_n \Rightarrow X$ but X_n does not converge to X almost surely. [7]

4. Let $\{\mu_n\}_{n\geq 1}$ be a tight sequence of probability measures and μ a probability measure such that every subsequence of $\{\mu_n\}_{n\geq 1}$ which converges weakly to a probability measure, converges to μ. Show that $\mu_n \Rightarrow \mu$. [22]

5.(a) Let $\{X_n\}_{n\geq 1}$, X be random variables with characteristic functions $\{\varphi_n\}_{n\geq 1}$ and φ respectively. Show that $X_n \Rightarrow X$ implies $\varphi_n(t) \to \varphi(t)$ for every real t. [7]

(b) Let $\{X_n\}_{n\geq 1}$, X be random variables such that $X_n \Rightarrow X$. Show that

$$E(|X|) \leq \liminf_n E(|X_n|).$$ [8]

p.t.o.

6.(a) State the Central Limit Theorem for a sequence of independent random variables. [6]

(b) Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables with mean 0 and let $\delta > 0$. Assume

$$\lim_{n} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} E(|X_k|^{2+\delta}) = 0 ,$$

where $s_n^2 = \mathrm{Var}\,(X_1+\ldots+X_n)$.

Show that $\dfrac{X_1+\ldots+X_n}{s_n} \Rightarrow N(0,1)$. [8]

(c) Suppose $\{X_n\}$ is a sequence of independent random variables with mean 0 and $K > 0$ is a constant such that $P(|X_n|\leq K)=1$ for each n. Show that if $s_n \to \infty$, then

$$\frac{S_n}{s_n} \Rightarrow N(0,1).$$ [7]

INDIAN STATISTICAL INSTITUTE
M.Stat.(S- and M-stream) I Year : 1987-88

OPTIMIZATION TECHNIQUES I
Semestral II Examination

Date : 11.5.88. Maximum Marks : 100 Time : $3\frac{1}{2}$ Hours

Note : This paper carries a total of 110 marks. Answer as many questions or parts thereof as you can. If your score exceeds 100 marks, your final score will be treated as 100.

1.(a) Show that $(1, 1, \frac{1}{2}, 0)$ is an optimal solution of the LP problem :

$$\text{Maximise } Z = 2x_1 + 4x_2 + x_3 + x_4$$
$$\text{subject to } \quad x_1 + 3x_2 + x_4 \le 4,$$
$$2x_1 + x_2 \le 3,$$
$$x_2 + 4x_3 + x_4 \le 3,$$
$$x_i \ge 0, \quad 1 \le i \le 4.$$

(b) Find all the optimal solutions of this problem.

[10+10 = 20]

2.(a) In the network N_1 of Figure 1 with the unique production centre x and destination y,

(i) determine all the minimal cuts,

(ii) determine all the integral maximum flows.

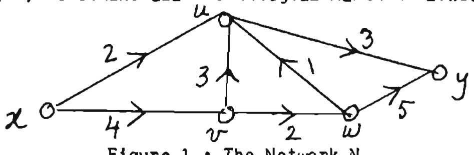


Figure 1 : The Network N_1.

(b) Prove that every m x n Latin rectangle based on n symbols with $n \ge m$ can be extended to a Latin square of order n based on the same set of n symbols.

[(8+5)+7 = 20]

p.t.o.

3.(a) In the network N_2 of Figure 2 prove that there is no feasible flow with the supplies $s(x_1) = 7$, $s(x_2) = 2$ and with the demands $d(y_1) = 1$, $d(y_2) = 8$.

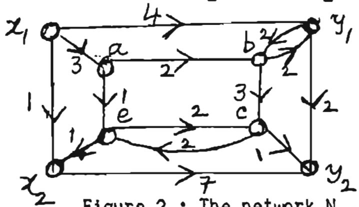


Figure 2 : The network N_2.

(b) Find a feasible flow in the network N_2^1 where N_2^1 is obtained from N_2 by changing only the capacity of the link (c, y_2) to 2; with the same supplies and demands.

[10+10 = 20]

4.(a) Find the value of the two person zero-sum game with pay-off matrix A of order n, where A is the diagonal matrix with entries $d_1, \ldots, d_n$ $(d_i > 0)$)

(b) Prove or disprove that there is a (0,1)-matrix A with the row-sum vector $s = (6,6,4,2,2,2)$ and the column-sum vector $d = (5,4,4,4,3,1,1)$.

(c) Find the minimum number of lines of the matrix which together cover all the positive entries of the matrix A. Justify your answer.

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 10 & 0 \\ 2 & 0 & 0 & 1 & 3 & 0 \\ 4 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 10 & 0 & 3 & 1 & 0 & 0 \end{pmatrix}$$

[10+10+10 =30]

5.(a) Formulate the flow problem in a network as an LP problem and write down its dual.

(b) State and prove Hall's theorem on the system of distinct representatives for a finite family of finite sets.

[8+12 = 20]

1987-88 | 562

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

SAMPLE SURVEYS I
Semestral II Examination

Date : 9.5.88. Maximum Marks : 100 Time : 3 Hours

Note : Do as many questions as you can. The paper carries 110 marks but the maximum you score can not exceed 100 marks. Marks allotted are given at the end of each question.

1. Prove or disprove the following statements :

(i) The sampling strategy T_0 = (SRSWOR, sample mean) is the best unbiased strategy for $\bar{Y}$ in the classes $T_1 \equiv (\text{SRSWOR}, \sum_{r=1}^{n} a_r y_r)$ and $T_2 \equiv (\text{SRSWOR}, \sum_{i=1}^{n} \beta(i) y_i)$, the coefficients a_r and $\beta(i)$ depending on order and identity of units respectively.

(ii) For a given sampling design, Horvitz-Thompson estimator is admissible in the unbiased subclass $T_5 = \sum_{i \in s} \beta(i,s) y_i$.

(iii) If $\rho < 0.7$ and $c_x/c_y > 1.5$ then ratio estimator would be better than sample mean (both based on SRSWOR).

(iv) If $y_i = \alpha + \beta x_i$ ($i = 1,..N$) then the sampling strategy $T = (\text{PPXWR}, \frac{\bar{X}}{n} \sum_{i=1}^{n} \frac{y_i}{x_i})$ would be better than the strategy T_0 = (SRSWR, sample mean) for estimating population mean $\bar{Y}$.

(v) Ratio estimator is unbiased under Midzuno-Sen sampling scheme.

(10+10+3+7+5) = [35]

2. (i) Describe double sampling for regression method of estimation for $\bar{Y}$.

(ii) Show that under a linear cost function

$$C = C_0 + nC_1 + mC_2$$

the sampling strategy based on double sampling and regression estimator is better than the strategy based on a single simple random sample and sample mean provided

$$\rho^2 > 4(C_1/C_2)/[1+C_1/C_2]^2.$$

(5+15) = [20]

p.t.o.

3.(i) Describe multi-stage sampling procedure.

(ii) A rural Block consisting of 120 villages was divided into 10 Areas having 12 villages each. A SRSWOR of 3 Areas was selected and then 4 villages were selected from each of the selected Areas using SRSWOR. The following table gives the no. of land-holdings of size less than 5 acres for each of the selected villages.

Sampled Areas Sr.No.	no. of land holdings (less than 5 acres) for sampled villages			
A_1	50,	20,	40,	70
A_2	15,	30,	18,	40
A_3	100,	25,	70,	45

(a) Give unbiased estimates for total no. of land holdings of size less than 5 acres in the entire Block; average no. of holdings per Area; average no. of holdings per village.

(b) Give unbiased estimates for the variances of the above estimates. (5+20) = [25]

4.(i) What do you mean by sampling and non-sampling errors.

(ii) Name the major types of non-sampling errors and indicate the sources of their occurance.

(iii) Describe Hansen-Hurwitz technique of estimating population mean unbiasedly in the presence of non-response.

(5+5+10)=[20]

5. Practical record note books and Home-assignments. [10]

1987-88 | 482/542

INDIAN STATISTICAL INSTITUTE

M.Stat. (S- and M-stream) I Year : 1987-88

ANALYSIS II

Semestral II Examination

Date : 9.5.88. Maximum Marks : 100 Time : 3 Hours

Note : Unless otherwise stated, all the linear spaces considered below are over the complex field.

ANSWER ANY FIVE.

1.(a) Let X be a normed linear space. Show that X is a Banach space iff every absolutely convergent series is convergent.

(b) Suppose $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on a linear space X. Show that these norms are equivalent iff there exist positive real numbers K_1 and K_2 such that $K_1\|x\|_1 \leq \|x\|_2 \leq K_2\|x\|_1$ for all $x \in X$.

(c) Let $\{T_\alpha : \alpha \in \Lambda\}$ be a collection of bounded linear transformations of a Banach space B into a normed linear space X. Suppose that for each $x \in B$, $\{T_\alpha x : \alpha \in \Lambda\}$ is bounded in X. Then prove that $\{\|T_\alpha\| : \alpha \in \Lambda\}$ is bounded. (6+7+7)

2.(a) Let f and g be complex (Lebesgue) measurable functions on [0,1] and let p,q (>1) be real numbers such that $\frac{1}{p} + \frac{1}{q} = 1$. Show that

$$\int_0^1 |f(x)g(x)|dx \leq \left\{\int_0^1 |f(x)|^p dx\right\}^{1/p} \left\{\int_0^1 |g(x)|^q dx\right\}^{1/q}$$

(b) Define $L_p[0,1]$ for $1 \leq p < \infty$ and show that it is a Banach Space. (7+13)

3.(a) Let B be a Banach space and Y a normed linear space. Suppose $T : B \to Y$ is a one-one, bounded linear transformation such that its range T(B) is non-meager in Y. Show that T is onto and a homeomorphism.

(b) If M is a closed linear subspace of a Hilbert space H, then show that M contains a unique vector of smallest norm. (13+7)

p.t.o.

4.(a) Define a complete orthonormal set in a Hilbert space. Let $\{e_i : 1 \leq i \leq n\}$ be a finite orthonormal set in a Hilbert space H. Show that for any vector x in H

$$\sum_{i=1}^{x} |(x,e_i)|^2 \leq \|x\|^2.$$

Hence, show that if $\{e_i : i \in I\}$ is an arbitrary orthonormal set then $\Sigma|(x,e_i)|^2$ is well-defined and $\Sigma|(x,e_i)|^2 \leq \|x\|^2$

(b) Show that the functions e_n, $n = 0, \pm 1, \pm 2, \ldots$ defined by

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$$

form an orthonormal set in $L_2[0,2\pi]$.

Hence deduce that for any function $f \in L_2[0,2\pi]$

$$\sum_{n=-\infty}^{\infty} |c_n|^2 \leq \int_0^{2\pi} |f(x)|^2 \, dx$$

where, $c_n = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) \, e^{-inx} \, dx.$

(12+8)

5.(a) Let H be a Hilbert space. Show that for each $f \in H^*$ there is a unique vector y in H such that

$$f(x) = (x,y), \ \forall x \in H.$$

(b) Define the adjoint of a bounded linear operator on a Hilbert space H.

Show that self-adjoint operators in $\mathcal{B}(H)$ form a real Banach space.

(c) Let T be an operator on a Hilbert space H such that $(Tx,x)=$- for all x. Show that $T \equiv 0$.

Hence, show that $T \in \mathcal{B}(H)$ is self-adjoint iff (Tx,x) is real for all $x \in H$.

(7+6+7)

6.(a) Let P be a projection on a Hilbert space H with range M and null space N. Show that $M \perp N$ iff P is self-adjoint. In that case, prove that $N = M^\perp$.

(b) Show that for any arbitrary operator T on a Hilbert space H, $I + TT^*$ is non-singular, where I is the identity operator.

(c) Let M be a closed linear subspace of a Hilbert space H. Show that M is invariant under an operator T iff $M^\perp$ is invariant under T^*.

(6+9+5)

INDIAN STATISTICAL INSTITUTE
M.Stat. (M-stream) I Year : 1987-88

STATISTICAL METHODS AND INFERENCE II (THEORY)
Semestral II Examination

Date : 6.5.88. Maximum Marks : 100 Time : 3 Hours

Note : Assumptions made or results used must be stated clearly. Unless otherwise stated, probability (density) function to be taken as $f(.,\theta)$, θ is a real parameter.

Answer any five questions, each carrying 18 marks, 10 marks for assignments.

1. State and prove Neyman-Pearson Lemma for testing a simple hypothesis against a simple alternative.
2. Let $X_1, X_2, \ldots X_n$ be i.i.d. $N(\theta,1)$. Find UMP test of $H_0:\theta \leq 0$ against $H_1: \theta > 0$. Write the power function of the UMP test.
3. (a) Define monotone likelihood ratio (MLR) family of distributions.

   (b) Show that, if the distribution of X has MLR, any test of the form

$$\varphi(x) = \begin{cases} 1 & \text{if } x > x_0 \\ \nu & \text{if } x = x_0 \\ 0 & \text{if } x < x_0 \end{cases}$$

   has monotone power function for testing $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$.
4. Let Θ_B denote the set of all θ in the boundary of H_0 and H_1.

   (a) Show that, if every test of H_0 against H_1 has continuous power function, and if φ is size α, UMP, α-similar on Θ_B, then φ is UMP unbiased.

   (b) If T is boundedly complete sufficient statistic for θ in Θ_B, then show that every test similar on Θ_B has Neyman structure.
5. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$. Find UMP unbiased test of $H_0: \mu \leq 0$ against $H_1: \mu > 0$ when σ^2 is unknown.
6. Find UMP unbiased test of $H_0: \lambda \leq \mu$ against $H_1: \lambda > \mu$ based on X, which is distributed as Poisson with mean λ, and Y, which is distributed independently as Poisson with mean μ.

p.t.o.

7. Find a test, based on likelihood ratio principle, of $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ based on a sample of n observations from $N(\mu_1, \sigma^2)$ and m observations from $N(\mu_2, \sigma^2)$ where the observations are mutually independent.

8. (a) What is a locally best test for testing $H_0: \theta \leq 0$ against $H_1: \theta > 0$, and how can you construct it using Neyman-Pearson Lemma ?

(b) Find the form of a locally best test of H_0 against H_1 where θ is the median of a Cauchy distribution.

9. A sample of n observations are available. Describe how you would test that the observations are from $N(\mu, \sigma^2)$ where both μ and σ are known, using (a) χ^2 goodness of fit and (b) Kolmogorov-Smirnov one-sample test. Which one of these two would you like to use and why ?

INDIAN STATISTICAL INSTITUTE
M.Stat. (M-stream) I Year : 1987-88

STATISTICAL METHODS AND INFERENCE II (PRACTICAL)
Semestral II Examination

Date : 6.5.88. Maximum Marks : 100 Time : 3 Hours

Note : Write clearly your assumptions and hypothesis, test procedure, computed values of test statistics, table values at 5% level and your inference.

Answer any three questions, each carry 30 marks, 10 marks for practical records.

1. Based on 250 observations of scores in a certain test, assumed to follow normal distribution with mean μ and known variance $\sigma_0^2 = 81$, sample mean was observed to be 54.2.

(a) Test $H_0: \mu \leq 55$ against $H_1: \mu > 55$.

(b) Obtain a (two sided) 95% confidence interval for μ.

2. The following gives the counts of the number of European red mites on 150 randomly selected apple leaves.

Number of mites per leaf	Number of leaves observed
0	70
1	38
2	17
3	10
4	9
5	3
6	2
7	1
8 or more	0
Total	150

Test if the number of mites on a leaf follows a Poisson distribution.

3. Scores of 15 students in Statistics (S) and Mathematics (M) in a certain examination are

(S,M) = (17,41), (56,85), (45,42),
(82,71), (65,46), (61,92),
(31,42), (54,54), (57,69),
(58,52), (58,41), (29,17),
(84,73), (93,91), (60,70).

Using signed-ranks test, decide if the students' Math. and Stat. abilities are same.

p.t.o.

4. Following gives the heart weight in grams of 12 female and 15 male cats. Does the heart of a male cat weigh more than that of a female cat ? Use run test to decide.

Heart-weight in grams

Males: 12.7, 5.6, 9.1, 7.6, 12.8, 8.3, 11.2, 9.3, 9.4, 8.0, 14.9, 10.7, 13.6, 9.6, 11.7.

Females: 7.4, 7.3, 17.1, 9.0, 7.6, 9.5, 10.1, 10.2, 10.1, 9.5, 8.7, 7.2.

5. A random sample of persons were classified by profession and state of origin as given below.

Profession	State of origin Bengal	Bihar	Assam	Punjab
Lawyer	53	12	19	15
Doctor	28	52	23	19
Engineer	64	23	72	21
Teacher	14	25	33	38

Are the professions independent of states of origin ?

6. Following 11 three-digit numbers were taken, from a random number table, for sample selection.

128, 319, 069, 951, 811, 314,
216, 680, 403, 385, 146.

Using Kolmogorov-Smirnov test procedure, test if these numbers are actually random.

1987-88 | 572

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

DESIGN OF EXPERIMENTS
Semestral II Examination

Date : 4.5.88. Maximum Marks : 100 Time : $3\frac{1}{2}$ Hours

GROUP A 1 Hr. 30 Minutes.

Note : Answer any three questions. All questions in this group carry equal marks.

1. What is a resolvable BIB Design ? Prove that for a resolvable BIB Design $b \geq v + r-1$.
   Show that for a resolvable BIB Design with $b = v+r-1$, the intersection number between two distinct blocks is either 0 or a constant c. Find c in terms of the parameters of the Design. Give an example of such a design. [24]

2. The observation corresponding to Block 1 and treatment 1 is missing in a Randomised Block Design with v treatments and r blocks.
   Describe in brief how you will analyse the data. Find the expressions for the variances of the BLUE's of (i) $\tau_1 - \tau_2$ and (ii) $\tau_2 - \tau_3$. How will you test the hypotheses (i) H_0 ($\tau_1 = \tau_2$) and (ii) H_0 ($\tau_2 = \tau_3$) ? [24]

3.(a) What is a factorial experiment ? Enumerate the advantages of a factorial experiment in comparison with a series of experiments with one factor at a time.
  (b) Explain the concept of confounding and partial confounding in the context of 2^n experiments, giving examples. Describe what you mean by balancing in a factorial experiment. Give an example of a balanced 2^3 experiment in 8 blocks, each of size 4.
  (c) Construct a 2^6 Design in 4 blocks of size 16 each, in which no main effects and no two factor and no three factor interactions are confounded. [8+8+8] = [24]

4.(a) A half replicate of a 2^7 design is to be laid out in 4 blocks of 16 plots each, in which all main effects and all two factor interactions remain orthogonally estimable. Write down the method of construction clearly and the composition of the principal block in full.

p.t.o.

4.(b) Construct a $\frac{1}{16}$th replicate of a 2^7 design in which all main effects are orthogonally estimable, assuming all interactions of all orders to be absent. [12+12]=[24]

GROUP B 2 Hours

5. The table below gives the yields and the layout of a 2^5 factorial experiment on beans. A single replication in 4 blocks of 8 plots each was used. The five factors and their levels are as follows :

(i) Spacing of rows (S) at 18 inches or 24 inches

(ii)Dung (D) at none of 10 tons per acre

(iii)Nitrochalk (N) at none or 0.4 cwt per acre.

(iv)Super phosphate (P) at none of 0.6 cwt per acre.

(v) Muriate of Potash (K) at none or 1.0 cwt per acre.

As usual, absence of a symbol denotes the lower level of the corresponding factor and presence of the symbol denotes its higher level.

Block 1				Block 2			
s	36.2	sdp	49.8	n	68.0	k	63.6
snk	60.5	dk	51.3	sdnk	92.5	dpk	63.6
np	36.3	sdnpk	61.3	sp	29.9	dnp	60.8
dn	67.3	pk	49.6	sd	54.7	snpk	47.0

Block 3				Block 4			
nk	71.2	(1)	66.5	p	56.7	npk	48.0
snp	45.7	sdn	70.5	sdnp	64.6	sn	23.5
dp	76.7	spk	74.3	d	74.8	sk	39.3
sdk	73.7	dnpk	77.0	dnk	73.7	sdpk	56.3

Analyse the data. [28]

1987-88 532

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

MULTIVARIATE ANALYSIS IA
Semestral II Examination

Date : 2.5.88. Maximum Marks : 100 Time : 3 Hours

Note : Answer as many as you can.

1. Define p-dimensional normal variable. If (x_1,x_2) and (y_1,y_2) are two independently and identically distributed bivariate normal variables with mean vector null and dispersion matrix I, obtain the conditional distribution of

$$(x_1, x_2, y_1, y_2) / (x_1 = y_1).$$ [10]

2. Define Wishart variable. If $W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$ has Wishart distribution, obtain the distribution of $w_{22.1} = w_{22} - w_{21}w_{11}^{-1} w_{12}$. [10]

3. Given a random sample of size n from a p-variate normal population with mean vector μ and dispersion matrix Σ, derive a test for testing the hypotheses $A\mu = b$. [10]

4.(a) Describe multivariate Gauss-Markoff Model $(\underset{n\times p}{Y} , XB, \Sigma \otimes I)$.

(b) Describe briefly (without derivations and proofs) how you would test the hypotheses $H'\beta i = \xi i$(given) for $i=1,\ldots,p$ where βi is the i^{th} column of B. State clearly the assumptions under which you can test. [5+10=15]

5. The following table gives the estimates of the means and the common dispersion matrix of three characters (x_1,x_2,x_3) for two groups of female desert locusts in two different phases P_1 and P_2 based on samples of sizes 20 and 72 respectively.

Character	Means P_1 $n_1=20$	Means P_2 $n_2=72$	Dispersion matrix based on 90 d.f. x_1	x_2	x_3
x_1	25.80	28.35	4.7350	0.5622	1.4685
x_2	7.81	7.41	0.5622	0.1431	0.2174
x_3	10.77	10.75	1.4685	0.2174	0.5702

(a) Find the linear discriminant function based on the three characters $(x_1\ x_2\ x_3)$ between the two groups of locusts.

p.t.o.

5.(b) Test for the significance of the difference in the mean vectors of the two groups.

(c) Examine whether it is worthwhile to include x_3 in addition to x_1 and x_2 for purposes of discrimination between the two groups.

(d) Are the data given consistent with the hypothesis that the linear discriminant function between the two groups is

$$y = -3x_1 + 7x_2 + 5x_3 \quad ?$$

(e) To which group would you assign a locust with measurements $x_1 = 27.06$, $x_2 = 8.03$ and $x_3 = 11.36$? [10+5+8+7+5 = 35]

6. Write short notes on

(a) Principal Components.

(b) Factor Analysis.

(c) Canonical Correlations and canonical variables.

[3x5=15]

7. Assignments. [20]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

SAMPLE SURVEYS I

Periodical Examination

Date : 4.3.88. Maximum Marks : 100 Time : 3 Hours

Note : Do as many questions as you can. The paper carries 110 marks but the maximum you score can not exceed 100 marks. The marks allotted are given at the end of each question.

1.(i) Define proportional and optimum allocations.

(ii) Consider the following strategies for estimating population mean $\bar{Y}$

$T_0 \equiv$ (SRSWOR, sample mean); T_p = (SSS with p-allocation $\bar{y}_{st}$)

$T_N \equiv$ (SSS with N-allocation, $\bar{y}_{st}$)

Show that for large populations

$V(T_N) \leq V(T_p) \leq V(T_0)$. (8+12)= [20]

2.(i) Define ratio, regression, difference and product estimators, for population mean, based on SRSWOR.

(ii) Derive large sample approximation to the bias and MSE of the ratio estimator defined above.

(iii) Identify the situations in which above estimators are better than sample mean. (5+12+8) =[25]

3.(i) What do you mean by PPS - sampling ?

(ii) The following table gives area under wheat crop (x) and the wheat yield (y) for 5-villages

Village No.	1	2	3	4	5
x (In hectares)	40.1	305	120.5	160	208
y (In quintals)	417.8	3660	1807	190.8	2275

(a) Draw a sample of 3 villages with probability proportional to area under the crop and with replacement.

(b) Obtain an unbiased estimate of the yield for all the villages and obtain the estimated standard error of the estimate. (5+15) = [20]

p.t.o.

4.(i) Define Horvitz-Thompson estimator for general sampling designs and show that it is UMVU in the unbiased subclass of T_2 - class of linear estimators.

(ii) Let $a(i)$, $i = 1, \ldots N$, be the coefficient associated with i-th unit of the population. Then show that in the class of unbiased sampling strategies

$$T \equiv (\text{SRSWOR}, \sum_{i=1}^{n} a(i)\, y_i) \text{ for population mean, the}$$

strategy $T_0 \equiv$ (SRSWOR, sample mean) is UMVU.

(iii) Let n draws be made according to SRSWOR and y_r denote the y-value for the unit selected at r-th draw. Show that sample mean is the best in the class of linear unbiased estimators $T = \sum_{r=1}^{n} a(r)\, y_r$ for population mean.

(10+5+10)=[25]

5. The fifty villages in a rural block were divided into two strata consisting of 32 and 18 villages respectively. It was decided to select an overall sample of twenty percent of the villages according to proportional allocation. The number of 'poor' - households in the selected villages were found as follows :

Stratum No.	No. of poor households
1	20, 50, 15, 75, 18, 25
2	30, 10, 15, 12

(a) Estimate the average number of poor-households in a village of the block.

(b) Obtain estimated standard error of the estimate.

(c) Obtain an estimate of the gain due to stratification.

(5+5+10) = [20]

INDIAN STATISTICAL INSTITUTE
M.Stat.(S- and M-stream) I Year : 1987-88

ANALYSIS II
Periodical Examination

Date : 4.3.88. Maximum Marks : 100 Time : 3 Hours

Note : Answer as many as you can. The maximum you can score is 100.
Unless otherwise stated, X below denotes a metric space.

1. Let d_1 and d_2 be two metrics on X. Show that d_1 and d_2 are equivalent iff for any sequence $\{x_n\}$ and any point x in X,

$$[\ d_1(x_n,x) \to 0\ ] \text{ iff } [\ d_2(x_n, x) \to 0\ ]$$

Show by means of an example that in a metric space the closure of the open sphere $S_r(x)$ need not be the closed sphere $\bar{S}_r(x)$.

Let $A \subseteq X$. Show that $x \in \bar{A}$ iff there is a sequence $\{x_n\}$ in A converging to x.

Let X be a complete metric space, and let $\{F_n\}$ be a decreasing sequence of non-empty closed subsets of X such that $d(F_n) \to 0$. Prove that $\bigcap_{n=1}^{\infty} F_n \neq \phi$.

Is the intersection of infinitely many open subsets of a metric space open ? Justify your answer. [4+2+5+6+3]

2. When is a function from a metric space into another metric space said to be uniformly continuous ?

Let $A \subseteq X$ and define a function f by $f(x) = d(x,A)$. Show that f is uniformly continuous. Also show that the closure of A, $\bar{A}$ is given by

$$\bar{A} = \{x \in X : d(x,A) = 0\}.$$

Let Y be a metric space and let A be a subspace of X. Prove that if f and g are continuous functions of X into Y such that $f(x) = g(x)$ for every $x \in A$, then $f(x) = g(x)\ \forall\ x \in \bar{A}$.

Let Y be a complete metric space and let A be a dense subspace of X. Prove that if $f : A \to Y$ is uniformly continuous, then f can be extended uniquely to a uniformly continuous function $g : X \to Y$. [2+7+4+7]

p.t.o.

3. Define a compact metric space.
Show that a compact subset of a metric space is closed.

If X is compact, then show that X is sequentially compact. Is the converse true ?

State Ascoli-Arzela Theorem.

Let $\{x_n\}$ be a sequence in X converging to a point x. Show that the set $A = \{x_n : n \geq 1\} \cup \{x\}$ is a compact subset of X. [2+3+8+3+4]

4. Let $\mathcal{U}$ be an open cover of a compact metric space X. Show that there is a real $\varepsilon > 0$ such that for every $x \in X$ there is a V in $\mathcal{U}$ such that $S_\varepsilon(x) \subseteq V$.

Hence, or otherwise, show that any continuous function of a compact metric space into a metric space is uniformly continuous.
Let (Y,d) be a metric space and X a subspace of Y. Suppose there is a complete metric ρ on X which is equivalent to d (on X). Show that X is a G_δ subset of Y. [6+5+9]

5. State and prove Baire Category Theorem.

Can $\mathbb{R}$ (the real line with the usual metric) be written as $\mathbb{R} = \bigcup_{n=1}^{\infty} F_n$, where each F_n is closed and nowhere dense ?
Justify.

Show that the set of rationals is an F_σ but not a G_δ subset of $\mathbb{R}$. [10+5+5]

6. Let A be a dense subspace of X. Suppose every Cauchy sequence in A converges to some point in X. Prove that X is a complete metric space.

Let f be a real-valued uniformly continuous function on a bounded set $E \subseteq \mathbb{R}$. Show that f is bounded.

Let X be a compact metric space and Y a metric space. Suppose $f : X \to Y$ is a continuous, one-one and onto function. Show that f is a homeomorphism. [8+6+6]

1987-88 | 521P

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

STATISTICAL METHODS AND INFERENCE II :
TESTS OF HYPOTHESIS (PRACTICAL)

Periodical Examination

Date : 2.3.88. Maximum Marks : 100 Time : 3 Hours

Note : Answer both questions.

1.(a) On the basis of 16 i.i.d. observations from $N(\mu,\sigma_0^2)$, σ_0^2 is given variance, write the critical region of UMP unbiased size 0.05 test for $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$, μ_0 is a given value of μ, and draw its power function.

(b) When the data are : 93, 89, 112, 8, 93, 11, 16, 32, 31, 37, 46, 35, 30, 8, 23, 33, and $\sigma_0^2 = 800$, test $H_0 : \mu = 50$ against $H_1 : \mu \neq 50$ at 5% level of significance and construct the corresponding 95% confidence interval for μ.

[5+25+5+10 = 45]

2. Based on 14 i.i.d. Bernoulli (p) observations, construct the best test of size 0.05 for testing $H_0 : p \leq 1/3$ against $H_1 : p > 1/3$ and draw its power function.

[20+25 = 45]

Practical records. [10]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

STATISTICAL METHODS AND INFERENCE II :
TESTS OF HYPOTHESIS (THEORY)

Periodical Examination

Date : 2.3.88.   Maximum Marks : 100   Time : 3 Hours.

Note : Answer all questions.

1. Define and discuss each term with an example.(a) hypothesis, simple and composite hypothesis, (b) Type I and Type II errors, (c) size and level of significance, (d) power function, (e) MP and UMP tests, and (f) unbiased test. [4x6 = 24]

2. A sample space consists of m (finite) elements $x_1, x_2, \ldots, x_m$. P_1 and P_2 are two alternative probability distributions on the sample space such that

$$P_j\ [\ X = x_i] = p_{ij} > 0 \text{ for all } i \ \& \ j.$$

(a) What element of the sample space is to be included in the critical region for testing

$$H_0 : P = P_1 \text{ against } H_1 : P = P_2$$

if the critical region is to consist of only one element ? Give your reasons.

(b) What if two elements were to be included in the critical region ?

(c) What elements are to be included in the size α critical region ? [6+3+5 = 14]

3. State and prove Neyman Pearson Lemma. [10+10= 20]

4. Let $X_1, X_2, \ldots X_n$ be i.i.d. $N(\mu, 1)$. Find the UMP size α test for $H_0: \mu \leq \mu_0$ against $H_1 : \mu > \mu_0$ (μ_0 given) and show that the test is UMP size α. [20]

5.(a) Show that the confidence region obtained from an acceptance region of a UMP test is uniformly most accurate.

(b) Using the size α test function ϕ given by

$$\phi(x) = \begin{cases} 1 & \text{if } x \leq a\theta_0 \text{ or } x > b\theta_0 \\ 0 & \text{if } a\theta_0 < x \leq b\theta_0 \end{cases}$$

where a and b are known constants, a < b; θ_0 is a given value of θ, $\theta > 0$. Construct a $100(1-\alpha)\%$ confidence interval for θ. [8+4 = 12]

6. Assignments. [10]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

PROBABILITY II
Periodical Examination

Date : 29.2.88. Maximum Marks : 100 Time : $3\frac{1}{2}$ Hours

1.(a) Let $(\Omega, \mathcal{A})$ be a measurable space and f a real valued function on Ω. When is f $\mathcal{A}$- measurable ? (Give the definition). [3]

(b) Let $\mathcal{B}$ be the Borel σ-field on $\mathbb{R}$ and $\mathcal{A}$ any σ-field on $\mathbb{R}$. Show that $\mathcal{B} \subseteq \mathcal{A}$ iff every continuous function on $\mathbb{R}$ is $\mathcal{A}$-measurable. [7]

(c) Let $\mathcal{F}$ be a field of subsets of Ω and let $f : \Omega \to \mathbb{R}$ be $\sigma(\mathcal{F})$-measurable. Show that there is a countable sub-collection $\mathcal{C}$ of $\mathcal{F}$ such that f is $\sigma(\mathcal{C})$-measurable. [13]

2. Outline the main steps in the definition of the integral of a measurable function defined on a measure space $(\Omega, \mathcal{A}, \mu)$. [No details necessary - a clear outline of the main steps is enough. However, you must state clearly any important fact that is needed to make the definition work]. [10]

3.(a) State the monotone class theorem. [5]

(b) Consider the unit interval with the Borel σ-field and Lebesgue measure λ. Show that given any Borel set B and $\varepsilon > 0$, there is a sequence of open intervals $I_1, I_2, \ldots$ such that $B \subseteq \bigcup_{k=1}^{\infty} I_k$ and $\lambda(\bigcup_{k\geq 1} I_k - B) < \varepsilon$. [17]

4. Recall that $\{f_n\}_{n\geq 1}$ is uniformly integrable if

$$\lim_{\alpha\to\infty} \sup_n \int_{\{x: |f_n(x)| \geq \alpha\}} |f_n| d\mu = 0.$$

Show that if $\{f_n\}$ is uniformly integrable then $\{f_n\}$ is uniformly absolutely continuous, i.e., for each $\varepsilon > 0$ there is $\delta > 0$ such that for any A such that $\mu(A) < \delta$ we have

$$\int_A |f_n| d\mu < \varepsilon \text{ for all } n \geq 1.$$

[10]

p.t.o.

5.(a) State Fubini's theorem for the product of two spaces. (State all hypotheses carefully).

Let E be a measurable set in the product σ-field. What does Fubini's theorem say when specialized to I_E ?

[8]

(b) Let μ be Lebesgue measure on $\mathbb{R}$ and γ be counting measure on $\mathbb{R}$ (both on the Borel σ-field), i.e.,

$$\gamma(A) = \infty \quad \text{if A is infinite}$$
$$= \text{no. of elements in A, if A is finite.}$$

Let $E = \{(x,y) : x = y\}$

Show that $\int \gamma(E_x)d\mu \neq \int \mu(E^y)d\gamma$.

Why does this not contradict Fubini's theorem.

[7]

6.(a) State Fatou's lemma. [3]

(b) Let $\{f_n\}$, $\{g_n\}$, $\{h_n\}$ be three sequences of measurable functions on a measure space $(\Omega, \mathcal{A}, \mu)$ such that

$$f_n \to f \quad \text{a.e.}(\mu) \ ; \quad g_n \to g \quad \text{a.e.}(\mu), \text{ and } h_n \to h \quad \text{a.e.}(\mu)$$

Assume that $f_n \leq h_n \leq g_n$ a.e.(μ) for all $n \geq 1$, and that

$$\int f_n d_\mu \to \int f d_\mu$$

and $\int g_n d_\mu \to \int g d_\mu$

Show that

$$\int h d_\mu - \int f d_\mu \leq \liminf \int h_n d_\mu - \int f d_\mu$$

[Hint : Apply Fatou's lemma to a suitable sequence].

[8]

(c) Show that, in the above,

$$\int h_n d_\mu \to \int h d_\mu$$ [9]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

DESIGN OF EXPERIMENTS I
Periodical Examination

Date : 26.2.88. Maximum Marks : 100 Time : 3 Hours

Note : Answer any five questions. Each question carries 16 marks.

1. Describe with examples from the fields of industry and agriculture the role of randomisation and replication in the planning of an experiment.

2.(a) What do you mean by a uniformity trial and what are its uses ?
  (b) How can you compare the efficiency of a Latin Square Design and a Randomised Block Design vis-a-vis a Completely Randomised Design with the same number of experimental units ?
  (c) Describe a suitable randomisation procedure for a Latin Square Design.

3. Suppose in an experiment laid out in a Latin Square Design, the observation in one of the plots is missing. Describe with adequate justifications a method for analysing the available observations.

4. What do you mean by a set of mutually orthogonal Latin Squares ? Show that the maximum number of mutually orthogonal latin squares of order v is $\leq$ v-1.
Prove that if v is a prime number or a prime power, there exists a complete set of mutually orthogonal latin squares of order v.

5. What is a BIB Design ?
Prove Fisher's inequality for a BIB Design. How will you test the hypothesis of equality of all treatment effects from the yield values obtained in a planned BIB layout.

6.(a) What is a crossover design ? How will you analyse a crossover design ?
  (b) Suppose a Randomised Block Design with v treatments and r blocks is replicated p times. How will you analyse the data obtained from such an experiment ?

Practical Notebook. [20]

INDIAN STATISTICAL INSTITUTE
M.Stat.(S-and M-stream) I Year : 1987-88

OPTIMIZATION TECHNIQUES I
Periodical Examination

Date : 24.2.88. Maximum Marks : 100 Time : $3\frac{1}{2}$ Hours

Note : This paper carries a total of 110 marks. Answer as many questions or parts thereof as you can. If your score exceeds 100 marks your actual score will be treated as 100 marks.

1.(a) State the duality theorem for the Standard Maximum Linear Programming (SMLP) problem after defining all the relevant mathematical terms in the statement.

(b) Write down the dual of the following General Maximum Linear Programming (GMLP) problem.

$$\text{Maximise } Z = 4x_1 + 7x_2 + 9x_4 - 10x_5$$

$$\text{Subject to } x_1, x_3, x_4 \geq 0,$$

$$2x_1 - 3x_2 + 7x_3 - 8x_4 \leq 3,$$

$$-3x_1 - 5x_2 + 19x_4 - 5x_5 \leq 7,$$

$$4x_1 + 9x_2 - 8x_4 + 7x_5 \leq 10.$$

(c) Prove or disprove that the above GMLP problem has an optimal solution. (7+6+12=25)

2.(a) State and prove the equilibrium theorem for an SMLP problem.

(b) Prove or disprove that x where $x' = (0, \frac{4}{7}, \frac{12}{7}, 0, 0)$ is an optimal solution of the following standard minimum linear programming problem.

$$\text{Min } z = x_1 + 6x_2 - 7x_3 + x_4 + 5x_5$$

$$\text{Subject to } 5x_1 - 4x_2 + 13x_3 - 2x_4 + x_5 = 20,$$

$$x_1 - x_2 + 5x_3 - x_4 + x_5 = 8,$$

$$x_i \geq 0, \quad 1 \leq i \leq 5.$$

(c) Find all the optimal solutions of the primal problem. (10+10+10 = 30)

p.t.o.

3.(a) Maximise $Z = \sum_{i=1}^{4} x_i$, $x_i \geq 0,\ 1 \leq i \leq 4,$

Subject to $\begin{pmatrix} 2 & 3 & -1 & 4 \\ 1 & -2 & 6 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 8 \\ 3 \end{pmatrix},$

by finding all the basic feasible solutions of the above problem.

(b) Reduce the feasible solution x where $x' = (\frac{17}{7}, \frac{17}{14}, \frac{1}{2}, 0)$ of the above CMLP problem of 3(a) to a basic feasible solution. (15+10=25)

4.(a) For a CMLP problem with usual notation if $c_j - z_j \leq 0$ for all j for a basic feasible solution x, then prove that x is an optimal solution of the above problem (Define the mathematical terms involved like z_j, y_{ij} in the proof).

(b) Using the simplex method Tabular forms

Maximise $2x_1 + x_2$

Subject to
$$3x_1 + 5x_2 \leq 15,$$
$$6x_1 + 2x_2 \leq 24,$$
$$x_1, x_2 \geq 0.$$

(c) Prove that if a CMLP problem has an optimal vector, then it has a basic optimal vector.

(10+12+8 =30)

1987-88 | 531

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year : 1987-88

MULTIVARIATE ANALYSIS IA
Periodical Examination

Date : 22.2.88. Maximum Marks : 100 Time : 3 Hours

1.(a) Define p-dimensional normal variable.
(b) Show that its mean vector and dispersion matrix exist.
(c) Obtain its characteristic function.
(d) Obtain the maximum likelihood estimates for the mean vector and the dispersion matrix based on a random sample of size n.
(e) State the distributions of the above maximum likelihood estimates and show that they are independent.
(f) Obtain the mean vector and dispersion matrix of the bivariate distribution of (X,Y) whose density is given by

$$f(x,y) = \{\exp[-\tfrac{1}{2}(2x^2+y^2+2xy-22x-14y+65)]\}/2\pi$$

Also obtain the conditional distribution of X given Y.

[3+5+7+15+10+10 =50]

2.(a) Define wishart matrix (variate).
(b) Let $W \sim W_p(k,\Sigma)$. Let $W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$.

Obtain the distribution of $w_{22.1} = w_{22}-w_{21}w_{11}^{-1} w_{12}$

and show that it is distributed independent of w_{11} and w_{12}.
(c) Obtain the distribution of $\frac{\ell'\Sigma^{-1}\ell}{\ell' W^{-1}\ell}$ where ℓ is a fixed vector.
(d) Show that $\frac{|W|}{|\Sigma|}$ is distributed as product of p independent central χ^2-variables with degrees of freedom k-p+1,....,k.

[4+10+8+8 = 30]

3. Define Hotellings T^2-statistic and show that it is distributed as constant times F-variable. [10]

p.t.o.

4.(a) State the problem of Linear regression.

(b) Consider the random vector $(x_1, x_2, x_3 \ldots x_p) = (x_1, x_2, X)$ with mean vector (μ_1, μ_2, μ) and dispersion matrix

$$S = \begin{pmatrix} s_{11} & s_{12} & s_1' \\ s_{21} & s_{22} & s_2' \\ s_1 & s_2 & s \end{pmatrix}$$

Write down (need not derive), for i = 1,2 the Linear regression function of x_i on X. Let e_i be the residual.

(c) Define Partial Correlation Coefficient $\rho_{12.(3\ldots.p)}$

Show that it is equal to correlation between e_1 and e_2.

[5+5+10=20]

[1987-88/542(b)]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
SEMESTRAL-I BACKPAPER EXAMINATION
Statistical Methods and Inference I: Estimation

Date: 31.12.87 Maximum Marks: 100 Time: 3 Hours

Note: Answer any three questions from Group A and one question from Group B. Ten marks are reserved for assignment and practical records.

Group A
(answer any three questions)

1. Derive an expression for optimum allocation of sample sizes, given variance, in stratified sampling with SRSWOR in each stratum. [15]

2.(a) Show that Fisher-information is additive when observations are i.i.d.

(b) Show that, Fisher-information based on n i.i.d. observations is same as Fisher-information based on a statistic if and only if the statistic is sufficient. [6+9=15]

3.(a) State and derive Chapman-Robbin lower bound for variance of an estimator.

(b) Compute the lower bound for estimating the parameter of a Bernoulli distribution based on a single observation. [6+9=15]

4.(a) Show that if a sufficient statistic is complete, then it is unique UMVUE of its expectation.

(b) Show that the largest of n i.i.d. observations from $U(0,\theta)$ is sufficient and complete. [5+10=15]

Group B
(answer any one)

5. Following is a hypothetical population of 22 equal size experimental plots, stratified into 3 strata on the basis of input, along with yield (in 100 kg.) rice per plot.

Stratum 1		Stratum 2		Stratum 3	
Plot No.	Yield	Plot no.	Yield	Plot no.	Yield
1	16.7	1	18.2	1	21.3
2	11.9	2	16.7	2	15.8
3	13.2	3	13.2	3	20.2
4	14.5	4	17.8	4	19.9
5	14.9	5	18.6	5	21.4
		6	19.2	6	20.6
		7	18.8	7	20.1
		8	18.6	8	19.8
				9	22.0

5.(a) Draw a SRSWOR sample of size 2 from stratum 1, a SRSWR sample of size 3 from stratum 2, and a SRSWOR sample of size 3 from stratum 3.

(b) Estimate the stratum means and stratum variances.

(c) Estimate the population mean and its variance. [9+27+9=45]

6. The number of deaths in fatal accidents in a town was recorded for six fatal accidents as 2,4,2,1,1,3. Assume truncated (at zero) Poisson distribution (with parameter λ) of the number of deaths in fatal accidents in the town.

Estimate λ by the method of maximum likelihood and compute its approximate error. [35+10=45]

Assignment and practical record. [10

:SS:

[1987-88/513(b)]

INDIAN STATISTICAL INSTITUTE

M.Stat.(M-stream) I Year
SEMESTRAL-I BACKPAPER EXAMINATION
Linear Algebra and Linear Models

Date: 23.12.87 Maximum Marks: 100 Time: 3 hours

Note: Answer as many as you can.

1. State and Prove Caley-Hamilton theorem. [10]

2. State and Prove Fisher-Cochran theorem. [15]

3. State and Prove Gauss-Markoff theorem. [10]

4. In an investigation for studying the effect of smoking on physical activity, 21 individuals were classified into one of the three groups by smoking history and randomly assigned to one of the three stress tests B1, B2 and B3. The time until maximum oxygen uptake was recorded. The data is given below :

Time until maximum oxygen uptake

Smoking history	Stress test B_1	B_2	B_3
A_1: None	12.6 13.5 11.2	16.2 17.8	22.5 19.3 18.9
A_2: Moderate	10.9	15.5 13.8 16.2	21.0 15.9
A_3: Heavy	9.2 7.5	13.2 8.1	16.2 16.1 17.8

Analyse the data by assuming no interactions. [30]

5. Write short notes on :
   (a) Tukey's test for non-additivity
   (b) General regression problem
   (c) Eigen values and singular values of matrices [15]

6. Assignments. [30]

:SS:

[1987-88/162]

INDIAN STATISTICAL INSTITUTE
M.Stat.(S-stream) I Year
SEMESTRAL-I EXAMINATION
Computer Programming and Applications

Date: 25.11.87 Maximum Marks: 50 Time: 3 Hours for Part A and B together

GROUP A

Note: Answer Part-A seperately

1. For each of the following, show the effect and comment on it:

(a)
```
COMMON       A,B(4),C
EQUIVALENCE (A,B(1),C)
```
(b)
```
DIMENSION    A(5)
COMMON     B,C,D
EQUIVALENCE (D,A(1))
COMMON  E,F
```
[2x5]

2.(a) Describe the two ways of passing argument values between a calling program and a subprogram.

(b) What are the advantages and disadvantages of the above two ways in (a)? Illustrate. [2x5]

3.(a) Write a FUNCTION subprogram to calculate N! (factorial N).

(b) Using the above subprogram develop another subprogram to calculate the binomial coefficient for given N and K (K<N):

$$\binom{N}{K} = \frac{N!}{K!\,(N-K)!}$$

(c) Now, calculate

$$PROB = \binom{n}{k} p^k q^{n-k}$$

where p and q are probabilities of success and failure respectively and PROB gives the probability of exactly K successes in n trials of an experiment. Taking n = 100 and K = 1,2,....10 write a program to tabulate PROB for each K, assuming p = 0.58. [3+5+7]

4. Given the following program:

```
      E = 3.0
      F = 4.0
      G = 5.0
      J = E+F/G
    5 GOTO (2,1,3),J
    1 H = H+E+F+G
      GOTO 4
    2 H = E * F * G
      GOTO 4
    3 H = E ** F
      J = J + 1
      GOTO 5
    4 WRITE ( 6,10) H
   10 FORMAT ( 2X., F8.3)
      STOP
      END
```

p.t.o.

Trace the flow of the program execution, in the form of a flow - chart and find the value of H.

5. Detect and correct the mistakes, if any, in the following statements :

(a) REAL *2 A,B,K,D

(b) DATA X,Y,Z/3.0 * 1/

(c) IMPLICIT A- F

(d) DOUBLE PRECISION NUMBER (4,7) X,Y, TEST (8,15)L, M

(e) DIMENTION A(5,5), C(5,N), K(M,M) [5x2]

GROUP B

Maximum Marks : 50

Note: Answer this part in separate answer book. Answer any part of any question. Maximum you can score is 50.

1.(a) Assume that the computer rounding rule is to "chop" and retain only first 4 significant digits. Find the rounding error and relative error introduced through conversion of the following numbers to floating point numbers.

(i) 1.16702 (ii) -6.66

(iii) 1742.52 (iv) -16.6299

(v) 0.00067418 [10]

(b) Give an example which shows that

$$fl\ [fl\ (a+b) + fl(c)] \neq fl\ [fl\ (a)+fl(b+c)]$$ [5]

2.(a) Given a real valued function $f(x)$ and $(n+1)$ distinct points $x_0, x_1, \ldots x_n$, show that there exists exactly one polynomial of degree $\leq n$ which interpolates $f(x)$ at $x_0, \ldots\ x_n$. [10]

(b) Prove or disprove the following statement:
"Addition of distinct points will always increase the degree of the interpolation polynomial." [5]

3. Find the real root of the polynomial

$$x^2 + x - 3$$

by Miiller's method. [10]

4. Find a lower triangular matrix L such that $A = L \cdot L'$ and hence compute the inverse of A and its determinant, where

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 8 & 12 & 16 \\ 3 & 12 & 27 & 36 \\ 4 & 16 & 36 & 64 \end{bmatrix}$$ [25]

5. Find the characteristic polynomial of the matrix

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 3 & 0 \\ 0 & 3 & 1 & 4 \\ 0 & 0 & 4 & 1 \end{bmatrix}$$ [5]

: 88:

1987-88/532

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
SEMESTRAL-I EXAMINATION
Analysis IB

Date: 23.11.87 Maximum Marks: 100 Time:

Note: Each question carries 15 marks. Answer as many questions as you can. The maximum you can score is 100.

1. If $p : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is defined by

$$p(x,y) = x.y,$$

then $[p'(a,b)](x,y) = bx + ay$.

2. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a continuously differentiable function. Show that f is not one-one.

3. Let $f(x_1, \ldots\ldots, x_n) = x_1^2 + \ldots\ldots + x_n^2$. Find the maximum or minimum of f subject to $x_1 + x_2 + \ldots\ldots + x_n = 1$

4. Show that there can be no non constant analytic function $f : G \longrightarrow \mathbb{C}$ satisfying any one of the following conditions

(i) $\{ f(z) : z \in G \} \subseteq \mathbb{R}$

(ii) $\{ f(z) : z \in G \} \subseteq \{ z : z = r \}$

5. State the residue theorem and verify using the residue theorem (if necessary) if the canchy Integral formula holds for

(a) $$\int_\gamma \frac{e^z}{z-1} dz = 2\pi i e$$

(b) $$\int_\gamma \frac{z}{z^4-1} dz = \frac{1}{2}\pi i ,$$

where $\gamma = \{ z : z = 2e^{it}, \quad 0 \leq t \leq 2\pi \}$

6. Find the Taylor series expanssion of the function $f(z) = \sum_{n=0}^{\infty} z^n$ about the points (a) $z = 1/2$ and (b) $z = -1/2$

In each case find the radius of convergence.

7. Classify the singularities of

$f(z) = \frac{1}{z^4+z^2}$ and find the Laurant expanssion for $f(z)$ about $z=0$.

8.(a) Prove the Fundamental theorem of Algebra

(b) Show that whenever f is bounded in a deleted neighbourhood of and isolated singularity, the singularity is removable.

:ss:

[1987-88/512]

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
SEMESTRAL-I EXAMINATION
Statistical Methods and Inference I: Estimation(Theory)

Date: 20.11.87 Maximum Marks: 100 Time: 3 hours

Note: Answer any one question from Group A, and any four of the nine questions from Group B. Ten marks are reserved for assignments.

GROUP A

(answer any one question)

1.(a) Based on a SRSWOR, show that the sample proportion is unbiased for the population proportion.

(b) Find an unbiased estimator of the variance of the sample proportion and show its unbiasedness. [6+12=18]

2.(a) Derive an expression for optimum allocation of sample sizes, given cost, in stratified sampling with SRSWOR in each stratum.

(b) Hence find an expression for Neyman allocation. [14+4=18]

GROUP B

(answer any four questions)

3.(a) Define Fisher-information based on a single observation whose distribution depends on a single parameter.

(b) Show that Fisher-information is additive when observations are i.i.d.

(c) Show that, Fisher-information based on n i.i.d. observations is same as Fisher-information based on a statistic if and only if the statistic is sufficient. [3+6+9=18]

4.(a) State and prove (discrete case will do) the factorization theorem.

(b) Find a sufficient statistic for θ when $X_1, X_2, \ldots, X_n$ are i.i.d. $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$.

(c) Show that m.l.e. is a function of sufficient statistic. [8+6+4=18]

5.(a) State and derive Chapman-Robbin lower bound for variance of an estimator.

(b) Compute the lower bound for estimating the parameter of a Bernoulli distribution based on a single observation. [8+10=18]

6.(a) State and derive Rao-Cramer inequality for mean square error of an estimator.

(b) When can the equality be attained by an unbiased estimator?

(c) Compute the Rao-Cramer lower bound for an unbiased estimator of the parameter of a Poisson distribution. [8+5+5=18]

p.t.o.

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

(b) Hence find UMVUE of θ based on n i.i.d observations from $U(0, \theta)$, starting with an unbiased estimator of θ based on one observation. [9+9=18]

8.(a) Show that a statistic is an MVUE of a parametric function at a given point if and only if its covariance, with every unbiased estimator of zero, is zero at that given value of the parameter.

(b) Show that $T(0) = 1$, $T(x) = 0$ for $x \neq 0$ is UMVUE of $(1-\theta)^2$ when it is known that $U(x) = kx$, for arbitrary k, is the class of all unbiased estimators of zero for the distribution.

$$P_\theta[X = -1] = \theta\ ,\ P_\theta[X=x]=(1-\theta)^2\theta^x,\ x = 0, 1, 2, \ldots\ldots .$$

[9+9=18]

9.(a) Show that if a sufficient statistic is complete, then it is unique UMVUE of its expectation.

(b) Show that the largest of $X_1, X_2, \ldots, X_n$, which are i.i.d. $U(0,\theta)$ is sufficient and complete. [6+12=18]

10.(a) Find UMVUE of θ^r, r is an integer, based on n i.i.d. observations from the integer valued random variable with distribution

$$P_\theta[X=x] = C(\theta)\,\theta^x a(x),\ x = s,\ s+1,\ s+2, \ldots, a(x)>0,\ x\geq s.$$

(b) Hence find UMVUE of the Poisson parameter based on n i.i.d. observations. [12+6=18]

11.(a) State Basu's theorem.

(b) Hence find UMVUE of $\Phi(a-\mu)$ based on n i.i.d. observations from $N(\mu,1)$, for fixed a. [10+8=18]

Assignments [10]

:SS:

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
SEMESTRAL-I EXAMINATION
Statistical Methods and InferenceI: Estimation (Practical)

Date: 20.11.87 Maximum Marks: 100 Time: 3 hours

Note: Answer any two questions. Ten marks are reserved for practical records.

1. Following is a hypothetical population of 22 equal size experimental plots, stratified into 3 strata on the basis of input, along with yield (in 100 kg.) of rice per plot.

Stratum 1		Stratum 2		Stratum 3	
Plot No.	Yield	Plot No.	Yield	Plot No.	Yield
1	16.7	1	18.2	1	21.3
2	11.9	2	16.7	2	15.6
3	13.2	3	13.1	3	20.2
4	14.6	4	17.9	4	19.9
5	14.8	5	18.6	5	21.4
		6	19.2	6	20.6
		7	18.8	7	20.1
		8	18.6	8	19.8
				9	22.0

(a) Draw a SRSWOR sample of size 2 from stratum 1, a SRSWR sample of size 3 from stratum 2, and a SRSWOR sample of size 3 from stratum 3.

(b) Estimate the stratum totals and stratum variances.

(c) Estimate the population total and population variance. [9+27+9=45]

2. Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ and T_n be the m.l.e. of σ^2. Given $\eta > 0$, and $\epsilon > 0$, find smallest N_0 such that

$$P\left[\left| \frac{T_n}{\sigma^2} - 1 \right| < \epsilon \right] > 1 - \eta \text{ for all } n > N_0$$ by selecting suitable values of η and ϵ. Hence discuss N_0 as a function of η and ϵ. [45]

Data for questions 3 and 4

The Number of deaths in fatal accidents in a town was recorded for six fatal accidents as 2,4,2,1,1,3. Assume truncated (at zero) Poisson distribution (with parameter λ) of the Number of deaths in fatal accidents in the town.

3. Estimate λ by the method of maximum likelihood and compute approximate error in your estimate. [35+10=45]

4. Compute UMVUE of λ and UMVUE of the variance of the UMVUE of λ. [25+20=45]

Practical records. [10]

1987-88 / 522

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

Date: 18.11.87 Maximum Marks: 100 Time: $3\frac{1}{2}$ Hours

Note: Answer all questions

1. **Either** [20]

(a) Show that for a collection of random variables $\{X, X_n : n \geq 1\}$ defined over the same probability space if

$$\sum_{n=1}^{\infty} P(\{w: |X_n(w)-X(w)| > \tfrac{1}{k}\}) < \infty \text{ for every integer } k>0,$$

then

$$P(\{w: \lim_{n\to\infty} X_n(w) = X(w)\}) = 1.$$

Or

(b) Use the result in (a) to prove Borel's strong law of large numbers.

2. **Either** [20]

(a) Prove the central limit theorem. (You must state carefully the results assumed for proving the CLT)

Or

(b) Stating very briefly the usual assumptions for a Poisson process, derive a differential equation for the probability $P_0(t)$ of no events up to time t and hence find the density of the waiting time for the first event.

Or

(c) Find the characteristic function of the cauchy density

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

3. Answer any three from the following. [25]

(i) Consider the following model for a batsman. The probability that a batsman gets out to a given ball is $p (0 < p < 1)$. If the batsman is not out to a given ball, he scores $i = 0, 1, 2, 3, 4$ and 6 with, say, probability $p_i = \frac{5}{6}$ if $i=0$ and $=\frac{1}{30}$ if $i>0$. Assuming independence, find the probability generating function of his total score in an innings.

(ii) Let X_i, $1 \leq i \leq n$, be independent random variables each having exponential distribution with parameter λ. Denote the corresponding order statistics by $X_{(i)}$, $1 \leq i \leq n$, where $X_{(1)} \leq \cdots \leq X_{(n)}$. Show that the n variables $X_{(1)}$, $X_{(2)}-X_{(1)}, \ldots, X_{(n)}-X_{(n-1)}$ are independent with $X_{(k+1)}-X_{(k)}$ having exponential distribution with parameter $(n-k)\lambda$. $(k=0,1,\ldots, n-1.\ X_0 \equiv 0)$

p.t.o.

(iii) The function

$P(s_1, s_2) = \exp(-\lambda_1 - \lambda_2 - \lambda + \lambda_1 s_1 + \lambda_2 s_2 + \lambda s_1 s_2)$, $\lambda_1, \lambda_2, \lambda > 0$, is the generating function of a bivariate Poisson distribution. Denote the corresponding random vector by (X_1, X_2). Show that X_1, X_2 are independent iff $Cov(X_1, X_2) = 0$.

(iv) Suppose $X \sim$ exponential (λ). Find a function f such that $f(X) \sim u[0,1]$.

4. Answer any one from the following: [10]

(i) Suppose X is a random variable having the following density :

$$f(x) = \frac{C}{1 + |x|^k}, \quad -\infty < x < \infty, \ k > 1.$$

For what values of r(r is a positive integer) is $E|X|^r < \infty$?

For what values of t is $E(e^{tX}) < \infty$?

(ii) Prove that as $n \rightarrow \infty$

$$\frac{1}{\Gamma(\frac{n}{2})} \sqrt{\left(\frac{n}{2}\right)^n} \int_0^{1+t\sqrt{\frac{2}{n}}} x^{\frac{n}{2}-1} e^{-\frac{nx}{2}} dx \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx,$$

for every real number t.

5. Answer any one of the following: [10]

(i) Suppose X is a continuous random variable and Y is independent of X. Show that X+Y is a continuous random variable.

(ii) Let X_i i.i.d. $B(1, \frac{1}{2})$, $i \geq 1$. Define a new random variable S by

$$S = \sum_{i=1}^{\infty} \frac{X_i}{2^i}.$$

Show that S converges everywhere. Find the distribution of S. [You can solve this problem immediately if you do the following calculation:

Consider $Y \sim U(0,1)$. Define a collection of random variables $\{Y_i : i \geq 1\}$ by

$Y_1 = 0$ if $0 < Y \leq \frac{1}{2}$, $Y_2 = 0$ if $0 < Y \leq \frac{1}{4}$ or $\frac{1}{2} < Y \leq \frac{3}{4}$

$= 1$ if $\frac{1}{2} < Y \leq 1$ $\quad = 1$ if $\frac{1}{4} < Y \leq \frac{1}{2}$ or $\frac{3}{4} < Y \leq 1$.

In general, $Y_i=0$ if $\frac{k-1}{2^i} < Y \leq \frac{k}{2^i}$, k odd and $1 \leq k \leq 2^i$

$=1$ if $\frac{k-1}{2^i} < Y \leq \frac{k}{2^i}$, k even and $1 \leq k \leq 2^i$

Find out the marginal distribution of $Y_i (i \geq 1)$ and show that they are independent.]

6. Assignment. [15]

:ss:

1987-88/512

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
SEMESTRAL-I EXAMINATION
Linear Algebra and Linear Models

Date: 16.11.87 Maximum Marks: 100 Time: 3 hours

Note: Answer as many as you can.

1. Let A be an idempotent matrix of order n and rank r such that

$$A = A_1 + A_2 + \dots + A_k$$

where the matrix A_i is of rank r_i, for $i = 1,2\dots k$. Show that

$$\sum_{i=1}^{k} r_i = r \iff A_i^2 = A_i \text{ , for } i = 1,2\dots k.$$

2. Let $A = ((a_{ij}))$ be a positive definite matrix of order n and $B = ((b_{ij}))$ be its inverse. Then show that

(a) $b_{ii} \geq \frac{1}{a_{ii}}$ with equality if and only if $a_{ij} = 0$ for $j \neq i$

(b) $|A| \leq a_{11} \cdot a_{22} \cdots a_{nn}$. [10]

3.(a) Describe standard Gauss-Markoff model $(Y, X\beta, \sigma^2 I)$.

(b) Obtain least squares estimator for the parameter vector β.

(c) Define estimability for a linear parametric function $p'\beta$ and obtain a necessary and sufficient condition for the same.

(d) Show that for an estimable linear parametric function its least squares estimator is BLUE.

(e) Show that $t'Y$ is BLUE of a linear parametric function if and only if t belongs to the column space of the design matrix X.

[3+5+7+5+5=25]

4. Define and derive linear regression function of the variable X_0 on the variables $X_1 \dots X_p$ and multiple correlation coefficient $\rho_{0.12\dots p}$ interms of expectations, variances and covariances of the variables. [10]

5. In an experiment to study the effect of glass type and phosphor type on the brightness of a television tube the following results are obtained. The measured variable is the current in micro amperes necessary to produce a certain brightness, the larger the current, the poorer the tube-screen characteristics.

p.t.o.

5. Contd.....

Analyse the data and give your comments

Current (in microamperes) necessary to produce a certain brightness on the t.v. screen

Glass type	Phosphor type A	Phosphor type B	Phosphor type C
1	280 290 285	300 310 295	270 285 290
2	230 235 240	260 240 235	220 225 230

[25]

6. Assignments [30]

1987-88 541(S)

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) I Year
PERIODICAL SUPPLEMENTARY EXAMINATION

Statistical Methods and Inference I: Estimation

Date: 10.11.87 Maximum Marks: 100 Time: 3 hours

Note: Answer all questions. Ten marks are reserved for practical records and assignments.

1. Select a SRSWOR of size $(n-1)$ from a population of N units.
   (a) Select another unit from the remaining $N-n+1$ units. Let y_n be the value of a study variable associated with this (last selected) unit. Compute expected value and variance of y_n.
   (b) Select a further SRSWOR of size r from the remaining $N-n$ units. Based on this sample of size r, construct unbiased estimators of population mean and variance, and show their unbiasedness.
   (15+20=35)

2. For a given budget, discuss gain due to stratification after deriving necessary variance formulae. (25)

3. Based on n i.i.d. observations from $(N(\mu,\sigma^2)$, find MLE of μ and σ^2. (15)

4. Based on n i.i.d. observations from Bernoulli distribution, with θ as the probability of success, find MLE of θ when it is known that θ can take only three values: 0.25, 0.50, 0.75. (15)

Practical records and assignments (10)

:sx:

1987-88 | 541

INDIAN STATISTICAL INSTITUTE
M.Stat. (M-stream) I Year : 1987-88

PERIODICAL EXAMINATION

Statistical Methods and Inference 1 - Estimation

Date: 11.9.1987 Maximum Marks: 100 Time: 3 hours

Note: Answer all questions.

I. A population consists of 13 units. Describe how you will select, using a random number table or random numbers generated by a calculator,

1. a simple random sample of size 2 without replacement, [4]

2. one unit so that the chance of selecting unit number i to the chance of selecting unit number j is in the ratio i to j, (i, j = 1, 2, ..., 13). [6]

II. A population consists of N units.

(A) (For questions 3-6 and 8). A simple random sample is drawn without replacement.

3. What is the probability that a given unit, say unit number i, is included in the sample ? [2]

4. What is the probability that a given pair of units, say unit numbers i and j, are included in the sample ? [2]

5. Show that the sample mean is an unbiased estimator of the population mean. [5]

6. Compute the variance of the sample mean. [8]

7. What would be the variance of the sample mean, had the sample been drawn with replacement ? [3]

8. Suggest an unbiased estimator, based on the sample, of the variance computed in (6), and show its unbiasedness. [6]

9. What would be an unbiased estimator of the variance of the sample mean, had the sample been drawn with replacement ? [3]

Contd..... 2/-

II.(B) (For questions 10-15).

The population is stratified into two strata, stratum h having N_h units, $N_1 + N_2 = N$. A simple random sample of n_1 units are drawn with replacement from stratum 1, and a simple random sample of n_2 units are drawn without replacement from stratum 2; $n_1 + n_2 = n$.

10. When is the sample mean of the combined sample unbiased for the population mean ? [5]

11. Find a linear combination of the two sample means to get an unbiased estimator of the population mean; do not derive but argue why it should be unbiased. (See also 10) [3]

12. Write the variance of the estimator, as obtained in 11, of the population mean; do not derive but argue with results of 6 and 7. [4]

13. Write an unbiased estimator of the variance, obtained in 12; do not derive but argue with results of 8 and 9. [6]

14. If the cost function is $C = c_1n_1 + c_2n_2$, find optimum n_1 and n_2 to minimize the variance, as written in 12, for a given cost. [10]

15. Find an expression for variance, as written in 12, for Neyman allocation. [3]

III. 16. A box contains 10 bullets. 4 are selected at random, without replacement, and fired. 3 bullets were found to be good and the other was defective. Find the maximum likelihood estimator of the number of defective bullets in the box. [10]

17. Based on a random sample of size n from normal distribution with mean μ and variance 1, find the maximum likelihood estimator of μ when it is known that $A \leq \mu \leq B$, A and B are given constants. [10]

IV. Practical records and assignments. [10]

---

bcc:

1987-88	461/551

INDIAN STATISTICAL INSTITUTE
M.Stat.(S) and M.Stat.(M) I Year: 1987-88
PERIODICAL EXAMINATION

Computer Programming and Applications

Date: 9.9.87 Maximum Marks: 100 Time: 3 hours

Note: Part-I and Part-II are to be answered in seperate answer books.

Part-I

1. Write a Fortran expression for each of the following :

a) $4x^2y - 3xy + 7yz^3$

b) $\dfrac{\frac{a}{b}+6}{x-\frac{y}{z}}$

c) $\sqrt{|\cos(a-nb)|}$

d) $(a+b)^3 \leq 50$

e) each side a,b,c of a triangle should be less than the sum of the other two sides. [5]

2. Let X = 3.1, Y = 4.6, L = 2 and M = -3. Execute each of the following segments of a program and find the values of A and J at the end of each segment execution.

a)
```
J = L + 3*M
J = J** 2+Y
A = X+J
```

b)
```
J = 8.0/M-Y
A = 8/M+3*2
```

c)
```
A = L/M * X/Y
J = L/(M+X) *Y
```

d)
```
J = SQRT (X * Y)
A = J * Y
J = A *M
A = A+ABS(FLOAT(J)
```
[8]

3. Execute the following segment of a program

```
      N = 0
      DO 21   I = 1,10,1
      J = 1
      DO  2   M = 5,1
      K = M
2     N = N + 1
21    CONTINUE
```

what are the values of I,J, N and K at the end of execution when

a) DO66 control is assumed and

b) DO77 control is assumed. [10]

4. Give a step by step flow chart to generate a sequence of numbers

1,1,2,3,5,8,13,21,34,55

observe that, the terms from 3rd term onwards are equal to sum of the just previous two terms. [8]

5. Explain the following:

(a) Explicit and Implicit variable declaration

(b) Three examples of difference between FORTRAN IV and FORTR... 77

(c) "Programming logic depends on Input data organisation". Give examples only.

6. The elements of a 4 x 5 matrix are

7	5.2857	24	14	.3242
5.2857	5.0070	2.9386	10.5714	.0527
5.0070	5.0000	0.0700	10.014	.0013
5.0000	5.0000	0	10	0

Give suitable FORTRAN statements to input the matrix element values column-wise. Choose a suitable format.
Prepare the data records according to the chosen format.

PART-II

Note: You can answer any part of any question.
Maximum you can score is 50.

1.(a) Convert base-10 number 123.625 to binary form, stating clearly the algorithm(s) used. From the binary form convert to octal and hexadecimal form. [5+5+2+2=14]

(b) Assuming a computer with four decimal place mantissa, add the following numbers, first in ascending order (from smallest to highest) and then in descending order (from highest to smallest). In doing so, round off the partial sums.
Compare your results with correct sum $x = 0.107101023 \times 10^5$.

$0.1580.10^0$	$0.7555.10^2$
$0.2653.10^0$	$0.7889.10^3$
$0.2581.10^1$	$0.7767.10^3$
$0.4288.10^1$	$0.8999.10^4$
$0.6266.10^2$	

[4+4+3=11

2.(a) Let $|\delta_i| \leq u$ for $i = 1,2\ldots n$ and $nu \leq 0.01$, then show that

$$1 - nu \leq \prod_{i=1}^{n} (1 + \delta_i) \leq 1 + 1.01nu$$

Also, $\prod_{i=1}^{n} (1 + \delta_i) = 1 + 1.01n\theta u$ for some $|\theta| < 1$ [10]

(b) Let $Q(x)$ be the result of computing the value of the polynomial

$$P(x) = \sum_{k=0}^{n} a_k . x^k$$

by the following algorithm

$$q_n = a_n$$

$$q_i = fl\,(q_{i+1} . x + a_i), \text{ for } i = n-1, n-2, \ldots, 0.$$

Prove that

$$Q(x) = \sum_{k=0}^{n} [1 + 1.01\ (2k + 1)\ \theta_k \cdot u]\ a_k x^k$$

where $|\theta_k| < 1$ for all k and u is the unit round off for floating point translation.

Assume that $x, a_0, a_1 \ldots a_n$ and all intermediate answers are floating point numbers and $n \leq \frac{1}{2}(0.01/u - 1)$ [15]

3.(a) Write short notes on condition of a function and stability of a computational process. [6]

(b) Evaluate

$$f(x) = \sqrt{x+1} - \sqrt{x}$$

at x = 12345 with six decimal arithmetic with round off, using the following computational steps :

$$x_0 = 12345$$

$$x_1 = x_0 + 1$$

$$x_2 = \sqrt{x_1}$$

$$x_3 = \sqrt{x_0}$$

$$x_4 = x_2 - x_3$$

Compare your results with the actual value

$$f(12345) = 0.0045000326$$

Is the process described above is stable ? If, not, explain, why not. And, in this case, can you suggest anyother stable process to compute f(x) ? [7+2+2+3+5=19]

: ss:

1987-88 | 531

INDIAN STATISTICAL INSTITUTE
M.Stat. (M-stream) I Year: 1987-88

PERIODICAL EXAMINATION

Analysis IB

Date: 7.9.1987 Maximum Marks: 100 Time: 3 hours

Note: The total marks for this test is 95, attempt as many questions as you can. There will be 15 marks for home work assignment and the maximum you can score is 100.

1. Use Taylors formula to express $x^2 + xy + y^2$ in powers of $(x-1)$ and $(y-2)$. [10]

2. Given that

$$x = u^2 + v^2 \qquad u = t+1$$
and
$$y = e^{uv} \qquad v = e^t \; ,$$

find $\frac{dx}{dt}$ and $\frac{dy}{dt}$ at $t = 0$. [8]

3. Let $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ be differentiable in an open set S in $\mathbb{R}^n$. Then at each point $\underline{x}$ in S for which $\Delta f(\underline{x}) \neq 0$, the vector $\Delta f(\underline{x})$ points in the direction of maximum increase of f. [10]

4.(a) Let $f(x,y) = \dfrac{x^2 - y^2}{x^2 + y^2}$, $(x,y) \neq (0,0)$.

Find $\lim\limits_{(x,y) \to (0,0)} f(x,y)$.

(b) Show that f or $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ with coordinate functions $f_1, \ldots, f_m$ and $(y_1, \ldots, y_m) \in \mathbb{R}^m$, $\lim\limits_{\underline{x} \longrightarrow \underline{x}_0} f(\underline{x}) = \underline{y}_0$

if and only if

$\lim\limits_{x \longrightarrow x_0} f_k(x) = \underline{y}_k \quad k = 1, \ldots, m.$ (4+6) = [10]

p.t.o.

5. Show that, $f : (S,d) \to (S,d)$ is continuous if and only if $f^{-1}(u)$ is open for each open set u in S. [10]

6. Define $P : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$P(\underline{x}, \underline{y}) = \langle \underline{x}, \underline{y} \rangle$$

(a) Find $P'(a,b)$ and $DP(a,b)$, the Jacobian matrix.

(b) if $f,g : \mathbb{R} \to \mathbb{R}^n$ are differentiable and $h : \mathbb{R} \to \mathbb{R}$ is defined by

$$h(t) = (f(t), g(t)),$$

show that

$$h'(a) = \langle Df(a)^t, g(a) \rangle + \langle f(a), Dg(a)^t \rangle$$

(Note that $Df(a)$ is nx1 matrix, it's transpose $[Df(a)]^t$ is a 1xn matrix)

(c) $f : \mathbb{R} \to \mathbb{R}^n$ is differentiable and $\|f(t)\| = 1$ for all t, show that

$$\langle (Df(t))^t, f(t) \rangle = 0.$$ (8+8+4) = [20]

7. For $f : \mathbb{R}^n \to \mathbb{R}$, let $\partial_{\underline{u}} f(\underline{c})$ denote the directional derivative of f at $\underline{c}$ in the direction $\underline{u}$. Let

$$S' = \{ (x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1 \}.$$

(a) Show that $\partial_{t\underline{u}} f(\underline{c}) = t \, \partial_u f(c)$ for $t \in \mathbb{R}$.

(b) If f is differentiable then $\partial_u f(\underline{c}) = f'(\underline{c})(\underline{u})$

(and in particular
$$\begin{aligned} \partial_{u+v} f(\underline{c}) &= f'(c)(u+v) \\ &= f'(c)(u) + f'(c)(v) \\ &= \partial_u f(c) + \partial_v f(c). \end{aligned}$$)

(c) Let $g : S' \to \mathbb{R}$ be any continuous function satisfying $g(0,1) = g(1,0) = 0$ and $g(-x) = -g(x)$.

Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(\underline{x}) = \begin{cases} \ddot{\underline{x}} \; g(\underline{x}/\|\underline{x}\|) & \underline{x} \neq 0 \\ 0 & \underline{x} = 0 \end{cases}$$

Contd..... 3/-

Contd..... Q.No.7.(c)

Show that f is differentiable at (0,0) implies g is identically zero.

(Hint: Show that Df(0,0) must be zero by first considering (h,k) with h = 0 and then with k = 0)

(:) If $g \neq 0$, then show that f possesses directional derivatives at (0,0) in all directions, however

$$\partial_{\underline{u}+\underline{v}} f(0) \neq \partial_{\underline{u}} f(0) + \partial_{\underline{v}} f(0) \quad \text{for all } u,v.$$

Conclusion: Part (b) is not always true.

(3+6+9+9) = [27]

. bcc:

1987-88 | 521

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) : 1987-88
I Year
PERIODICAL EXAMINATION

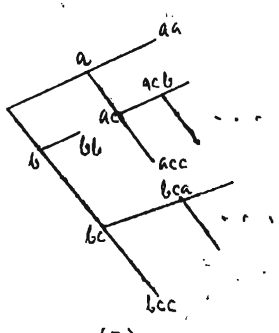
Probability - I

Maximum Marks- 50 Time: $2\frac{1}{2}$ hrs.

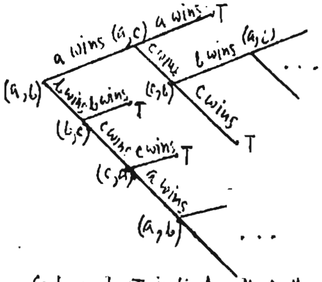
Note: Answer question 5 and three others questions.

Consider Polya's urn scheme. You start with an urn containing b black and r red balls, from which one ball is drawn at random, its colour noted and it is then returned alongwith c more balls of the same colour, and the process is repeated. Show by induction that the probability of a black ball at any trial is $\frac{b}{b+r}$. [13]

2. Three players a, b, c take turns at a game according to the following rules. At the start a and b play while c is out. The loser is replaced by c and at the second trial the winner plays against c while the loser is out. The game continues in this way until a player wins twice in succession, thus becoming the winner of the game. Assume that the game is such that at each trial each of the two players has probability 1/2 of winning. Each of the following two graphs show the successive winners :



(I)



Each node T indicates that the game has terminated.

(II)

Give a sample space together with the probability of each outcome of this random experiment. [13]

p.t.o.

3. Answer any two of the following:

(a) One mapping is selected at random from all the mappings of $\{1, 2, \ldots, n\}$ into itself. Find the probability that exactly k of the elements assume value 2.

(b) Two sets A_1 and A_2 are chosen with replacement from the collection of all subsets of $\{1, 2, \ldots, n\}$. Prove that the probability that A_1 and A_2 are disjoint equals $3^n/4^n$. [Hint. First consider the probability that A_1 has r elements and A_1 and A_2 are disjoint.]

(c) In Fisher's famous example, a lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup. To test this assertion we design the following experiment. We mix 8 cups of tea, 4 in one way and 4 in the other, and present them to the lady for judgement. She is informed of what has been done and asked to taste the 8 cups and then divide the 8 cups into two sets of 4. Assuming that the lady divides the set of 8 cups into two sets of 4 at random, find the probability that she makes six or more corrct guesses. [Note that she makes six correct guesses if and only if she guesses right for 3 cups of each kind]

[12]

4. Answer all parts :

(a) Two dice are thrown r times. Show that the probability that each of the six combinations (1,1), (2,2), ..., (6,6) appears at least once equals

$$\sum_{n=0}^{6} (-1)^n \binom{6}{n} \left(\frac{36-n}{36} \right)^r$$

(b) If the events $A_1, \ldots, A_n$ are mutually independent and if $P(A_i) = p_i$, $i=1, \ldots, n$, find the probability that none of the events $A_1, \ldots, A_n$ occurs.

p.t.o.

(c) A simplified system for inspection of certain articles involves two independent checks. As a result of the k-th checking ($k = 1,2$), an article complying with the standard is rejected with the probability β_k and a defective article is accepted with the probability α_k. An article is accepted when it passes both checks. Find the probability that a defective article is accepted. [14]

5. Consider n independent tosses of a coin with probability of falling head p. Denote by X the random variable giving the number of heads in the first (n-1) tosses and by Y the same for the last (n-1) tosses. Find

(i) the distributions of X and Y

and

(ii) $P(X \neq Y)$. [12]

: SS:

INDIAN STATISTICAL INSTITUTE
M.Stat.(M-stream) : 1987-88

PERIODICAL EXAMINATION

Linear Algebra and Linear Models

31.8.87 Maximum Marks: 100 Time: 3 hours

Note: Answer as many as you can.

1. Let $V = R^n$ and let S be a subset of V. Then

a) define inner product in V.

b) define Orthogonal compliment, $S^{\perp}$.

c) Show that $S^{\perp}$ is a subspace whether S is a subspace or not.

d) Show that $(S^{\perp})^{\perp} \supseteq S$ and equality holds if and only if S is a subspace.

e) Show that $S + S^{\perp} = V$ if S is a subspace.

f) If S and T are two subsets of V both containing null vector $\emptyset$, show that $(S + T)^{\perp} = S^{\perp} \cap T^{\perp}$

g) If S and T are two subspaces of V these show that $(S \cap T)^{\perp} = S^{\perp} + T^{\perp}$.

h) Define Orthogonal projection operator onto the subspace S of V. Show that it can be represented by a matrix P which is (i) symmetric (ii) idempotent and (iii) $M(P) = S$.

[3+2+2+5+2+4+2+10 = 30]

2. Show that

a) $r(AB) = r(B) - d\{M(B) \cap N(A)\}$

b) $r(AB) + r(BC) \leq r(B) + r(ABC)$ [10+5=15]

3.a) Define rank factorization.

b) Show that $r(A+B) = r(A)+r(B) \Longleftrightarrow M(A)\cap M(B) = \{\emptyset\}$ and $M(A')\cap M(B') = \{\emptyset\}$ [2+8 = 10]

4.a) Define spectral decomposition of a real symmetric matrix.

b) Define singular value decomposition of a real matrix. Derive it for a matrix A. [5+10 = 15]

p.t.o.

5. Compute a generalized inverse of

$$\begin{bmatrix} 4 & 3 & 2 & 8 \\ 3 & 4 & 1 & 7 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

[15]

6. Prove or disprove the following statements.

a) Let $\alpha \in M(\alpha_1 \ldots \alpha_u)$. Then α is a unique leniar combination of $\alpha_1 \ldots \alpha_u \iff \{\alpha_1 \ldots \alpha_u\}$ is independent

b) The system of equations $A_{m \times n} x = b$ is consistent for all $b \iff r(A) = m$

c) The system of equations $A_{m \times n} X = b$ has unique solution whenever it is consistent $\iff r(A) = n$.

d) A is full column rank $\iff$ there exists a matrix B such that $BA = I$.

e) A is full row rank $\iff$ there exists a matrix B such that $AC=I$.

f) $M(A) \subsetneq M(I - A) \iff A^2 = A$.

g) $r(ABC) = r(AC)$ if B is non-singular.

h) Every eigen value of a positive definite matrix is positive.

i) Every eigen value of a real symmetric matrix is real.

j) Every eigen value of a real skew symmetric matrix is complex with real part zero.

h) Every eigen value of an idenpotent matrix is either 0 or 1.

k) rank of a triangular matrix = the number of nonzero diagonal elements.

l) B is left inverse of A then AB is idempotent

m) $T_r(AA') = T_r(A^2) \iff A = A'$.

n) $M(A) = M(A') \iff A = A'$.

[15x2 = 30]

:ss: