INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) Part III: 1980-81
PERIODICAL EXAMINATIONS
Differential Equations

Date: 15.9.80 Maximum Marks: 100 Time: 3 hours

Note: Answer any five questions.
Each question carries 20 marks.

1. Solve.
$x'' + 2n \cos \alpha \cdot x' + n^2 x = a \cos nt.$
$x(0) = 0$, $x'(0) = 0$.
Here a, α are given real numbers, n given positive integer.

2. Suppose that L_1, L_2 are two constant coefficient differential operators with characteristic polynomials p_1, p_2 respectively. Assume that p_1, p_2 have no common roots. Let L be the operator with characteristic polynomial $p_1 \cdot p_2$, that is $Lx = L_1(L_2 x)$. Prove that every solution of $Lx = 0$ can be written as a sum $x = x_1 + x_2$ where $L_1 x_1 = 0$ and $L_2 x_2 = 0$.
Show by an example that the above conclusion is false without the assumption on common roots.

3. Let a_1 a_2 be two continuous functions of t, ϕ_1 ϕ_2 be two solutions of

$$x'' + a_1(t)\, x' + a_2(t)\, x = 0$$

[Warning: This is not a constant coefficient equation].
Let W be the Wronskian of ϕ_1, ϕ_2 show that

$$W(t) = W(t_0) \cdot \text{Exp}\left[-\int_{t_0}^{t} a_1(s)\,ds\right].$$

4. Find a differentiable function ϕ on $[0, 4\pi]$ such that $\phi(0) = 1$, $\phi'(0) = 0$ and ϕ satisfies

$$x'' + x = 0 \quad \text{on} \quad 0 \le t \le 2\pi$$
$$x'' - x = 0 \quad \text{on} \quad 2\pi < t \le 4\pi$$

Show that your ϕ is differentiable at 2π. Does $\phi''(2\pi)$ exist?

5.a) Consider $x_1(t) = t^2$, $x_2(t) = t\,|t|$ for $-\infty < t < \infty$. Show that x_2 is differentiable. Is there a linear differential operator L with constant coefficients so that x_1, x_2 are solutions of $Lx = 0$.

b) Calculate the Wronskian of the n functions $1, t, t^2, \ldots t^{n-1}$. Are they Linearly independent over $(-\infty, \infty)$? Give reasons.

p.t.o.

5. Consider the equation with constant coefficients

$$Lx = x^{(n)} + a_1 x^{(n-1)} + \dots + a_n x = 0.$$

a) If all roots of the characteristic polynomial have strictly negative real parts then show that every solution tends to zero as $x \to \infty$.

b) If all roots of the characteristic polynomial have non-positive real parts and those roots with zero real parts have multiplicity one then show that all solutions are bounded on $[0, \infty)$,

----

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

PERIODICAL EXAMINATIONS

Stochastic Processes II

Date: 29.9.80. Maximum Marks: 100 Time: $3\frac{1}{2}$ hours

<u>Note:</u> Answer as many questions as you can. Marks are given in brackets.

1. Consider a birth and death process on $\{0,1\}$ with birth and death rates as follows: $\lambda_0 = \lambda$, $\mu_1 = \mu$. Assume that $\lambda, \mu > 0$.

   a) Write down the forward differential equations.

   b) Hence, or otherwise, determine $p_{ij}(t)$, $i,j = 0,1$, $t \geq 0$.

   c) Find the stationary initial distribution. [7+8+5]=[20]

2. Consider a pure birth process on $\{0,1,\ldots\}$ with birth rates as follows: $\lambda_i = i\lambda$, $i \geq 0$. Assume that $\lambda > 0$. Find $p_{ij}(t)$, $i,j = 0,1,\ldots$, $t \geq 0$. Determine $E(X_t \mid X_0 = i)$ and $\sigma^2 (X_t \mid X_0 = i)$, $i \geq 0$. [12+4+4]=[20]

3. a) Formulate the queuing process $M|M|\infty$ as a birth and death process. Specify the birth and death rates as well as the infinitesimal generator matrix.

   b) Find $p_{ij}(t)$ for $M|M|\infty$.

   c) Prove that $M|M|\infty$ is positive recurrent.

   d) Find the stationary initial distribution for $M|M|\infty$.

   e) Determine $E(X_t \mid X_0 = i)$ and $\sigma^2 (X_t \mid X_0 = i)$, where X_t is the numbers of persons in the 'queue' at time t. [5+10+5+5+5]=[30]

4. Consider the branching process as follows. A collection of particles act independently in giving rise to succeeding generations of particles. Suppose that each particle, from the time it appears, waits a random length of time having exponential distribution with parameter α and then splits into two identical particles with probability p and disappears with probability $1-p$. Assume $0 < p < 1$. Let X_t be the number of particles present at time t.

   a) Formulate the branching process X_t, $t \geq 0$ as a birth and death process by specifying birth and death rates.

   b) Determine the probability of extinction given that there are i particles present at time 0, $i \geq 1$. [5+15]=[20]

p.t.o.

5. Suppose d particles are distributed into two boxes. A particle in box 0 remains in that box for a random length of time that is exponentially distributed with parameter λ before going to box 1. A particle in box 1 remains there for an amount of time that is exponentially distributed with parameter μ before going to box 0. Let X_t denote the number of particles in box 1 at time t.

   a) Determine $p_{id}(t)$ for the X_t - process, $i = 0,1,\ldots,d$.

   b) Find $E(X_t \mid X_0 = i)$, $i = 0,1,\ldots, d$.

   c) Find the stationary initial distribution for the X_t - process. [8+4+3]=[20]

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

Date. 6.10.80 Maximum Marks 100 Time: 3 hours

Answer all questions

1. Let $x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m}$ be i.i.d. $N(\theta, \sigma^2)$, θ is unknown σ^2 is known. Let the prior be $N(\mu, \eta^2)$

(i) Write down the posterior distribution of θ given $x_1, \ldots, x_n$

(ii) Write down the normal approximation to the posterior of θ given $x_1, \ldots, x_n$ as worked out in the class. Compare the mean and variance of the exact and approximation posterior.

(iii) Write down the joint distribution of $\bar{x}_1 = \sum_1^n x_i/n$ and $\bar{x}_2 = \sum_1^m x_{n+i}/m$

(iv) If $n = m = 10$, $\sigma^2 = 1$, $\mu = 0$, $\eta^2 = 25$ and $\bar{x}_1 = 10$, find a 95% (posterior) fonfidence interval for θ and a 95% (posterior) tolerance interval for $\bar{x}_2$.

(v) Prove that for any proper prior, not necessarily normal
$$Var(\bar{x}_2 | \bar{x}_1) \geq V(\theta | \bar{x}_1)$$
4+11+5+10+6 = 36

2. i) Define a location and scale parameter family of densitis. Write down the joint vague prior for the two parameters.

ii) Is $N(\theta, \theta^2)$, $\theta > 0$ a scale parameter family ?

iii) Indicate briefly the invariance argument leading to the vague prior $\pi(\theta) = \theta^{-1}$ for a scale parameter family.

iv) Let $x_1, \ldots, x_n$ be i.i.d. $N(\theta, 1)$ what prior will you choose in the following cases ?

(a) θ is an unknown real number.

(b) θ is between 10 and 20 .

(c) θ is between 10 and 20 and likely to be near the centre of the interval (10,20). 4+2+5+3 = 14

3. i) Define the Kullback-Leibler information numbers.

ii) Define identifiability of a parameter. Is it true that two values θ_1 and θ_2 of a parameter θ are not identifiable (i.e. one gets the same distribution of data under θ_1 and θ_2)

iff I $(\theta_1, \theta_2) > 0$

iii) State a necessary and sufficient condition for the identifiability of the natural parameter of an exponential density.

. 2 + 5 + 3 = 10

4.a) Let x be B(n,p). Assuming a uniform prior write down the posterior density. Hence determine an upper bound p_u such that the posterior probability of p being less than p_u is 0.95 given x = 0. Calculate p_u nemerically for n = 10 and n = 30.

b) A particle starts at time t=0 at a point x=0 on a line. At time t=1 it jumps to x=1 with probability p and to x=-1 with probability 1 - p. At times t = 2,3,...,n it jumps one unit to the right (x increases by 1) with probability p and one unit to the left (x decreased by 1) with probability (1-p). The steps at times t=1,2,3,...,n are independent. Let the position of the particle after n steps be X_n. Suggest a good estimate of p, based on X_n.

c) Suppose x is poisson with mean λ and Y is Poisson with mean μ, X, Y are independent.
Write down the conjugate distribution for λ and that for μ. Assuming λ and μ are independent random variables, write down the joint posterior distribution for λ and μ (given x, Y) and hence write down the posterior distribution of $\lambda/(\lambda+\mu)$.

12 + 6 + 12 = 30

Viva 10

1980-81: 341

INDIAN STATISTICAL INSTITUTE

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

PERIODICAL EXAMINATIONS

Sample Surveys

Date: 27.10.80 Maximum Marks: 100 Time: 3 hours

Note: Answer any four questions each carrying 25 marks. [Marks allotted to different parts of the questions are indicated in the margin].

1. For 14 college students (serial number i) their Mathematics scores (X_i) and Statistics scores (Y_i) in an annual examination are available as (i, X_i, Y_i) with values (1, 84, 71), (2, 67, 58), (3, 73, 91), (4, 64, 59), (5, 60, 71), (6, 36, 44), (7, 71, 80), (8, 61, 78), (9, 53, 41), (10, 65, 52), (11, 43, 37), (12, 61, 56), (13, 70, 76), (14, 59, 58).

   Draw an SRSWOR of size 5 and hence get unbiased estimates for the mean X and mean Y scores for all the 14 students. Also obtain unbiased estimates for the variances of the two estimates and of the covariance between the two estimates. Present a theoretical proof of the unbiasedness of the last-mentioned estimate.

   (4+2+2+4+4+3+6) = [25]

2. Given below are the values of (i, X_i) for 10 students with serial numbers i and their 'standing' heights X_i (in centimeters) as (1, 155), (2, 160), (3, 162), (4, 159), (5, 165), (6, 168), (7, 170), (8, 166), (9, 171), (10, 163). Using these it is intended to estimate the "average sitting height" of these 10 students on taking a sample of size two from these ten and actually measuring the 'sitting heights' (Y_i) for the sampled students. Choose such a PPSWOR sample of size two using X_i as a size-measure and obtain two distinct unbiased estimates from your sample, supposing that the sitting heights of the students you draw are both 145 cm. Indicate which of the two estimates is more efficient and give adequate reasons for your claim.

   (5+5+5+10) = [25]

3. Explain the principles of stratification and clustering. Obtain a formula to unbiasedly estimate (from a stratified sample) the gain due to stratification over a comparable simple random sampling procedure. Indicate a situation where stratified sampling may fare worse than simple random sampling. Establish relevant results to find an optimal stratification point for a population with a uniform density over $(0, \theta)$ required to be split up into two strata with the intention of estimating the population mean from a simple sample of size n with Neyman's optimal allocation.

   (5+8+4+8) = [25]

p.t.o.

4.(a) Show that you cannot obtain a uniformly best estimator for a finite population mean in the entire class of all unbiased estimators. Can you get one in the restricted class of homogeneous linear unbiased estimators ? Give reasons for your answer in details.

(b) Briefly indicate the role of standard errors in inference making in survey sampling.

(5+12+8) = [25]

5. Assuming the population size to be an integral multiple of sample-size obtain a simple measure of efficiency of systematic sampling relative to simple random sampling. Indicate in physical terms a situation when the former may fare better than the latter. Suppose you wish to unbiasedly estimate the total income of 23 households, serially numbered, from a systematically drawn sample with a single random start and with a sampling interval of 4. Give two different procedures for the purpose. Explain why you cannot unbiasedly estimate the variance of your estimate in either case. Also explain how you may do so if you are permitted to take two independent random starts with either method.

(5+5+5+5+5) = [25]

6.(a) Suppose 7 households are serially numbered as i = 1,..., 7. With the ith household (i = 1,..., 7) perform a Bernoullian trial with i/7 as the probability of success. Draw into your sample every household with which you score a success. Explain how you may obtain an unbiased estimator for the average size of these 7 households for such a sample along with an unbiased estimator for the variance of your estimator. What are the average and variance of the effective sample-size for this procedure ?

(b) Give the formulae for Hansen - Hurwitz and Horvitz - Thompson estimators for the PPSWR sampling scheme along with those for the respective unbiased estimators for the variance of each.

(3+2+3+2+5+2+2+3+3) = [25]

INDIAN STATISTICAL INSTITUTE

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

1980-81: 355

Elective-4: Economics-4

PERIODICAL EXAMINATION

Date: 17.11.80 Maximum Marks: 100 Time: 2 hours

Note: Answer two questions, choosing one from each group. Marks are indicated in the margin. Use a separate answer booklet for each group.

GROUP - A

1. Solve the following L.P. problem using the simplex method:

Maximize

$$x_1 - x_2 + x_3 - 3x_4 + x_5 - x_6 - 3x_7$$

Subject to

$$\begin{aligned} 3x_3 + x_5 + x_6 &= 6 \\ x_2 + 2x_3 - x_4 &= 10 \\ -x_1 + x_6 &= 0 \\ x_3 + x_6 + x_7 &= 6 \end{aligned}$$

and $x_j \geq 0$, $j = 1, \ldots, 7$. [50]

2.(a) Consider the following L.P. problem:

Min $p_1 x_1 + p_2 x_2 + \cdots + p_n x_n$

S.t. $a_{i1} x_1 + a_{i2} x_2 + \cdots + a_{in} x_n \geq C_i$, $i = 1, \ldots, m \leq n$

$x_j \geq 0$, $j = 1, \ldots, n$.

Where: p_j = price per unit of food j ;

x_j = quantity consumed of food j ;

a_{ij} = amount of "nutrient" i available per unit of food j ;

C_i = minimum amount of "nutrient" i required by an "average" individual.

Set up the dual to this "diet" problem and discuss in detail the economic interpretation of the dual problem.

[40]

p.t.o.

Contd.....Q.No.2

(b) Prove the following result: Suppose the set of feasible solutions to a given L.P. problem forms a convex polyhedron. Then the objective function assumes its optimum value at an extreme point of the convex polyhedron generated by the set of feasible solutions. If it assumes the optimum at more than one extreme point, then it takes on the same value for every convex combination of those particular values.

[10]

Group - B

3.(a) Describe the Mahalanobis two-sector model. [25]

(b) Discuss the relevance of the assumption about foreign trade in the model. How would the model be affected by a relaxation in the assumption ?

[25]

4. Discuss the features of a planning model. How are they relevant for a developing economy ? Illustrate your answer with the approach of the First Five Year Plan for India.

[50]

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) Part-III: 1980-81
Elective-4: Mathematics-4

1980-81 : 75

PERIODICAL EXAMINATION

Date: 17.11.80 Maximum Marks: 100 Time: 3 hours

Note: All questions carry equal marks.

1. Show that a finite multiplicative subgroup of a field is cyclic.

2. Let H be the parity check matrix of a linear code. Show that the coset whose syndrome is v contains a vector, whose weight is w, iff some linear combination of w columns of H equal v.

3. In a binary linear space $\mathscr{C}$ with $\mathscr{C} \subseteq \mathscr{C}^{\perp}$, there is a basis $\{v_1, \ldots, v_n\}$, where each v_i ($i = 1, \ldots, n$) has weight congruent to zero mod 4. Show that the weight of every element of $\mathscr{C}$ is congruent to zero mod 4.

4. If G is a group of order p^n, p prime, and H is a proper subgroup of G, show that the normaliser of H contains H as a proper subgroup.

5. Define a BCH code of designed distance δ. Show that it has minimum distance atleast δ.

   Give an example where the minimum distance is equal to δ.

6. Construct a one-error correcting cyclic code of length 15. Determine its dimension and the generator matrix.

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

Stochastic Processes - 2

SEMESTRAL-I EXAMINATION

1980-81 : 322

Date: 12.12.80 Maximum Marks: 100 Time: $3\frac{1}{2}$ hours

Note: All questions carry equal marks. For full marks, answer any Six completely.

1. Let $\{X_t, t \geq 0\}$ be a Poisson process with parameter $\lambda > 0$ and $X_0 = 0$. Let n be a positive integer and let r be a positive integer less than n. Let $T = \inf. \{t > 0 : X_t = r\}$. Find the conditional density of T given that $X_s = n$.

2. Let $\{X_t, t \geq 0\}$ be a birth and death process on $\{0, 1\}$ with birth and death rates as follows: $\lambda_0 = \lambda = \mu_1 > 0$. Let N_t be the number of times the system has changed states up to time t, $t \geq 0$. Find the distribution of N_t for a fixed $t > 0$. Does it depend on the initial distribution ?

3. A telephone exchange has m (≥ 1) channels. Calls arrive in the pattern of a Poisson process with parameter $\lambda > 0$; they are accepted if there is an empty channel, otherwise they are lost. The duration of each call is exponentially distributed with parameter $\mu > 0$. The lifetimes of separate calls are independent random variables. Let X_t, $t \geq 0$, denote the number of busy channels at time t. Find the stationary initial distribution of the Markov Chain $\{X_t, t \geq 0\}$.

4. Let $\{X_t, -\infty < t < \infty\}$ be a Gaussian process having mean zero. Let $Y_t = X_t^2$, $-\infty < t < \infty$. Prove that $r_Y(s, t) = 2(r_X(s, t))^2$.

5. Let $\{W_t, -\infty < t < \infty\}$ be Wiener process with parameter σ^2. Set

$$X_t = \int_t^{t+1} (W_s - W_t)ds, \quad -\infty < t < \infty .$$

Prove that $\{X_t, -\infty < t < \infty\}$ is a weakly stationary process. Find its covariance function.

p.t.o.

6. Find the covariance function of the stationary solution on $(-\infty, \infty)$ of the stochastic differential equation $X_t'' + 3X_t' + 2X_t = W_t'$, where W_t' is white noise with parameter σ^2.

7. Let X_t, $0 \leq t < \infty$, be the unique solution of the stochastic differential equation $X_t'' + X_t' = W_t'$ satisfying the initial conditions $X_0 = 0$ and $X_0' = 0$. Here again W_t' is white noise with parameter σ^2. Find explicitly the best linear and nonlinear predictors of $X_{t_1+\tau}$ in terms of $\{X_t, 0 \leq t \leq t_1\}$. Here t_1 and τ are positive constants. Determine the mean square error of prediction in each case.

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

1980-81 : 312

Difference and Differantial Equations

SEMESTRAL-I EXAMINATION

Date: 16.12.80 Maximum Marks: 100 Time: 3 hours

Note: Answer any five questions. Each question carries 20 marks.

1. Solve the following difference Equations:

a) $x_{k+2} - x_{k+1} - 2x_k = k^2$

b) $8x_{k+2} - 6x_{k+1} + x_k = 2^k$

2. Solve

a) $x\,dt = (e^{3t} + 1)\,dx$

b) $t\,dx + x\,dt + tx\,(dt + dx) = 0$

3. Solve

a) $(\text{Cos}\,t\ \text{Cos}\,x - \text{Cot}\,t)\,dt - \text{Sin}\,t\ \text{Sin}\,x\,dx = 0$

b) $(e^x + te^x)\,dt + te^x\,dx = 0$

4. Solve on $(0, \infty)$

$$t^3 x''' + 2t^2 x'' - tx' + x = 0$$

5. State Bessel's Equation. Compute one solution of the Bessel's Equation of order 0. You have to give explicit formula for the coefficients in the power series expansion. Show also that the power series converges on the entire real line.

6. Let I be an interval and $t_0 \in I$. Let b be a continuous function on I. Define $(D-a)^{-k} b$ to be the function

$$(D-a)^{-k} b(t) = e^{at} \int_{t_0}^{t} \frac{(t-s)^{k-1}}{(k-1)!} e^{-as} b(s)\, ds$$

Here a is a fixed number and k is a fixed integer ≥ 1.

a) Show that $(D-a)^k [(D-a)^{-k} b] = b$

b) If Q is a continuous function on I with k continuous derivatives which satisfies

$$Q(t_0) = Q'(t_0) = \dots = Q^{k-1}(t_0) = 0$$

then show that

$$(D-a)^{-k} [(D-a)^k Q] = Q.$$

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

Sample Surveys

SEMESTRAL-I EXAMINATION

Date: 19.12.80 Maximum Marks : 100 Time: 3 hours

Note: Usual symbols and notations will be assumed throughout. Answers should be brief and to the point.
Answer any four questions each carrying 20 marks. Records of practical and home assignments to be submitted in the Examination Hall will carry 20 marks.

(Figures in the margin indicate marks assigned to part questions)

1. (a) From an SRSWOR of size n taken from a finite population of size N the sample means $\bar{y}$, $\bar{x}$ and $\bar{r}$ for variates y, x and $r = \frac{y}{x}$ are calculated and the population total X of x is known. Show that an unbiased estimator for the population total Y of y is given by

$$\bar{r}\,X + \frac{n(N-1)}{n-1}\,(\bar{y} - \bar{r}\,\bar{x})$$

(b) Obtain a sampling design for which the ratio of sample means $\frac{\bar{y}}{\bar{x}}$ is unbiased for the ratio of population means $\frac{\bar{Y}}{\bar{X}}$ for two variates y and x (assume that the values of x are all positive and known).
Explain how you may implement a scheme of sampling corresponding to the design you suggest. From a sample so chosen show how you may unbiasedly estimate the variance of your estimate $\frac{\bar{y}}{\bar{x}}$. [10+3+3+4 = 20]

2. (a) Ignoring f.p.c.'s derive results to reasonably claim the following relationship

$$V_{opt} \leq V_{prop} \leq V_{ran}$$

(Symbols as usual)

(b) If you wish to divide the right triangular population with density f (y)=2(1-y), for $0 \leq y \leq 1$, into two strata show how you may proceed to find the optimal point of stratification in choosing a stratified sample following Neyman's formula so as to estimate the population mean from the stratified sample. [10+10 = 20]

3. (a) For choosing an every k-th systematic sample of size n from a finite population of size N (with $k = \frac{N}{n} \doteq$ an integer) with a random start define (1) s^2_{wsy} = the variance among units that lie within the same systematic sample and (2) $\rho_w \equiv$ the correlation coefficient between pairs of units that are in the same systematic sample, giving formulae for them.

Under what conditions on the natures of the values of s^2_{wsy} and ρ_w respectively can you claim systematic sampling to be more efficient than simple random sampling? Obtain explicit results to justify your claims. [1+1+4+4 = 10]

(b) Suppose from a finite population of N first stage units (fsu) a PPSWR sample of fsu's is chosen in n draws and each fsu so drawn, (repeated or not) is independently sub-sampled taking a systematic sample of second stage units (ssu) from it.

From the survey data based on such a sample chosen, show how you may unbiasedly estimate (i) the population total of ssu' values and (ii) the variance of your estimate in (i). [3+7 = 10]

4. (a) From a given finite population an initial (first phase) SRSWOR is chosen to get information about the unknown numbers of individuals belonging to L (> 2) different pre-defined strata which together constitute the population. From this first phase sample show how one may draw a stratified sample (second phase) and use it to unbiased estimate the population mean. Obtain a formula for the variance of your estimate. [3+2+5 = 10]

(b) Show how you may compare the relative efficiencies of sample mean, ratio estimator and regression estimator (based on an SRSWOR) for a finite population mean when the mean of an auxiliary variate is known. [10]

5. The table below gives the sitting and standing heights of 10 boys serially numbered in increasing order of their standing heights as :

Serial number of boys	Standing Height (in centimetre) (x)	Sitting Height (in centimetre) (y)
1	156.2	131.8
2	157.0	128.4
3	161.6	130.1
4	162.2	129.6
5	162.5	125.2
6	162.7	144.1
7	163.0	139.2
8	164.2	141.6
9	165.5	137.8
10	166.1	135.9

Using the standing heights as size-measures draw (describing your procedure in details) a PPSWOR sample of size two. From this sample obtain Des Raj and symmetrized Des Raj estimates for the average sitting height of these 10 boys. Give a theoretical proof for the . efficiency of the latter estimate.

(4+4+4+8 = 20)

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

1980-81: 332

Inference

SEMESTRAL-I EXAMINATION

Date: 23.12.80 Maximum Marks: 100 Time: 3 hours

Note: Written 90 marks. Viva 10 marks.

1.(A) In each of the following cases which alternative will you regard as true ?

i) A prior is necessary to analyse the data because

(a) we often have a lot of prior information.

(b) we need a prior to get a posterior and make inference conditional on given data. [2]

ii) An improper prior is used because

(a) total uncertainty is infinity.

(b) it gives results similar to those in the Neyman-Pearsonian theory.

(c) it gives rise to a posterior which is the limit of posteriors arising from plausible proper priors. [2]

(B) i) Suppose π is a prior such that

$$1 < \int_{-\infty}^{\infty} \pi(\theta)\, d\theta < \infty .$$ Will you call π improper ?

Justify your answer. [3]

ii) θ is a parameter for which your prior belief is quantified as a normal prior $N(\mu, \sigma^2)$. To get information on θ you may either observe X which is $N(\theta, 1)$ or Y which is $N(\theta, 2)$. Which of X, Y will you observe ? Justify your answer from a Baysian point of view.

(2+6) = [8]

2.(i) Write down Fisher's Inequality defining clearly the class of estimates for which it is valid.

[4]

(ii) State and prove the Cramer-Rao inequality. You must state your assumptions

[4]

Assuming

$$f_\theta(\underline{x}) = e^{c(\theta) + d(\theta)\, T(\underline{x})} \cdot \Psi(\underline{x})$$

and $m(\theta) = E_\theta(T(\underline{x}))$

show that T attains the Cramer-Rao lower bound for $m(\theta)$. [6]

p.t.o.

Contd..... Q.2

(iii) State the Rao-Blackwell theorem.
Give a sufficient condition which makes it easy to find a complete sufficient statistic.

(3+3) = [6]

Let $X_1\ X_2, \ldots, X_n$ be i.i.d. $(n \geq 3)$ with density

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

Find the best unbiased estimate of θ.
Also find a function $m(\theta)$ for which the Cramer-Rao lower bound is attained.

(6+3) = [9]

3.(a) You have a single r.v. X_1 having density

$$\theta . e^{-\theta x}, \quad \theta > 0, \quad x > 0.$$

i) Find the Uniformly Most Powerful test of

$$H_0 : \theta = \theta_0 \quad Vs \quad H_1 : \theta > \theta_0 . \qquad [8]$$

ii) Is the test found in (i) unbiased for the problem of testing

$$H_0 : \theta = \theta_0 \quad Vs \quad H_1 : \theta \neq \theta_0 ? \qquad [4]$$

iii) Write down the form of a Uniformly Most Powerful Unbiased (UMPU) test of

$$H_0 : \theta = \theta_0 \quad Vs \quad H_1 : \theta \neq \theta_0 . \qquad [3]$$

(b) Let $X_1\ X_2, \ldots, X_n$ be i.i.d. $N(\theta, \sigma^2)$, σ^2 known.
Write down the UMPU test of

$$H_0 : \theta = \theta_0 \quad Vs \quad H_1 : \theta \neq \theta_0 .$$

Suppose that $\alpha = .05$ and $\bar{X} = \frac{12\sigma}{\sqrt{n}}$ so that H_0 is rejected.
The client wants to know the probability that H_0 has been rejected wrongly. The statistician tells him the probability is less than .05. Comment on this.

(2+4) = [6]

4.(i) State the conditions under which the maximum likelihood estimate (m.l.e) is a consistent solution of the likelihood equation.

[4]

(ii) Let $X_1, \ldots, X_n$ be i.i.d. with density

$$f_\theta(x) = \frac{1}{2} e^{-|x-\theta|} \quad -\infty < x < \infty, \quad -\infty < \theta < \infty .$$

What is the m.l.e. ?

[2]

p.t.o.

5. Let $X_1, \ldots, X_n$ be i.i.d. with density

$$f_\theta(x) = \frac{1}{\theta} \qquad 0 \le x \le \theta$$
$$= 0 \qquad \text{otherwise}$$

i) Following is a sample of size 10 from $f_\theta(x)$:
1.1, 2.3, 3.8, 7.0, 5.5, 6.2, 1.3, 2.6, 4.9, 4.1.
Draw a rough sketch of the likelihood function and hence or otherwise show that Max $(X_1, \ldots, X_n)$ is the m.l.e.
Find an interval estimate $J = (\hat{\theta}_n, \hat{\theta}_n + d)$ such that if $\theta \notin J$ then the likelihood $L(\theta)$ of θ is less than or equal to $L(\hat{\theta}_n)/20$, where θ_n is the m.l.e. of θ.

(2+1+1) = [4]

ii) Compute the posterior distribution of θ using the vague prior for a scale parameter and hence discuss the reasonableness of the m.l.e.

(1+2) = [3]

iii) Is it true that $E_\theta \left(\frac{d}{d\theta} \log f_\theta(x) \right) = 0$? [1]

iv) Can you find C such that

$$P_\theta \left\{ \hat{\theta}_n < \theta < C.\hat{\theta}_n \right\} = .95 \quad ?$$ [3]

6. The following data gives values of two variables X and Y at 16 time points t:

t	X	Y	t	X	Y
1	2	-.43	9	7	1.27
2	5	1.26	10	8	5.65
3	8	1.26	11	2	2.83
4	3	.11	12	5	2.27
5	6	1.25	13	13	5.96
6	9	3.26	14	10	4.09
7	14	3.03	15	1	1.35
8	11	3.98	16	7	2.82

Let H_{t_o} be the hypothesis that

$$Y = \alpha_1 + \beta_1 X + e_1 \quad \text{for} \quad t \le t_o,\ e_1 \sim N(0, \sigma^2)$$
$$= \alpha_2 + \beta_2 X + e_2 \quad \text{for} \quad t > t_o,\ e_2 \sim N(0, \sigma^2)$$

Contd..... Q.No.6

Which of the two hypothesis will you favour, H_7 or H_8 ?

$\sum_1^7 X_t = 47$, $\sum_1^8 X_t = 58$, $\sum_8^{16} X_t = 64$, $\sum_9^{16} X_t = 53$

$\sum_1^7 Y_t = 9.74$, $\sum_1^8 Y_t = 13.72$, $\sum_8^{16} Y_t = 30.27$, $\sum_9^{16} Y_t = 26.29$

$\sum_1^7 X_t^2 = 415$, $\sum_1^8 X_t = 536$, $\sum_8^{16} X_t^2 = 582$, $\sum_9^{16} X_t^2 = 461$

$\sum_1^7 Y_t^2 = 24.74$, $\sum_1^8 Y_t^2 = 40.58$, $\sum_8^{16} Y_t^2 = 124.85$, $\sum_9^{16} Y_t^2 = 109.01$

$\sum_1^7 X_t Y_t = 95.11$, $\sum_1^8 X_t Y_t = 138.89$, $\sum_8^{16} X_t Y_t = 254.47$, $\sum_9^{16} X_t Y_t$

210.60

[8]

INDIAN STATISTICAL INSTITUTE
B.Stat. Part III, 1980-81
Elective-4: Economics
SEMESTRAL-I EXAMINATION

1980-81 | 354

Date : 26.12.80 Maximum marks 100 Time 3 hours

Note : Answer 3 questions, answering atleast one question from each group.
[Use separate answerscript for each group]

Group - A

1. (a) State and prove the complementary slackness theorem for linear programming. [13]
   (b) Discuss the economic interpretation of this theorem in terms of the problem of comparative advantage in international trade. [20]

2. Prove that the following results are true for any general L.P. problem :
   (a) "If either the primal or the dual problem has a finite optimum solution, then the other problem has a finite optimum solution and the extremes of the linear functions are equal.
   If either the primal or the dual has no optimal solution, then the other has no feasible solutions." [25]
   (b) "If both the primal and its dual have feasible solutions then both have optimal solutions." [8]

3. Illustrate the simplex algorithm for solving an L.P. problem by means of the following transportation problem :

	Transportation costs per ton : Factory		Capacities :	Requirements
	A	B	C	Requirement
Locality 1	Rs. 10.00	Rs. 20.00	Rs. 30.00	25
2	15.00	40.00	35.00	115
3	20.00	15.00	40.00	60
4	20.00	30.00	55.00	30
5	40.00	30.00	25.00	70
Capacity	50	100	150	300

The objective is to find a pattern of shipments from factories to localities that involves the least possible total transportation cost consistent with capacities and requirements. [33]

P.T.O.

Group – B

1. Discuss critically the market conditions assumed for a competitive economy. How does social cost benefit analysis seek to improve upon the conditions ? [32]

2. Why is the market rate of interest used for discounting in profitabil[illegible] analysis of a commercial firm ? Explain why a different rate of inte[illegible] need be considered for cost benefit analysis. [32]

3. State the reasons why world price is advocated for use in cost benefit analysis. How far would you agree with them ? [32]

1980-81:352

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons) Part-III: 1980-81
Elective-4 : Mathematics
SEMESTRAL-I EXAMINATION

Date: 26.12.80 Maximum Marks: 100 Time: 3 hours

Note: Answer any 5. All questions carry equal marks. Max. 100

1. How many non-isomorphic Z modules are there of order 2^5 ? Explain how you would construct them all.

2. Draw the state diagram for a machine with the same input and output alphabet $\{0,1\}$ which gives an output 1 abc.... when ever an input abcd.... is given.

b) Simplify in a Boolean algebra
$[(f+g+h)(f+\bar{g})\bar{h}][fg+\bar{f}h+\bar{g}h+\bar{h}]$

3. a) Show that if $p > q$ are two primes, a group of order pq is solvable

b) If q does not divide p-1 (p,q as in a)) show that a group of order pq is cyclic.

4. If $\varphi_1, \varphi_2, \ldots \varphi_n$ are distinct automorphisms of a finite field F, show that it is impossible to find elements $a_1, \ldots a_n$ in F, not all zero, such that $\sum_{i=1}^{n} a_i \varphi_i(\beta) = 0$ for all β in F.

5. Derive a formula for the number of ireducible polynomials of degree m over GF(p), using Mobius inversion. Clearly state the results, you use.

6. Let α be the cube of a primitive element of $GF(2^6)$. Let $m_j(x)$ denote the minimal polynomial of α^j over GF(2), and let $g(x) = m_1(x)m_3(x)$. Show that the cyclic code generated by $g(x)$ is a BCH code of length 21, dimension 12, and corrects all double errors.

1980–81 : 361

INDIAN STATISTICAL INSTITUTE
B. Stat. (Hons). III year
Multivariate Analysis
PERIODICAL EXAMINATION

2 March 1981 — Maximum marks : 100 — Time : 3 hours

1. The following table gives the mean-vector and the dispersion-matrix of four statistical variables — scores X, Y, Z in three tests and the efficiency-rating U — for a population of industrial workers

variables	mean-vector	dispersion-matrix X	Y	Z	U
X	60	17.42	21.21	18.95	11.20
Y	50		30.28	26.33	15.44
Z	55			31.39	18.26
U	70				19.67

(a) By the forward step-wise procedure, derive the best linear formula for predicting U on the basis of X, Y and Z [35]

(b) Calculate the multiple correlation coefficient of U on X, Y and Z [5]

(c) If the variables that are included as predictors in the first, second and third stages are denoted as 1, 2 and 3 respectively, calculate the partial correlation coefficients $\rho_{U2.1}$ and $\rho_{U3.12}$ [10]

If every pair out of k statistical variables has the same correlation coefficient ρ, show that $\rho \geq -\frac{1}{k-1}$ [10]

If you can detect any inconsistency in the following values of ordinary and multiple correlation coefficients, explain the nature of the inconsistency [10]

$\rho_{12} = -0.62$, $\rho_{13} = +0.31$, $R_{1.23} = +0.58$

4. Given k statistical variables $X_1, \ldots, X_k$, if you are told that the best linear formula for predicting X_i on the basis of the remaining $(k-1)$ variables is $\alpha + \beta S_i$ where S_i is the sum of the remaining $(k-1)$ variables and α and β are the same for every value of $i = 1, \ldots, k$, what can you say about the mean-vector and the dispersion matrix of $(X_1, \ldots, X_k)$? [20]

5. Four statistical variables X_1, X_2, X_3, and X_4 have the structural forms

$$X_1 = 1 + \epsilon_1$$
$$X_2 = 2 + 2\epsilon_1 + \epsilon_2$$
$$X_3 = 3 + 3\epsilon_1 + 2\epsilon_2 + \epsilon_3$$
$$X_4 = 4 + 4\epsilon_1 + 3\epsilon_2 + 2\epsilon_3 + \epsilon_4$$

where ϵ_i's are independent statistical variables each with mean zero and variance unity.

Find the best linear formula for predicting X_4 on the basis of X_1, X_2 and X_3 and the associated multiple correlation coefficient.

[Hint: You may try in two different ways. The straight forward way is to calculate the dispersion matrix. An alternative way is to eliminate ϵ_1, ϵ_2 and ϵ_3 between X_1, X_2, X_3 and X_4 and to use the fact that the residual is uncorrelated with all the predictors] [20 + 10]

Note: The total value of the five questions is 120. You may attempt as many as you like. The maximum you can score is 100.

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

Design of Experiments

PERIODICAL EXAMINATION

1980-81: 371

Date: 9.3.81 Maximum Marks: 100 Time: 3 hours

Note: Answer any four questions. Each question carries 23 marks. 8 marks are alloted to your PRACTICAL RECORDS which must be submitted alongwith your answerscript.

1.(a) State and prove, giving all the details, a necessary and sufficient condition for a connected block design to be orthogonal. Hence show that

$$N^* = \left[\frac{N}{c\,\underline{k}'}\right]$$

is the incidence matrix of an orthogonal design, where N is that of a connected orthogonal design, $\underline{k}$ is the vector of its block-sizes and c is a positive integer.

(b) Give the chain definition of a connected block design and show that every disconnected design is an union of several connected designs. Justify with examples the following statement:

"Incomplete block designs are not always non-orthogonal"

2.(a) Deduce for a general block design with parameters $(v, b, \underline{r}, \underline{k})$ the reduced normal equations for the treatment effects under the usual fixed-effects additive model, and show that the adjusted treatment sum of squares is given by

$$\sum_{i=1}^{v} \hat{\tau}_i Q_i .$$

(b) Show that Rank C $\leq$ v-1, equality holding if and only if all the treatment and block contrasts are estimable. Also prove that $V(\underline{\ell}'\hat{\tau}) = \underline{q}'\underline{\ell}\,\sigma^2$, where $\underline{q}'\underline{Q} = \underline{\ell}'\hat{\tau}$, $\underline{\ell}'\tau$ being any estimable function.

3.(a) Define a balanced incomplete block design (BIBD) with parameters v, b, r, k and λ, and show that (i) $\lambda(v-1) = r(k-1)$, (ii) $b \geq v$, equality holding if and only if any two blocks have λ common treatments.

(b) Obtain the average variance, i.e.,

$$\binom{6}{2}^{-1} \sum_{\substack{i<j \\ i,j=1}}^{6} V(\hat{\tau}_i - \hat{\tau}_j)$$

for the following incomplete block design:

p.t.o.

Contd..... Q.3.(b)

Blocks	Treatments	Blocks	Treatments
1	(1, 5, 2, 6)	4	(4, 5, 2, 6)
2	(4, 6, 3, 5)	5	(1, 4, 5, 6)
3	(1, 3, 5, 6)	6	(2, 3, 5, 6)

[Hint: Recall the side restriction : $\Sigma \hat{\gamma}_i = 0$]

4.(a) Construct the following BIBD:

$$v = 12, \quad b = 22, \quad r = 11, \quad k = 6, \quad \lambda = 5.$$

(b) Give the analysis of variance of a block design obtained from a randomised block design for v treatments in v blocks after deleting treatment i from block i, i = 1, 2, ..., v.

5.(a) Give a method of construction, with proof, of the following series of BIBD:

$$v = b = s^2 + s + 1, \quad r = k = s^2, \quad \lambda = s(s-1),$$

where s is any prime power.

(b) Show that $(x^0, x^2, x^4, \ldots, x^{4n})$ generates a BIBD with the following parameters:

$$v = b = 4n + 3, \text{ a prime power,}$$

$$r = k = 2n + 1, \quad \lambda = n,$$

where x is a primitive element of GF(v).

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

1980-81: 391

Optimisation Techniques

PERIODICAL EXAMINATION

Date: 30.3.81 Maximum Marks: 100 Time: $3\frac{1}{2}$ hours

Note: Answer all questions. All questions carry equal marks.

1. Consider a system with state space $S = \{1, 2\}$ and action space $A = \{1, 2\}$. In state 1 action 1 gives a return of 1 and the system stays in state 1 with probability $\frac{1}{2}$, while action 2 gets a return of $\frac{6}{5}$ and the system stays in state 1 with probability $\frac{1}{4}$. In state 2 either action fetches a return of 0 and the system moves to state 1 with probability $\frac{3}{4}$. Let $f(1) = 1$, $g(1) = 2$. Compute the following quantities:

$$V_\beta(f^{(\infty)}), \quad V_\beta(g^{(\infty)}), \quad x(f), \quad x(g), \quad y(f), \quad y(g).$$

2. Consider the following canonical maximum problem:

$$\begin{array}{ll} \text{Maximize} & 2\xi_1 + 3\xi_2 \\ \text{subject to} & 4\xi_1 + 2\xi_2 + \xi_3 = 4 \\ & \xi_1 + 3\xi_2 = 5 \end{array}$$

Exhibit all the basic feasible solutions. Evaluate the value of the problem.

3. Formulate the general maximum problem and its dual. State and prove the Equilibrium Theorem for the general maximum problem.

4. Consider a system with state space $S = \{1, 2\}$. In state 1 there are three actions 1, 2, 3 with action i yielding a return of $\frac{1}{4}(5-i)$ and the system remaining in state 1 with probability $\frac{1}{2}(i-1)$. In state 2 there is only one action which yields a return of 0 and the system remains in state 2 with probability one. Let $f_i(1) = i$, $i = 1, 2, 3$. For each $i = 1, 2, 3$, find the set of β's for which $f_i^{(\infty)}$ is β - optimal. Find an optimal policy.

5. If A is a $m \times n$ matrix and b is a n-vector, then $xA \leq b$ has a nonnegative solution or $Ay \geq 0$, $by < 0$ have a non-negative solution. Deduce this directly from the Fundamental Duality Theorem.

6. Let f, g be rules. Show that there is $\beta_0 < 1$ and a rule h such that

$$V_\beta(h^{(\infty)}) \geq V_\beta(f^{(\infty)}),\ V_\beta(g^{(\infty)})$$

for all $\beta \in [\beta_0, 1)$.

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

1980-81: 3M1

Elective-5: Mathematics

PERIODICAL EXAMINATION

Date: 13.4.81 Maximum Marks: 100 Time: 3 hours

Note: You can answer any part of any question.

1.(a) Discuss whether the following equations are Elliptic, Hyperbolic, Parabolic, or degenerate.

$$u_{xx} - 3u_{xy} + u_{yy} - u_x = 0$$

$$2u_{xx} + 3u_{xy} + 4u_{yy} + u_x = 0$$

$$u_{xx} - 6u_{xy} + 9u_{yy} + u_x = 0$$

(2+2+8) = [12]

(b) Solve on R^2

$$u_{xx} + u = 0$$

$$u(0, y) = \text{Cos } y$$

$$u_x(0, y) = -\text{Sin } y$$

[8]

2.(a) $f \in C[-\pi, \pi]$, $f(-\pi) = f(+\pi)$, $x_0 \in (-\pi, \pi)$;

f is differentiable at x_0. Then show that

$$\sum_{-\infty}^{\infty} \hat{f}_n e^{inx_0} \quad \text{Converges to } f(x_0).$$

[10]

(b) $f \in C^1[-\pi, \pi]$ $f(-\pi) = f(+\pi)$

Starting from the Fourier series expansion of f exhibit number $a_1, a_2, \ldots ; b_0, b_1, b_2, \ldots$

so that the series

$$\frac{b_0}{2} + \sum_{n=1}^{\infty} [a_n \text{ Sin } nx + b_n \text{ Cos } nx]$$

Converges uniformly to $f(x)$ on $[-\pi, \pi]$.

[10]

3.(a) Suppose f, g are bounded piecewise continuous on $[-\pi, \pi]$.

Show that $\sum |\hat{f}_n \hat{g}_n| < \infty$.

Show that $f.g$ is again bounded piecewise continuous.

p.t.o.

Contd..... Q.No.3.(a)

Show that $\frac{1}{2\pi}\int_{-\pi}^{\pi} fg = \Sigma\, \hat{f}_n \hat{g}_n$ [10]

(b) Show that $\Sigma \frac{\text{Sin } nx}{\log(1+n)}$ is not the Fourier series of any function in $C[0, \pi]$. [5]

(c) Calculate the Fourier Coefficients of the function

$f(x) = \text{Sin } 56x + \text{Cos } 81x \qquad -\pi \leq x \leq \pi$. [5]

4.(a) Solve on R^2 :

$$u_{xx} - 4_{yy} = 1$$
$$u(x, 0) = x$$
$$u_y(x, 0) = x^2$$

Verify that your function is a solution. [10]

(b) Solve on R^2 :

$$u_{xx} - u_{yy} = 0$$
$$u(x, x) = \text{Sin } x$$
$$u(-x, x) = \text{Sin } x$$

Verify that your function is a solution. [10]

5.(a) For the following Eigenvalue problem find all real Eigen values and for each Eigen value all the real Eigen functions.

$$\begin{cases} \varphi'' + \lambda\varphi = 0 & -L < x < L \\ \varphi(L) + \varphi(-L) = 0 \\ \varphi'(L) + \varphi'(-L) = 0 \end{cases}$$

[10]

(b) Show that in the above problem there are no Complex Eigen values. [10]

6. Solve

$$u_t = u_{xx} + 2\beta u_x \qquad 0 < x < L, \; t > 0$$
$$u(0, t) = 0 \quad u(L, t) = 0$$
$$u(x, 0) = f(x)$$

Here $\beta > 0$, $L > 0$ are given. $f \in C^1[0, L]$ and $f(0) = f(L) = 0$. It is enough if you obtain a formal series solution. You do not have to show convergence of the series etc.

[20]

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

1980-81: 381

Nonparametric and Sequential Methods

PERIODICAL EXAMINATION

Date: 20.4.81 Maximum Marks: 100 Time: 4 hours

Note: Answer Q.No 1 or Q.No.2 and all the other questions.

1.(a) Define what is meant by

(i) Stochastically larger.

(ii) One-sided and two-sided alternatives in the context of the two-sample problem.

(b) If X is stochastically larger than Y and X and Y are independent, show that $P\{X \geq Y\} \geq \frac{1}{2}$.

(c) Define the Wilcoxon test for two samples and briefly motivate its use. Show that the one-sided Wilcoxon test is unbiased for one-sided alternatives.

[20]

2.(a) Compare briefly the Wilcoxon and Fisher's t-test for the two-sample problem.

(b) Write down the two sample Wilcoxon statistic as a linear function of a U-statistic and hence find its mean and variance under H_0.

(6+14) = [20]

3. Can you think of an alternative $F \neq G$ in the two-sample context such that the power of the Wilcoxon test can't be made close to one by making n large ?

[5]

4. Explain briefly how one gets the asymptotic distribution of the Wilcoxon signed rank statistic under H_0.

[12]

5.(a) State and prove the Fundamental Identity of sequential analysis.

(b) Suppose X_i's are i.i.d. and $P\{X_i = +1\} = p$, $P\{X_i = -1\} = 1-p$, $p \neq \frac{1}{2}$. Let $b < 0 < a$ be two integers. Let n be the first i such that $(X_1 + \dots + X_i)$ is a or b. Assuming $P\{n < \infty\}$ and using 5(a), calculate $P\{S_n = a\}$.

(13+7) = [20]

p.t.o.

6.(a) The following data gives gain in physical fitness (on a suitable scale) during a six weeks experiment for 12 matched pairs of students. One student from each pair was given usual diet alongwith vitamine B and other was given just usual diet. Carry out the sign test to test the hypothesis H : no treatment effect against the alternative K : the treated student show a larger increase in the fitness score than those not treated (use normal approximation and $\alpha = .05$).

Pair	1	2	3	4	5	6	7	8	9	10	11	12
Control x	8	26	-7	-1	2	9	0	-4	13	3	3	4
Treated y	14	28	2	4	5	14	3	-1	1	6	1	3

(b) In the above problem how many matched pairs should a experimentor choose so that the power of sign test is .9 at the alternative $P(Y > X) = 0.8$ (use normal approximation.)

(6+12) = [18]

7. The following data gives the gain in weight of a control group of young rats and a group of rats of the same age kept in an ozone environment for seven days. Test, at the 5% level of significance, the hypothesis that the ozone does not affect the gain in weight of young rats. Use normal approximation to the hypothesis distribution of Wilcoxon statistic.

x (treat.)	41.0	38.4	29.4	27.2	23.5	21.2	20.1	20.0	19.8	19.5
y (control)	42.2	41.5	40.3	40.0	36.2	34.1	30.1	27.0	19.6	19.2

Ranks of Y's in the combined sample are given to be 20, 19, 17, 16, 14, 13, 12, 9, 3, 1.

[5]

8. Let $X_1\ X_2 \ldots X_n$ be i.i.d. $N(\theta, 2)$.

Find SPRT with $\alpha = .01$ and $\beta = .01$ to test

$$H : \theta = 0 \quad Vs \quad K : \theta = 1$$

Calculate O.C. of above SPRT at $\theta = 0, 0.3, 1$.

[20]

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

1980-81 : 3M1

Differential Equations
(Elective Mathematics)

PERIODICAL EXAMINATION

Date: 2.5.81 Maximum Marks: 100 Time: 3 hours

1. Let f be a bounded continuous function on $[-\pi, \pi]$ $\hat{f}_n$ its n-th Fourier coefficient and

$$s_n = \sum_{-n}^{n} \hat{f}_k \, e^{ikx}$$

Show that $\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f - s_n \right|^2 \longrightarrow 0$ as $n \to \infty$

[15]

2. Calculate the Fourier series expansion of

$$\text{Sin}\,(56 - x) + \text{Cos}\,(x - 56)$$ [5]

3. Suppose f, g are continuous real valued periodic functions on R with period 2π. Define for $x \in R$

$$h(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)\, g(x-y)dy$$

(i) Show g is uniformly continuous on R.

(ii) Using (i) show that h is continuous.

(iii) Show that $\hat{h}_n = \hat{f}_n \hat{g}_n$. (10+10+10) = [30]

4. (i) Solve the boundary value problem on $[0, \pi] \times [0, \infty)$

$$u_{xx} - u_{yy} = 0 \qquad (0, \pi) \times (0, \infty)$$

$$u(x, 0) = \text{Sin}\, x \qquad 0 \le x \le \pi$$

$$u_y(x, 0) = 0 \qquad 0 \le x \le \pi$$

$$u(0, y) = u(\pi, y) = 0 \qquad 0 \le y < \infty .$$

(ii) Justify rigorously that your function is a solution.

(15+5) = [20]

p.t.o.

5.(i) $p \in C^2(R)$ $q \in C^1(R)$ are given. Solve on R^2 :

$$u_{xx} - 64u_{yy} = 0$$
$$u(x,0) = p(x)$$
$$u_y(x,0) = q(x)$$

(ii) Justify rigorously that your function is a solution.

(15+5) = [20]

6.(i) Solve on $[0,1] \times [0,\infty)$

$$u_t - u = u_{xx} - 2u_x \quad \text{on} \quad (0,1) \times (0,\infty)$$
$$u(0,t) = 0 \qquad t \geq 0$$
$$u(1,t) = 0 \qquad t \geq 0$$
$$u(x,0) = e^x \operatorname{Sin} 2\pi x \qquad 0 \leq x \leq 1 .$$

(ii) Justify rigorously that your function is a solution.

(20+5) = [25]

7. Describe all solutions of the following equation on R^2:

$$u_{xx} + 2u_{xy} + u_{yy} = 0 . \qquad [10]$$

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

1980-81: 382

Nonparametric and Sequential Methods

ANNUAL EXAMINATION

16.5.81 Maximum Marks: 100 Time: 4 hours

1. State the MLR property and under this assumption prove the monotonicity of the OC of an SPRT. Give an example where this result can be applied but Wald's approximations can't be used.

(16+4) = [20]

Or

2. Explain briefly what is meant by the optimum property of the SPRT of $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ for i.i.d. observations. Prove the approximate efficiency of the SPRT under suitable assumptions.

Suppose the yield of plot (i) under a certain variety of wheat is a random variable X_i which is $N(\theta, 1)$, and one wants to test $H_0 : \theta \leq \theta_0$ against $H_1 : \theta \geq \theta_1$. (Here θ_0 is the mean value for a variety normally used by cultivators and θ_1 is the value claimed for the new variety under trial). Will you recommend a sequential or a non-sequential test ? Justify your answer.

(16+4) = [20]

3. Write down the two equations of Wald and prove one of them.

Let $\{X_i\}$ be i.i.d., $P\{X_1 = -1\} = P\{X_1 = +1\} = \frac{1}{2}$, $S_m = X_1 + \ldots + X_m$, n = first $m \geq 1$ such that $S_m = 1$. Check whether Wald's equations hold here. (You may assume $n < \infty$ with probability one.)

(14+6) = [20]

4. Describe briefly the Kolmogorov and the Kolmogorov - Smirnov tests. In each case prove that the distribution of the test-statistic does not depend on the true distribution under H_0.

[10]

Or

5. Explain briefly how one constructs a non-parametric confidence interval for the difference in the location parameters of two random variables X and Y. Define the Hodges - Lehmann (point) estimate. (No proof is needed.)

[10]

6. Suppose you have two independent observations X_1 and Y_1 with distribution function $F(x)$ and $G(y)$. It is desired to test $H_0 : F = G$ against Lehmann's alternative $H_1 : G(x) = \{F(x)\}^p$ (where $p > 1$ is known)

(a) Which of X_1 and Y_1 is the stochastically larger of the two under the alternative ?

p.t.o.

Contd..... Q.No.6

(b) Assuming H_0 is true write down the distribution of the rank rector (R_1, R_2).

(c) Obtain the distribution of (R_1, R_2) under H_1 .

(d) Using (b) and (c) show that the two-sample Wilcoxon test of size $\alpha = \frac{1}{2}$ is the most powerful rank test of H_0 against H_1.

(2+2+5+4) = [13]

7. Assignment ... ... [7]

8. In a study of the pollution of Lake Michigan, the number of "odour periods" was observed for each year. The following results for the period 1950 - 1964, were attained.

Year	50	51	52	53	54	55	56	57	58	59	60	61
Number of odour periods	10	20	17	16	12	15	13	18	17	19	21	23

Year	62	63	64
Number of odour periods	23	28	28

It is desired to test the hypothesis that the degree of pollution has not changed with time against the alternative of upward linear trend. State the model, the null hypothesis and the alternative carefully. Suggest a suitable test statistic. Can you identify a feature of the problem which makes the distribution theory taught in the class inapplicable ?

[8]

9. To test whether children who cry more actively as babies later tend to have higher. IQs, a cry count was taken for a sample of children aged five days and was later compared with their Stanford-Bint IQ scores at age three with the results shown below:

Cry count	20	17	15	19	23	14	27	17
IQ score	90	94	100	103	103	106	108	109

Using Spearman's rank correlation coefficient test the hypothesis H of lack of association against the alternative K of positive association between cry counts and IQ scores (take $\alpha = .05$ and use normal approximation). (If you don't know what to do in case of ties ask the invigilator and get one mark deducted).

[12]

p.t.o.

10. The following data give the gains in weight of three groups of rats receiving three types of diet. Using large sample approximation for Kruskal-Wallis statistic test whether there is significant difference between the diets. (take $\alpha = .05$).

diet	gain in weight	Ranks in the combined sample
A	24.0, 25.0, 25.1, 25.7 28.1, 29.0, 29.7, 30.1	2, 7, 8, 11, 22, 25, 27, 30.
B	23.9, 24.2, 24.3, 24.7 25.4, 26.1, 26.6, 27.3, 27.6, 27.8, 30.0, 27.0	1, 4, 5, 6, 10, 12, 14, 16, 18, 19, 20, 29.
C	24.1, 25.2, 26.2, 26.7, 27.2, 28.0, 28.2, 28.7, 29.3, 29.8.	3, 9, 13, 15, 17, 21, 23, 24, 26, 28.

[10]

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

1980-81: 392

Optimisation Techniques

SEMESTRAL II EXAMINATION

19.5.81 Maximum Marks: 100 Time: 4 hours

Note: The paper carries 120 marks. You can score a maximum of 100. Attempt as many questions or parts thereof as you like.

1. Consider in the usual notation a standard maximum linear programming problem and its dual given by the triple (A, b, c). For nonnegative vectors x in $\mathbb{R}^m$ and y in $\mathbb{R}^n$, define

$$\varphi(x, y) = xc + yb - xAy.$$

Let $\bar{x}$, $\bar{y}$ be nonnegative vectors in $\mathbb{R}^m$, $\mathbb{R}^n$, respectively. Prove that $(\bar{x}, \bar{y})$ is a saddle point of φ if and only if $\bar{x}$ is an optimal solution for the maximum problem and $\bar{y}$ is an optimal solution for the dual.

[15]

2. For each $k \geq 1$, consider the canonical maximum problem given by the triple (A, b, c^k) and suppose the problem has value w_k. Consider also the canonical maximum problem given by (A, b, c) with value w. Prove that, if $c^k \to c$ as $k \to \infty$, then $w_k \to w$ as $k \to \infty$.

[15]

3. Maximize $8\xi_1 + 19\xi_2 + 7\xi_3$

subject to

$$3\xi_1 + 4\xi_2 + \xi_3 \leq 25$$

$$\xi_1 + 3\xi_2 + 3\xi_3 \leq 50$$

$$\xi_1, \xi_2, \xi_3 \geq 0.$$

[10]

4. Maximize $\xi_1 + 3\xi_2 + \xi_3$

subject to

$$5\xi_1 + 3\xi_2 \leq 3$$

$$\xi_1 + 2\xi_2 + 4\xi_3 \leq 4.$$

[10]

p.t.o.

5. Associate with a matrix game with pay-off matrix A the following general maximum linear programming problem: Find a nonnegative vector x and a real w such that w is maximum subject to $xA \geq wv$ and $xu = 1$, where u, v are vectors (in the appropriate spaces) with all coordinates unity.

   (a) Prove that solving the game is equivalent (in a sense to be made precise by you) to solving the general maximum problem.

   (b) Deduce from the Fundamental Duality Theorem for the general maximum problem that the game A has a solution in mixed strategies.

(10+10) = [20]

6. Find a solution of the matrix game with pay-off matrix

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$

[8]

7. Players I and II choose simultaneously integers x and y, respectively, between 1 and 4 (both numbers inclusive) and II pays I the amount $|x-y|$. Write down the pay-off matrix. Find a solution of the game. [Hint: Try eliminating some of the pure strategies].

(2+10) = [12]

8. In a dynamic programming problem with finite state and action spaces, let v_n be the optimal (undiscounted) return over n periods of play. Prove that $\lim_{n \to \infty} \frac{v_n}{n}$ exists and equals the optimal average return vector (over an infinite horizon).

[20]

9. Consider a system with state space $S = \{1, 2\}$ and action space $A = \{1, 2\}$. In state 1 action 1 gives a return of 1 and the system stays in state 1 with probability $\frac{1}{2}$, while action 2 gets a return of $\frac{6}{5}$ and the system stays in state 1 with probability $\frac{1}{4}$. In state 2 either action fetches a return of 0 and the system moves to state 1 with probability $\frac{3}{4}$. Describe an optimal policy when the horizon is 2N + 1 periods of play. Prove the optimality of your policy.

[10]

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

1980-81: 362

Multivariate Distribution and Tests

SEMESTRAL II EXAMINATION

23.5.81 Maximum Marks: 100 Time: 3 hours

Note: Answer Question No.1 which carries 40 marks and any two questions from the rest, each of which carries 25 marks. Ten marks are reserved for class room practical records.

1. On the basis of the mean-vector $\bar{x} \equiv (\bar{x}_1, \ldots, \bar{x}_p)$ and the corrected sums of products matrix $S \equiv ((s_{ij}))$ $i, j = 1, \ldots, p$ computed from a random sample of size n from a p-variate normal population with mean-vector $\mu = (\mu_1, \ldots, \mu_p)$, $\mu_i \neq 0$ $i = 1, \ldots, p$ how will you test the hypothesis that

$$\mu_1 : \mu_2 : \quad : \mu_p = C_1 : C_2 : \quad : C_p$$

where C_i's are given real numbers ?

[Hint: This is equivalent to testing the hypothesis that $\mu T = 0$ where T is a suitably chosen $p \times (p-1)$ matrix]

The following table gives the sample mean and the estimated dispersion matrix (corrected sums of products matrix divided by the degrees of freedom) of three anthropometric characters:

x_1 = maximum head length

x_2 = maximum head breadth

x_3 = frontal breadth

of Muslims of Nadia, based on a sample of size n = 50.

variables	means	estimated dispersion matrix x_1	x_2	x_3
x_1	182.19	41.30	5.47	7.48
x_2	139.80		26.89	10.65
x_3	102.37			17.81

Are these data consistent with the hypothesis that the mean values of x_1, x_2 and x_3 are in the ratio 4 : 3 : 2 ?

(10+30) = [40]

p.t.o.

2.(a) Explain the concepts of principal components of a random vector and that of canonical correlations between two random vectors. [Proofs are not required].

(5+8) = [13]

(b) $(Y_1, \ldots, Y_p)$ and $(Z_1, \ldots, Z_q)$ are random variables such that the coefficient of correlation between any two Y's is ρ_{yy}, that between any two Z's is ρ_{zz} and that between any Y and any Z is ρ_{yz}. Starting from first principles, obtain one linear compound of $Y_1, \ldots, Y_p$ and another linear compound of $Z_1, \ldots, Z_q$ such that the coefficient of correlation between these two linear compounds is a maximum. (First show that there is no essential loss of generality in assuming that each of the (p+q) random variables has mean zero and variance unity).

(2+10) = [12]

3. A random vector $Y = (Y_1, \ldots, Y_p)$ is said to follow a multivariate normal distribution if for every non-null real vector $\underline{\lambda} = (\lambda_1, \ldots, \lambda_p)$ the random variable $\lambda_1 Y_1 + \ldots + \lambda_p Y_p$ follows a univariate normal distribution with positive variance. Starting from this definition

(a) Obtain the joint probability density function of a multivariate normal distribution and show that it involves only the mean-vector and the dispersion matrix as parameters,

[8]

(b) Calculate the moment generating function both from the definition and from the joint probability density function derived in (a) above and hence show that the joint probability density function is uniquely determined,

[8]

(c) find the marginal distribution of $(Y_1, \ldots, Y_{p_1})$ where $p_1 < p$, and

[4]

(d) find the conditional distribution of $Y_{p_1+1}, \ldots, Y_p$ given $Y_1, \ldots, Y_{p_1}$.

[5]

4.(a) If the elements of the p x p symmetric matrix S follow the Wishart distribution with ν degrees of freedom and parameter matrix Σ, and $X(: 1 \times p)$ follows the p-variate normal distribution with mean vector μ and dispersion matrix Σ and X and S are independent, work out the probability density function of

(i) $T_0^2 = X \Sigma^{-1} X'$ and (ii) $T^2 = X (S/\nu)^{-1} X'$

indicating in each case the simplification when $\mu = 0$.

Contd..... Q.No.4.(a)

You are not required to prove any properties of the Wishart distribution that you may need in the derivation but you must state these properties briefly and clearly.

(9+9) = [18]

(b) How can you use this result in deriving an appropriate statistic for testing, on the basis of samples from two p-variate normal populations with a common but unknown dispersion matrix, the hypothesis that their mean vectors are equal ?

[7]

5.(a) An individual is known to belong to one, but not exactly which, of k different populations, with an apriori probability $\pi_i > 0$ of belonging to the i-th population, $i = 1,\ldots,k$, $\pi_1 + \ldots + \pi_k = 1$. Measurements on p characteristics of the individual denoted by $X = (X_1, \ldots, X_p)$ are available, and it is known that the probability density function of X in the i-th population is $f_i(x)$; $i = 1, \ldots, k$. The problem is to use the measurements X to assign the individual to one of these k populations. The loss in wrongly assigning the individual to the j-th population when it actually belongs to the i-th population in $\ell_{ij} > 0$ for $i \neq j = 1, \ldots, k$. Show that a best procedure, in the sense of minimising the expected loss, is to assign the individual to the j-th population if j is the smallest index for which $F_j(x)$ is the minimum of $F_1(x), \ldots, F_k(x)$ where

$$F_j(x) = \sum_{i=1}^{k} \pi_i \ell_{ij} f_i(x)$$ [15]

(b) Reduce this procedure to its simplest form when $k = 2$, $\pi_1 = \pi_2 = 1/2$, $\ell_{12} = \ell_{21} = \ell$ and the density function is p-variate normal with the same dispersion matrix but possibly different mean-vectors in the two populations.

[10]

1980-81: 3E2/482/573

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	
	cerealy	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.) III Year: 1980-81

1980-81: 312

Elective-5 : Mathematics

SEMESTRAL II EXAMINATION

27.5.81 | Maximum Marks: 100 | Time: 3 hours

Note: You can answer any part of any question.

1.(a) Calculate the Fourier Sine series expansion of

$$f(x) = 1 - \cos 2x \qquad 0 \le x \le \pi$$

(b) Explicitly calculate solution to the problem on $[0, \pi] \times [0,\infty]$:

$$u_{xx} - c^2 y_{yy} = 0 \qquad \text{on } (0, \pi) \times (0, \infty)$$

$$u(0, y) = u(\pi, y) = 0 \qquad 0 \le y < \infty$$

$$u(x, 0) = 1 - \cos 2x \qquad 0 \le x < \infty$$

$$u_y(x, 0) = 0 \qquad 0 < x < \infty$$

[10+15) = [25]

2.(a) Describe the Laplacian in Polar Coordinates.

(b) $f(x)$ and $g(x) \in C^1(R)$ periodic with period 2π. Explain a method of solving the Dirichlet problem for the region $a \le r \le b$ with the boundary functions

$$u(a\,e^{i\theta}) = f(\theta) \quad \text{and} \quad u(b\,e^{i\theta}) = g(\theta).$$

(10+10) = [20]

3.(a) Suppose u is a positive continuous function on the closed disc $\bar{S}(\underset{\sim}{0}, \rho)$

and harmonic in the open disc $S(0, \rho)$. Suppose P is a point

$P = re^{i\theta}$ $r < \rho$. Show that

$$\frac{\rho - r}{\rho + r}\, u(0) \le u(P) \le \frac{\rho + r}{\rho - r}\, u(0).$$

(b) Suppose u is a positive harmonic function on R^2. Using (a) show that u is a constant.

(15+10) = [25]

p.t.o.

4. Solve for $u(x, t)$ on $[0, \pi] \times [0, \infty)$:

$$u_t = u_{xx} + ru \quad \text{on} \quad (0, \pi) \times (0, \infty)$$

$$u(0, t) = u(\pi, t) = 0 \quad \text{for} \quad t \geq 0$$

$$u(x, 0) = \sin 2x + \sin 5x \qquad 0 \leq x \leq \pi .$$

[10]

5.(a) Explain whether the following equation is parabolic or dogenerate :

$$u_{xx} - 2u_{xy} + u_{yy} = 0.$$

(b) Solve by using separation of variables technique

$$x^2 u_{xx} + y^2 u_{yy} + 5x u_x - 5y u_y + 4u = 0.$$

(8+12) = [20]

6.(a) If $F(a, b, c) = \varphi(a) c^2 + 2\psi(a) bc + \phi(a) b^2$ then the show that the Euler's Equation in the problem of minimizing

$\int_{x_1}^{x_2} F(x, y(x), y'(x))\, dx$ is a second order linear differential equation.

(b) Find the Curve joining the points $(x_1, y_1)(x_2, y_2)$ that yields a surface of revolution of minimum area when revolved about the X-axis.

(10+10) = [20]

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

1980-81: 372

Design of Experiments

SEMESTRAL II EXAMINATION

30.5.81 Maximum Marks: 100 Time: 3 hours

Note: Submit your practical records alongwith your answer-scripts. 12 marks are alloted for the practical records. Marks alloted to a question are indicated in brackets [ ]. Answer any four questions.

1. Show that under a missing plot situation, (a) $SSE_{\Omega*(\underline{w}*)} = SSE_{\Omega}$ where $\underline{w}*$ is any solution of $\underline{w}* = Z\hat{\underline{\beta}}_{\Omega*(\underline{w}*)}$, and that (b) $\hat{\underline{\beta}}_{\Omega*(\underline{w}*)} = \hat{\underline{\beta}}_{\Omega}$, where the symbols have their usual significances. Hence obtain the estimates for missing values in cells (i, j) and (i', j'), $i \neq i' \in (1, 2, \ldots, v)$, $j \neq j' \in (1, 2, \ldots, r)$ in a randomised block design for v treatments in r blocks. Give also the expressions for $V(\hat{\tau}_i - \hat{\tau}_j)$, $i \neq j \in (1, 2, \ldots, v)$ for the resulting incomplete block design.

(5+5+4+8) = [22]

2. Develop the analysis of covariance under the usual fixed effects model $\tilde{\Omega}$: $\underline{y} = X\underline{\beta} + H\underline{\gamma} + \underline{e}$, etc. Apply the analysis to a latin square design with one concomitant variable. Obtain also $V(\sum_{i=1}^{v} l_i \hat{\tau}_i)$ under $\tilde{\Omega}$ for this design.

(10+8+4) = [22]

3.(a) Give a balanced confounding scheme for a $(2^5, 2^3)$ experiment, and construct a replication of the suggested design, indicating only how the other replications can be constructed.

(b) Identify all the confounded effects in one replication of a $(2^8, 2^4)$ experiment, of which the following constitute a block:

efg, bgh, abefh, cfh, bdf, bce, a, bcdefgh,

acegh, deh, cdg, adfgh, abcfg, abcdh, abdeg,

acdef.

(6+8+8) = [22]

p.t.o.

4.(a) Define the main effects and interactions for a 3^3 experiment and generalise them to those of a 3^n experiment. Show that any contrast belonging to an effect is orthogonal to any contrast belonging to another effect in a 3^n factorial.

(b) Construct a confounded design for a $(3^3, 3^2)$ experiment in two replications, and give the appropriate analysis of variance for the suggested design.

(4+4+8+6) = [22]

5. Write short notes on any <u>two</u> of the following:

(i) analysis of experiments with missing data using covariance technique;

(ii) method of differences for the construction of BIBD;

(iii) confounding in designs for factorial experiments.

(11+11) = [22]

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

1980-81: 322B

Stochastic Processes - 2

BACK-PAPER EXAMINATION

Date: 6.7.1981 Maximum Marks: 100 Time: 3 hours

Note: Answer all questions. All questions carry equal marks.

1. There are two transatlatic cables each of which can handle one telegraph message at a time. The time-to-breakdown for each has the same exponential distribution with parameter λ. The time to repair for each cable has the same exponential distribution with parameter μ. Given that at time 0 both cables are in working condition, find the probability that, if at time $t > 0$ two messages arrive simultaneously, they will find both cables operative.

2. Let $\{X_t, t \geq 0\}$ be the pure death process on $\{0,1,2,\ldots\}$ with death rate $\mu_i = i\mu$, $i \geq 0$. Here μ is a positive constant. Find $p_{ij}(t)$.

3. Find the mean and covariance functions of the process $X_t = W_t - t W_1$, $0 \leq t \leq 1$, where $\{W_t\}$ is Wiener process with parameter σ^2.

4. Let $\{V_t, t \geq 0\}$ be defined as the unique solution of the stochastic differential equation

$$m V_t' + f V_t = W_t'$$

satisfying the initial condition $V_0 = v_0$. Here W_t' is white noise with parameter σ^2, m, f are positive constants. Find the mean and covariance functions of the process $\{V_t\}$. What is the process called ?

5. For the above process, find the best linear and nonlinear predictors of $V_{t_1+\tau}$ in terms of $\{V_t, 0 \leq t \leq t_1\}$. Here t_1 and τ are positive constants. Determine the mean square error in each case.

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

1980-81: 332B

Inference

BACK-PAPER EXAMINATION

Date: 7.7.81 Written: 90 marks Viva: 10 marks

Time: 3 hours

1.(a) State the result proved in the class on normal approximation to the posterior. State all your assumptions clearly. [5]

(b) Suppose the posterior of θ is approximately normal with mean $\hat{\theta}$ and variance $(nI(\hat{\theta}))^{-1}$. What would be the approximate posterior distribution of $g(\theta)$, where $g(\theta)$ is continuously differentiable with a non-vanishing derivative ? (You don't have to justify your answer). [3]

(c) Under suitable regularity conditions a Neyman-Pearsonian uses $\hat{\theta} \pm 1.96 \frac{1}{\sqrt{nI(\hat{\theta})}}$ as an approximately 95% confidence interval for θ ; so does a Baysian. Explain briefly the difference in their interpretations. [5]

2. Give an example where the m.l.e. of a parameter is not consistent but a consistent estimate exists. Justify your answer. [15]

Or

Under suitable regularity conditions prove that a consistent solution of the likelihood equation is asymptotically normal. [15]

3.(a) Prove the Cramér-Rao inequality. [9]

(b) Prove the uniqueness of the best unbiased estimate of an estimable function. [6]

(c) Describe briefly how one uses the Rao-Blackwell theorem to get a best unbiased estimate. [9]

(d) $X_1, \ldots, X_n$ are i.i.d. $N(\theta, 1)$. Give an example of an estimable function for which the best unbiased estimate does not attain the Cramér-Rao lower bound. [9]

4. Let $X_1, \ldots, X_n$ be i.i.d. $N(0, \sigma^2)$. Find the UMP test of $H_0 : \sigma^2 = 1$ Vs $H_1 : \sigma^2 > 1$. [15]

p.t.o.

5. The following data give values of two variables X and Y at 18 time points t:

t	X	Y	t	X	Y
1	1	.45	10	11	3.98
2	4	1.39	11	7	1.27
3	2	-.43	12	8	5.65
4	5	1.26	13	2	2.88
5	8	1.26	14	5	2.27
6	3	.11	15	13	5.96
7	6	1.25	16	10	4.09
8	9	3.26	17	1	1.35
9	14	3.03	18	7	2.82

Let H_{t_0} to be the hypothesis that

$$Y = \alpha_1 + \beta_1 X + e_1 \quad \text{for } t \leq t_0$$
$$= \alpha_2 + \beta_2 X + e_2 \quad \text{for } t > t_0 ,$$

$e_1 \sim N(0, \sigma^2)$ $e_2 \sim N(0, \sigma^2)$

Which of the two dypothes's will you favour H_9 or H_{10} ?

$\sum_1^9 X_t = 52$ $\sum_1^{10} X_t = 63$ $\sum_{10}^{18} X_t = 64$ $\sum_{11}^{18} X_t = 53$

$\sum_1^9 Y_t = 11.58$ $\sum_1^{10} Y_t = 14.84$ $\sum_{10}^{18} Y_t = 30.27$ $\sum_{11}^{18} Y_t = 26.29$

$\sum_1^9 X_t^2 = 432$ $\sum_1^{10} X_t^2 = 553$ $\sum_{10}^{18} X_t^2 = 582$ $\sum_{11}^{18} X_t^2 = 461$

$\sum_1^9 Y_t^2 = 26.88$ $\sum_1^{10} Y_t^2 = 42.72$ $\sum_{10}^{18} Y_t^2 = 124.83$ $\sum_{11}^{18} Y_t^2 = 109.01$

$\sum_1^9 X_tY_t = 101.12$ $\sum_1^{10} X_tY_t = 144.90$ $\sum_{10}^{18} X_tY_t = 254.47$

$\sum_{11}^{18} X_tY_t = 210.69$

[14]

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

1980-81: 342B

Sample Surveys

BACK-PAPER EXAMINATION

Date: 8.7.81 Maximum Marks: 100 Time: 3 hours

Note: Usual symbols and notations will be assumed throughout. Answers should be brief and to the point.

Answer any four questions each carrying 20 marks. Records of practical and home assignments to be submitted in the Examination Hall will carry 20 marks. Figures in the margin indicate marks assigned to part - questions.

1.(a) For SRSWR in n draws from a finite population of N units how many 'distinct units' do you expect to find on an average ? Obtains the variance of this number in terms of n and N. (Give proofs in detail).

(b) Write down the formula for the Yates-Grundy estimator for the variance of Horvitz-Thompson estimator for a finite population total. Show that it is unbiased if the sample is chosen according to PPSWOR scheme but not necessarily so if the sample is chosen by PPSWR method.

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

2.(a) Define 'separate ratio estimator' and 'combined ratio estimator' for a finite population total based on a stratified random sample. Using large sample approximations obtain usual variance formulae for these estimators. From the consideration of bias alone, indicate which one of these two estimators you may use more safely in practice and why.

(b) If there are two strata only of sizes N_1, N_2 with variances S_1^2, S_2^2, then (ignoring f.p.c.) show that

$$V_{opt}(\bar{y}_{st})/V(\bar{y}_{st}) \geq 4\phi/(1+\phi)^2$$

where $\phi = \dfrac{n_1}{n_2} \Big/ \dfrac{n_1(opt)}{n_2(opt)}$, (symbols are usual ; 'opt' means Neyman's optimal rule).

(2+2+3+3+2+8) = [20]

3. What do you mean by a uni-cluster sampling design ? Show that for such a design a minimum variance (MV) estimator for a finite population total is available within the class of homogeneous linear unbiased estimators (HLUE). Also show that in the HLUE class no MV estimator exists if the design is not 'uni-cluster'. Again show that in the class of all unbiased estimators (UE) no MV estimator is available for any design which does not choose the entire population with probability one.

(2+8+6+4) = [20]

p.t.o.

4. For two-stage sampling establish the formula

$$V = V(\bar{y}) = \frac{N-n}{N}\frac{S_1^2}{n} + \frac{M-m}{M}\frac{S_2^2}{nm}.$$

(Symbols as usual and one to be explained by you).

Obtain a formula for an unbiased estimator for V for an appropriately chosen two-stage sampling procedure.

Assuming the usual simple cost function

$$C = C_1 n + C_2 n m_1$$

obtain the formula for optimum choice of m as m_{opt} to minimize V for a given C or vice versa in the form

$$m_{opt} = \frac{S_2}{\sqrt{S_1^2 - S_2^2/M}} \Big/ \frac{C_1}{C_2}.$$

What will be the corresponding choice of optimal n ?

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

5. The table below gives the information about whether TV sets are owned (Scored '1' if 'yes' and '0' if 'no') by the households (hh) serially numbered in five serially numbered buildings (B_i) in a city-block (No hh owns more than one TV set).

B1	TV	B2	TV	B3	TV	B4	TV	B5	TV
1	1	1	0	1	0	1	1	1	0
2	0	2	0	2	1	2	1	2	0
3	0	3	1	3	1	3	0	3	0
4	0	4	1	4	0	4	1	4	1
5	1	5	0	5	1			5	0
6	0			6	0			6	1
7	1								
8	1								
9	1								

Draw an SRSWOR of 2 buildings and from each selected building independently take an SRSWOR of 2 households. For the selected households note whether TV sets are owned or not and hence estimate unbiasedly the proportion of households, in the city-block, owning TV sets, obtain an unbiased estimate of the variance of your estimate. Note the error in your estimate.

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

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year : 1980-81

1980-81: 362B

Multivariate Distribution and Tests

BACK-PAPER EXAMINATION

Date: 9.7.81 Maximum Marks: 100 Time: 3 hours

Note: Answer Question No.1 which carries 40 marks and any two questions from the rest, each of which carries 25 marks. Ten marks are reserved for class-room practical records.

1.(a) How will you test, on the basis of samples from two p-variate normal populations with a common but unknown dispersion matrix, the hypothesis that the two mean vectors are identical ? Write down the statistic that you would use, and the name of the standard distribution which has to be referred to. No proof is required. [10]

(b) The table below given the estimates of the means and the common dispersion matrix of three characters:

x_1 = length of hind femur

x_2 = maximum width of head in the genal region

x_3 = length of pronotum at the peel

for two groups of female desert locusts one in the phase gregaria and the other in an intermediate phase between gregaria and solotaria. The estimate of the common dispersion matrix has been obtained by adding the two corrected sum of products matrices within the groups and dividing it by the sum of the two degrees of freedom.

	means		estimated dispersion matrix		
	phase : Gregaria sample size n_1 = 20	phase: intermediate sample size n_2 = 72	based on n_1+n_2-2 = 90 df		
character			x_1	x_2	x_3
x_1	25.80	28.35	4.7350	0.5622	1.4685
x_2	7.81	7.41		0.1413	0.2174
x_3	10.77	10.75			0.5702

Test the significance of the difference in the two mean-vectors. [30]

2.(a) Given two random vecters X_1 (: 1 x p_1) and X_2 (: 1 x p_2) with mean vecter $\mathcal{E}[X_1 \vdots X_2] = [\mu_1 \vdots X_2]$ and dispersion-matrix

$$\mathcal{D}(X_1 \vdots X_2) = \begin{bmatrix} \Sigma_{11} & \vdots & \Sigma_{12} \\ \Sigma_{21} & \vdots & \Sigma_{22} \end{bmatrix}$$, how would you predict the

p.t.o.

Contd..... Q No.2.(a)

value of X_2 when the value x_1 of X_1 is known, on the basis of a linear function of x_1 ? What good properties does this predictor possess? How would you examine whether using x_1 linearly is at all necessary in the case where $p_2 = 1$? [Proofs are not required] (4+6+5) = [15]

(b) Three random variables Y_1, Y_2 and Y_3 have the structural form

$$Y_1 = 1 + \xi_1$$

$$Y_2 = 3 + 2\xi_1 + \xi_2$$

$$Y_3 = 5 + 3\xi_2 + 2\xi_2 + \xi_3$$

where ξ_1, ξ_2 and ξ_3 are unobservable independent random variables, each with expected value zero and variance unity. Calculate the best linear function for predicting Y_3 on the basis of Y_1 and Y_2, the multiple correlation coefficient of Y_3 on Y_1 and Y_2 and the partial correlation coefficient between Y_1 and Y_3 eliminating Y_2. (4+3+3) = [10]

3. Let X_λ (:1 x p) for $\lambda = 1, \ldots, n$ be distributed independently and identically in the p-variate normal form with mean vector μ and dispersion matrix Σ. Let $\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$ and $S = (X_1 - \bar{X})'(X_1 - \bar{X}) + \ldots + (X_n - \bar{X})'(X_n - \bar{X})$ denote respectively the sample mean-vector and the corrected sums of products matrix.

(a) Show that $\sqrt{n}\,(\bar{X} - \mu)$ follows the p-variate normal distribution with mean vector zero and dispersion matrix Σ. [10]

(b) Show that the probability distribute of S is the same as that of $T = Y_1' Y_1 + \ldots + Y_{n-1}' Y_{n-1}$ where Y_λ(:1 x p) for $\lambda = 1, \ldots, (n-1)$ are independently and identically distributed in the p-variate normal form with mean vector zero and dispersion matrix Σ. [10]

(c) Show further that $\bar{X}$ and S are independent. [5]

4. Let X be a random vector with p components. The expected value of X in the i-th of two populationsis μ_i i = 1, 2 but the dispersion - matrix is the same, Σ, in both the populations.

p.t.o.

Contd..... Q.No.4

(a) Find a vector $\underline{\lambda}$ with p real components such that the square of the difference of the expected values of $L \equiv \lambda X'$ in the two populations divided by the common variance of L is maximum. Call this maximum value Δ^2.

[10]

(b) Give an algorithm for computing L and Δ^2. [8]

(c) Write a note on the problem of classification of an individual into one of two populations on the basis of multiple observations on the individual and explain the roles played by L and Δ^2 in this connection. [7]

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

1980-81: 372B

Design of Experiments

BACK-PAPER EXAMINATION

Date: 10.7.81 Maximum Marks: 100 Time: 3 hours

Note: Answer any four questions. Marks alloted to a question are indicated in brackets [ ].

1. Define a balanced incomplete block design (BIBD) with parameters (v, b, r, k, λ). Show that for such a design $b \geq v$, with equality holding if and only if any two blocks have λ treatments common. If for a BIBD, $b = v$ = even, show that $(k - \lambda)$ is a perfect square. Give the analysis of variance of a BIBD (v, b, r, k, λ). (2+8+5+10) = [25]

2. What is missing plot technique ? Explain its application to a latin square design with one missing observation. (5+10+10) = [25]

3. Explain confounding with an example. Give a balanced confounded design for a $(2^4, 2^2)$ experiment in 3 replications. Give also the appropriate analysis of variance for your design using the method of sums and differences due to Yates. (5+8+12) = [25]

4. (a) Identify all the confounded effects in one replication of a $(3^6, 3^3)$ experiment of which the following are the three independent treatment combinations of the key block:
(100001), (010111) and (001121).

(b) Give the analysis of variance of a $(3^n, 3^k)$ confounded design in r randomized blocks. (10+15) = [25]

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

(i) connectedness of a block design ;
(ii) construction of mutually orthogonal latin squares ;
(iii) analysis of covariance.

$(12\frac{1}{2} + 12\frac{1}{2})$ = [25]