INDIAN STATISTICAL INSTITUTE

B.STAT.(HONS) III YEAR BACKPAPER EXAMINATION 1994-95

SEMESTRAL - I EXAMINATION

STATISTICAL INFERENCE II

Date: 26.6.95 Maximum Marks:100 Time: 3 Hours

Note: You may use any result done in class either proved or unproved. Be sure to state them clearly. The paper carries 120 marks and the maximum you can score is 100.

1. Suppose $X_1, X_2, \ldots, X_n$ are iid observations from $N(\mu, \sigma^2)$. The parameter of interest is the density ϕ_{μ,σ^2}.

   (i) If $n = 1$ show that there is no unbiased estimator.

   (ii) If $n \geq 2$, show that $X_n - \bar{X}_n$ and $\bar{X}_n$ are independent.

   (iii) If $n \geq 2$, using (ii) or otherwise find an estimator $f(\cdot)$ such that $Ef(x) = \phi_{\mu,\sigma^2}(\cdot)$ for all x. [10+5+10 = 25]

2. Suppose $X_1, \ldots, X_n$ are i.i.d standard Cauchy random variables. Let $X_{(k,n)}$ be the k^{th} order statistic. Decide whether $E(X^2_{(k,n)}) < \infty$ for any k. What is the limiting distribution of the sample median as $n \to \infty$? State clearly any result that you are using. [20]

3. Consider the kernel density estimate where the kernel is of order p. Under suitable assumption derive the *AMISE* and the optimal bandwidth. [10]

4. Suppose $Y_i = X_i\beta + \epsilon_i$ $i = 1, 2, \ldots n$ (are) estimate β by minimising $\sum_{i=1}^{n} |Y_i - X_i\beta|$. Show that this can be written as a linear programming problem. [10]

5. Suppose $X_1, \ldots, X_n$, are iid observations from the distribution $F(x) = \frac{1}{2}\Phi(x) + \frac{1}{2}\Phi(\frac{x}{\tau})$ where Φ is the d.f. of $N(0,1)$. Consider the two estimators $\bar{X}_n$ = sample mean and $\tilde{X}_n$ = sample medium and compare then in terms of relative efficiency. State clearly any result that you are using. [20]

6. Derive the Sequential Cramer Rao inequality. State clearly any assumptions that you are making and any result that your are using. [20]

7. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$ where μ, σ^2 are both unknown. Suppose $\theta_1 = \theta_1(X_1, \ldots, X_n)$ and $\theta_2 = \theta_2(X_1, \ldots, X_n)$ are such that

   $\inf_{\mu,\sigma^2} P_{\mu,\sigma^2}\{0 \leq \theta_2 - \theta_1 \leq L\} = 1$ where L is a fixed constant. Then show that

   $\inf_{\mu,\sigma^2} P_{\mu,\sigma^2}\{\theta_1 \leq \mu \leq \theta_2\} = 0.$ [15]

INDIAN STATISTICAL INSTITUTE

B.STAT.(HONS) III YEAR BACKPAPER EXAMINATION 1994-95

SEMESTRAL - I EXAMINATION

STATISTICAL INFERENCE II

Date: 26.6.95 Maximum Marks: 100 Time: 3 Hours

Note: You may use any result done in class either proved or unproved. Be sure to state them clearly. The paper carries 120 marks and the maximum you can score is 100.

1. Suppose $X_1, X_2, \ldots, X_n$ are iid observations from $N(\mu, \sigma^2)$. The parameter of interest is the density ϕ_{μ,σ^2}.

   (i) If $n = 1$ show that there is no unbiased estimator.

   (ii) If $n \geq 2$, show that $X_n - \bar{X}_n$ and $\bar{X}_n$ are independent.

   (iii) If $n \geq 2$, using (ii) or otherwise find an estimator $f(\cdot)$ such that $Ef(x) = \phi_{\mu,\sigma^2}(\cdot)$ for all x. [10+5+10=25]

2. Suppose $X_1, \ldots, X_n$ are i.i.d standard Cauchy random variables. Let $X_{(k,n)}$ be the k^{th} order statistic. Decide whether $E(X^2_{(k,n)}) < \infty$ for any k. What is the limiting distribution of the sample median as $n \to \infty$? State clearly any result that you are using. [20]

3. Consider the kernel density estimate where the kernel is of order p. Under suitable assumption derive the $AMISE$ and the optimal bandwidth. [10]

4. Suppose $Y_i = X_i\beta + e_i$ $i = 1, 2, \ldots n$ and we estimate β by minimising $\sum_{i=1}^{n} |Y_i - X_i\beta|$. Show that this can be written as a linear programming problem. [10]

5. Suppose $X_1, \ldots, X_n$, are iid observations from the distribution $F(x) = \frac{1}{2}\Phi(x) + \frac{1}{2}\Phi(\frac{x}{\gamma})$ where Φ is the d.f. of $N(0,1)$. Consider the two estimators $\bar{X}_n =$ sample mean and $\tilde{X}_n =$ sample median and compare them in terms of relative efficiency. State clearly any result that you are using. [20]

6. Derive the Sequential Cramer Rao inequality. State clearly any assumptions that you are making and any result that your are using. [20]

7. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$ where μ, σ^2 are both unknown. Suppose $\theta_1 = \theta_1(X_1, \ldots, X_n)$ and $\theta_2 = \theta_2(X_1, \ldots, X_n)$ are such that

   $\inf_{\mu,\sigma^2} P_{\mu,\sigma^2}\{0 \leq \theta_2 - \theta_1 \leq L\} = 1$ where L is a fixed constant. Then show that

   $$\inf_{\mu,\sigma^2} P_{\mu,\sigma^2}\{\theta_1 \leq \mu \leq \theta_2\} = 0. \quad [15]$$

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1994-95
BACK-PAPER SEMESTRAL-II EXAMINATION
Introduction to Stochastic Processes

Date: 23.6.95 Maximum Marks: 100 Time: 3 hours

Group - A

Attempt any two of the following problems

1. Consider a Poisson process $\{X(t): t\geq 0\}$ with intensity $\lambda > 0$.

(a) Let S_r be the time to the occurrence of the r-th event. Obtain $E(S_r | X(t) = n)$ for $r \leq n$ and $r > n$.

(b) Let T be the time required to observe the first event and let $N(\frac{T}{K})$ be the number of events in the next $\frac{T}{K}$ units of time. Obtain the first two moments of $N(\frac{T}{K})\, T$.

(c) Suppose each occurrence of the event is registered with probability π, independent of other occurrences. Let $\{Y(t): t\geq 0\}$ be the process of the registered occurrences. Prove that the latter is also a Poisson process with intensity parameter $\lambda\pi$.

(9+8+8=25)

2. (a) What is an age replacement policy? Develop the relevant optimization equation and discuss its solvability.

(b) Cars arrive at a gate. Each car is of random length L having distribution function $F(.)$. The first car arrives and parks against the gate. Each succeeding car parks behind the previous one at a distance that is random and is distributed according to a uniform distribution on [0, 1]. Consider the number of cars that are lined up within a total distance x of the gate. Determine $\lim_{x\to\infty} \frac{E N(x)}{x}$ when F is (i) degenerate at c and (ii) exponential with mean 1.

(c) State and prove a result concerning the asymptotic distributional property of a renewal process.

(10+8+7=25)

3. (a) Suppose that $\{N_i(t): t\geq 0\}, i=1,2$ are two independent Poisson processes with rates $\lambda_i, i=1,2$ and

$N(t)=N_1(t)+N_2(t)$ is their sum. Suppose that T_1 and Z are the times to the first occurrences under $N_1(t)$ and $N(t)$ respectively. Obtain the probability distributions of (i) $N_1(T_2)$ and (ii) Z.

(b) In a parking lot with N spaces, the incoming traffic is of Poisson type with intensity λ, but only as long as empty spaces are available. The occupancy times are iid, the common distribution being exponential. Find the appropriate differential equations for the probabilities $P_n(t)$ of finding exactly n spaces occupied.

(c) Consider a renewal process $\{N(t), t \geq 0\}$ and define $\gamma_t = S_{N(t)+1} - t$ the residual life of the unit in use at time t, where S_n is the time to the n-th renewal, $n = 1, 2, \ldots$. Using renewal theoretic arguments, show that for $x > 0$,

$$P(\gamma_t > x) = 1 - F(t+x) + \int_0^t [1 - F(t+x-y)]\, dH(y)$$

where $H(t) = E\,N(t)$ and F is the distribution function of S_1. (8+7+10=25)

GROUP - B

Max. Marks: 50

1.(a) Define a stochastic process and a Markov Chain.

(b) Define a branching chain and its probability of extinction.

(c) Give example of a Markov Chain whose all elements are transient. Can it be finite ?

(3+3+4) = [10]

2.(a) Does there exist an irreducible Markov Chain with uncountably many states ? (justify).

(b) Can one irreducible Markov Chain have one state positive recurrent and other states null recurrent ? (Justify).

(c) Let a finite - state Markov Chain (irreducible) with transition matrix P have the property that $P^2 = P$. Prove that each row of P is identical.

(12+9+6) = [27]

3. Let the state space $\mathcal{S}$ of an (irreducible) Markov Chain be countably infinite. Let P(x,y) denote the transition probability of it and it is given that

$$\sum_x P(x,y) = \sum_y P(x,y) = 1$$

Prove that the chain can not have a stationary distribution.

[13]

:bcc:

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1994-95
BACK-PAPER SEMESTRAL-II EXAMINATION
Design of Experiments

Date: 21.6.1995 Maximum Marks: 100 Time: 3 hours

Note: Answer ALL questions. Keep your answers brief and to the point.

1. A chemist wishes to test the effect of 5 chemicals on the strength of yarn using 4 kinds of yarn.

   (a) Suggest a suitable experimental design for this experiment if it is known that the kinds could represent a potential source of variability. Justify your choice of design.

   (b) Give the model to be used and the ANOVA table showing the sources of variation, d.f., expressions for sum of squares and E (mean squares).

   (c) Suppose the observation for chemical 1 was missing for kind-of-yarn 3. Derive an expression for estimating this missing value.

(4+8+3) = [20]

2. An engineer wants to compare 5 different production processes for plastic sheets. He suspects that the raw material supplied by different suppliers are not of equal quality. He also knows that the machines used do not have uniform performances. 6 manufacturers, each of whom have 5 different machines, are each to run a complete replicate of the experiment. Each manufacturer has his own 5 suppliers of raw material.

   (a) Suggest, with reasons, a suitable design for this experiment.

   (b) For the suggested design give the model, and the ANOVA table showing the splitting of the total of and the different SS.

(4+10) = [14]

3. Consider the effect of different factors on the life of gears. The factors of interest are: (A) weight of load: light and heavy; (B) type of load: shock and steady and (C) speef of operation: low and high.

   (a) If an experimenter can run only 4 treatment combinations under similar conditions at one time, design a suitable experiment, including a suggested number of replications.

contd..... 2/-

Assume that interest lies mostly in main effects but some information is also required on the interactions.

(b) For the suggested design in (a), give the model and the ANOVA table showing d.f. and expressions for SS.

(c) Give the information on each effect as obtained from the design in (a).

(8+8+8) = [24]

4. Consider a 3^3 experiment in a CRD replicated r times. The 3 factors are A, B, C.

Suppose each factor has 3 equispaced quantitative levels.

(a) If we want to explore the possible curvature in the main effects, explain how you would do so.

(b) Obtain expressions for the, SS, Mean Square and E (Mean Square) for A_L, the linear effect of A.

(c) Show the model and the ANOVA table giving d.f. only showing the linear and quadratic components of the effects. (give the necessary assumptions used if any).

(5+9+6) = [20]

5. (a) What is the role of the concomitant variables in the analysis of covariance.

(b) Consider the ANOVA model:

$$E(Y) = X\beta + C\gamma$$

where $Y_{n\times 1}$ is the response vector, $X_{n\times p}$ is the design matrix, $\beta_{p\times 1}$ is the parameter vector, $C_{n\times k} = (C_1 --- C_k)$ where C_i represents observations on the ith concomitant variable and γ is the vector of regression coefficients.

(i) What is the adjustment to be made in estimating β under this model as compared to the corresponding ANOVA model ? Give the necessary derivations.

(ii) With necessary derivations, show how you would test the hypotheses $H_0 : \gamma = 0$ against $H_1 : \gamma \neq 0$. Explain the significance of H_0 in this context.

(iii) Show how you would test the hypothesis about some contrasts of the parameters of interest, say $H_0 : H\beta = m$.

(5+6+6+5) = [22]

:bcc:

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1994-95
BACK-PAPER SEMESTRAL-II EXAMINATION

Economics IV

Date: 19.6.1995 | Maximum Marks: 100 | Time: 3 hours

Note: Answer question no.1 and any FOUR from the rest of the questions.

1. Consider the model

$$y_1 = \beta_1 + \beta_2 x_t + e_t, \quad t = 1, 2, 3, 4, 5, 6, 7 \text{ and } 8,$$

where $x_t = 2t-9$, $E(e_t) = 0$, $E(e_t^2) = \sigma^2 |x_t|$ and

$E(e_t e_s) = 0$ for $t \neq s$. Let $e' = (e_1, e_2, \ldots, e_8)$, and

$E(ee') = \sigma^2 \Psi$.

(a) Specify Ψ.

(b) Find Ψ^{-1}.

(c) Find the covariance matrix of the LS estimate for $\beta = (\beta_1, \beta_2)$

(d) Find the covariance matrix of the GLS estimator for β.

(e) Find $E(\hat{\sigma}^2)$, where $\hat{\sigma}^2 = (y-Xb)'(y-Xb)/6$, where b is the LS estimator.

(2+2+4+7+5) = [20]

2. (a) Define Multiplicative Heteroscedasticity.

(b) Give details of EGLS estimation when the variances are unknown and the heteroscedasticity is of the multiplicative form.

(c) Describe a test for Multiplicative Heteroscedasticity.

(2+13+5) = [20]

3. Assume that the short-run production decision of a firm is given by the model

$$y_t = b_0 + b_1 x_t + u_t$$

where y_t = output, x_t = labour input.

Suppose further that whenever anything causes the firm to "overproduce" in the period t-1 (a fact identified by $u_{t-1} > 0$) the firm will tend to "underproduce" in period t (a fact indicated by $u_t < 0$).

contd...... 2/-

(a) Identify which assumption(s) of the linear regression model is (are) violated.

(b) Indicate the effects of these violations on the OLS estimate of the slope coefficient and its standard error.

(c) Discuss briefly the appropriate "corrective solution" in this case.

(2+10+8) = [20]

4.(a) Give an example of a system of simultaneous equations of which one equation is exactly identified, one is under-identified and the other is over-identified.

(b) Construct an example of an econometric model in which the order condition for identifiability is satisfied but the rank condition is not.

(c) Is the following statement true ? Give reasons for your answer:

"If the residuals in a regression model are not independently distributed with a common variance σ^2, the OLS estimates are always less efficient than the GLS estimates for all finite sample sizes."

(9+3+8) = [20]

5.(a) What do you understand by the identification problem in the context of simultaneous equation model ? Illustrate your answer with examples of your own.

(b) State the formal identifiability rules, with reference to your examples.

(c) What are the implications of the state of identification of a model for choice of estimation method ?

(d) Suggest ways for rendering any unidentified model exactly identified.

(10+5+3+2) =[20]

6. In the following simultaneous equation model

$$y_1 = \beta_1 y_2 + e_1$$
$$y_2 = \beta_2 y_1 + \lambda x + e_2$$

where all variables are measured from their respective means:

contd..... 3/-

(a) Find a consistent estimator for β_1;

(b) Prove that the OLS estimate of β_1 is consistent if any one of the following three conditions hold:

(i) $\beta_1\beta_2 = 1$, (ii) $V(x) = \infty$ and (iii) $V(e_1) = 0$ and $cov(e_1, e_2) = 0$.

(10+10) = [20]

7. Consider the three-variable model, expressed in deviations from the means

$$y_t = bx_t + cz_t + u_t.$$

Prove that the estimate $\hat{b}$ from applying OLS to the above model is equal to the estimate which we can obtain by regressing

y on x where

y = residuals from the regression $y = \alpha z + v$

x = residuals from the regression $x = \beta z + w$.

(where v and w are random terms satisfying the usual assumptions.)

[20]

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year : 1994-95
SEMESTRAL-II EXAMINATION
Economics IV

Date: 4.5.1995 Maximum Marks: 100 Time: 3 hours

Note: Answer question no.1 and any FOUR from the rest of the questions.

1. Consider the following model, expressed in deviations of the variables from their means:

$$y_1 = b_1 y_2 + u$$
$$y_2 = c_1 y_1 + c_2 x_1 + v.$$

Use the following observations (which are given in deviation form):

y_1	y_2	x_1
-4	2	-2
0	3	-1
3	2	0
1	-7	3

to estimate the first equation (a) by OLS, (b) by ILS, (c) by 2SLS and compare your results. Which method is most appropriate in this case and why ? [20]

2. Consider a simple linear regression model

$$y_t = \alpha + \beta x_t + e_t, \qquad t = 1, 2, \ldots, n;$$

with usual (classical) assumptions. Suppose that n sample observations are divided into G groups. The group means are $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_G$ and $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_G$. The corresponding number of observations are $n_1, n_2, \ldots, n_G$ such that $\sum_{g=1}^{G} n_g = n$.

If we denote x_{ig} and y_{ig} for the ith observations in the gth group, we have

$$y_{ig} = \alpha + \beta x_{ig} + e_{ig}, \qquad i = 1, 2, \ldots, n_g, \quad g = 1, 2, \ldots, G.$$

contd..... 2/-

Suppose now that instead of being given a complete enumeration of all observations in each group, we are only given their number and the mean values of x and y. Derive estimation formulas for the regression coefficient β using the group means, and compare the variance of the estimator with that of OLS estimator based on individual observations. When are the two variances equal ?

[20]

3. Define best linear unbiased predictor in the Generalized Linear Statistical Model. Simplify it when the disturbance term follows a first order autoregressive process. Derive its covariance matrix.

(3+12+5) = [20]

4. Describe DW test for autocorrelation giving important steps to derive the probability distribution of DW statistic under the assumption that $\rho = 0$.

[20]

5. Suppose

$$q = \alpha p + u \quad \ldots \quad (i)$$

and

$$q = \beta p + v \quad \ldots \quad (ii)$$

are two relations operating simultaneously, where q and p are observables, α and β are unknown constants, and u and v are nonobservable random variables with zero means, constant (unknown) variances (designated σ_u^2 and σ_v^2) and zero covariance.

(a) Show that, in the population of observations (q,p) generated by this system, the LS regression coefficient of q on p is equal to a weighted average of α and β, the weights being σ_v^2 and σ_u^2.

(b) If in addition it is known that $\sigma_v^2 = K\sigma_u^2$, where K is a known nonnegative constant, show how α and β might be estimated.

(c) Suppose equation (ii) is replaced by

$$p = \beta z + v \quad \ldots \quad (iii)$$

where z is an exogenous variable. Under the same assumption about the distribution of u and v as before (i.e., as in (a)) what can you say about the LS regression coefficient of q on p ?

(8+5+7) = [20]

p.t.o.

6. Show that

(a) One way of interpreting the LS estimators is in terms of the solutions to the problem of minimizing the variance of the estimator, given the constraint that the estimator is linear and unbiased.

(b) In the equicorrelated case, in which the covariance matrix (Ω) exhibits homoscedasticity (equal diagonal elements) and equal covariances (equal off-diagonal elements, which are not necessarily zero), the GLS estimator reduces to the OLS estimator.

(10+10) = [20]

7. Examine whether the following statements are true, false, or uncertain. Give a short explanation. If a statement is not true in general but is true under some conditions, state the conditions:

(a) Working with a model expressed in the first differences of the variables

$$(y_t - y_{t-1}) = b_0 + b_1(x_t - x_{t-1}) + u_t$$

is equivalent to a model in which the variables assume their current values and a trend appears as a separate regressor

$$y_t = a_0 + a_1 x_t + a_2 t + v_t .$$

(b) In the 2SLS method we should replace only the endogenous variables on the right hand side of the equation by their estimated values from the reduced form. We should not replace the endogenous variable on the left hand side by its estimated value from the reduced form.

[20]

cc:

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1994-95
SEMESTRAL-II EXAMINATION
Biology II

Date: 3.5.1995 Maximum Marks: 100 Time: 3 hours

Note: Answer any EIGHT questions.

1. Write short notes on (any three):
   (a) Structural modification of epidermis for protection mechanisms from disease of plants;
   (b) Phytoalexins;
   (c) Prosthetic group;
   (d) Michelis-Menten constant (Km).

2. Write brief description of factors affecting enzyme action.

3. Name the factors influencing crop production. How general classification of climatic zones is made on the basis of rainfall and evaporation.

4. What are plant nutrients ? Name essential plant nutrients with their sources. What are the criteria for essentiality of element in plant nutrition ?

5. Write short notes on (any three):
   (a) 2 : 1 type expanding mineral;
   (b) Soil acidity
   (c) CEC
   (d) Base saturation.

6. Name the different cultural operations in direct and transplanted rainfed rice. Give the water requirement at various growth stages of rice in the field.

7. What are the different yield attributing characters of rice ? Estimate the yield per hectare of rice from the following data:
   (i) Spacing - 20 x 20 cm.
   (ii) Average no. of effective tillers/hill - 9
   (iii) Average no. of grains/panicle - 160

contd..... 2/-

(iv) Average no. of unfilled grains/panicle - 20

(v) Test weight (1000 grain) - 22 gms.

8. Write short notes on (any four):

(a) Soil texture

(b) Soil water

(c) Critical temperature for rice at different growth stages.

(d) Bulky and concentrate organic manure

(e) Fertility status of soil.

(f) Nutrient uptake by plant.

9. Write short notes on (any three):

(a) Nucleosome,

(b) Molecular transmission of genetic information,

(c) Plasmid,

(d) Gene library,

(e) Transgenic animal.

10. What is recombinant DNA ? Illustrate the model for synthesis and introduction of recombinant DNA into an animal system.

:bcc:

INDIAN STATISTICAL INSTITUTE
B.STAT. (HONS.) III YEAR: 1994-95
SEMESTRAL II EXAMINATION
INTRODUCTION TO STOCHASTIC PROCESSES

Date: 2.5.95 Maximum Marks: 50+50 Time: 3 hours

GROUP A

Note: Attempt any two of the following problems

1. (a) Suppose $X_1, X_2, \ldots$ are the inter-occurrence times of a renewal process $\{N(t): t \geq 0\}$, with $EX_1 = \mu$. Discuss the limiting behaviour of $\frac{E(N(t))}{t}$ as $t \to \infty$.

(b) Let $\{N(t): t \geq 0\}$ and $\{M(t): t \geq 0\}$ be two independent Poisson processes with rates α and β respectively. What is the probability that the number of occurrences of the event under $M(t)$ within the interval between the first and the $(r+1)$-st occurrences of the event under $N(t)$ equals n.

(c) In (b) above, obtain the covariance between $N(t)$ and $N(t+s)$. (13+7+5=25)

2. (a) State and prove a result regarding existence of various moments of $N(t)$, where $\{N(t): t \geq 0\}$ is a renewal process.

(b) Suppose that passengers arrive at a metro station in accordance with a Poisson process with rate λ to avail of the first train which departs at time t. Compute the expected value of the sum of the waiting times of the passengers arriving in $(0, t)$.

(c) Suppose $\{N(t): t \geq 0\}$ is a Poisson process. Define another process $\{X(t): t \geq 0\}$ by $X(t) = N(t+s_0) - N(s_0)$ where $s_0 > 0$ is a fixed constant. Is $\{X(t): t \geq 0\}$ also a Poisson process? Justify your answer. (9+9+7=25)

3. (a) Consider an equipment which is comprised of two units A and B. The units are connected in parallel; component-A is activated first and the component-B is automatically activated the moment component-A fails. Thus, component-B is the cold standby for component-A. The equipment fails once component-B fails and at that point of time, it is replaced by an identical equipment.

Let $X(t)$ = 1 or 2 according as whether component-A or component-B is functional at time $t \geq 0$; write $p(t) = P(X(t)=1)$. Assume that the life of component-A is distributed exponentially with mean μ_A and the same for component-B is distributed exponentially with mean μ_B.

p.t.o.

3. (a1) Show that under the boundary condition $p(0) = 1$, $p(t)$ satisfies the following differential equation:

$$\frac{d}{dt}(p(t))+\left(\frac{1}{\mu_A}+\frac{1}{\mu_B}\right) p(t) = \frac{1}{\mu_B}.$$

(a2) Now, solve this equation to obtain $p(t)$ explicitly.

(a3) Let $Y(t)$ denote the residual life of the equipment in position at time t. Argue that $E(Y(t)|X(t)=1)=\mu_A+\mu_B$. Hence show that

$$EY(t) = \mu_B + \frac{\mu_A}{\mu_B\left(\frac{1}{\mu_A}+\frac{1}{\mu_B}\right)} + \frac{1}{\left(\frac{1}{\mu_A}+\frac{1}{\mu_B}\right)} e^{-\left(\frac{1}{\mu_A}+\frac{1}{\mu_B}\right)t}.$$

(a4) Observe that $S_{N(t)+1}=t+Y(t)$ where S_n is the time to the n-th replacement. Use this identity to verify if

$$H(t) = \frac{1}{\mu_A+\mu_B} t - \frac{\mu_A\mu_B}{(\mu_A+\mu_B)^2}\left(1-e^{-\left(\frac{1}{\mu_A}+\frac{1}{\mu_B}\right)t}\right),$$

where $H(t)$ is the mean number of renewals of the equipment by time t.

(b) Consider a Poisson process $\{N(t):t\geq 0\}$ and let $Y_0, Y_1, Y_2, \ldots$ be integer-valued iid random variables that are independent of $\{N(t):t\geq 0\}$. Now, consider the process $\{X(t):t\geq 0\}$ where $X(t) = \sum_{n=0}^{N(t)} Y_n$. Show that for $s < t$, $X(s)$ and $X(t)-X(s)$ are independent. Hence, argue to conclude that $\{X(t):t\geq 0\}$ is a Markov process with sationary transition probability law $(P_{i,j}(s,t))$ to be specified by you. (4.5)*4+7=25)

contd. .....3.

GROUP B (Full marks:50)

Stochastic Processes

Note: Attempt all questions.

1. Consider a two-state Markov Chain with general transition probability matrix. Let the state space be = {0,1} and X_n be the state of the process at time n.
   If $P(X_0 = 0) = 1$, calculate $P(X_n = 0)$ and $P(X_n = 1)$ explicitly. [9]

2. Let x be a recurrent state and let x lead to y. Prove that y is also a recurrent state. [18]

3. The following represents a transition matrix.

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

   In the above transition matrix of a Markov Chain with state space {0, 1, 2, 3}, calculate the transient and recurrent classes. [4]

4. Let a Markov Chain with finite state space (with cardinality n) have a transition matrix $P_{n\times n}$ such that each row sum and each column sum of it equals 1. Find one stationary distribution of it. When is the stationary distribution unique? [7]

5. (a) Prove analytically that a finite - state Markov Chain always contains one positive recurrent state.

   (b) Let $P_{n\times n}$ be a stochastic matrix. Using (a) or, otherwise prove that $x = xP$ always have a solution x $(\neq 0)$ with non-negative elements. [6+6]

:SS:

INDIAN STATISTICAL INSTITUTE
B.STAT. (HONS.) III YEAR: 1994-95
SEMESTRAL II EXAMINATION
DESIGN OF EXPERIMENTS

Date: 28.4.95 Maximum Marks:100 Time: 3 Hours

Note: Answer all questions. Keep your answers brief and to the point.

1. A chemist has to perform an experiment to investigate whether 3 batches of raw material are of the same quality. On any day he can take 3 observations. He assumes that days could represent a potential source of variability and so decides to use a Randomized Block Design (RBD) with days as blocks. The experiment will be run over 4 days.

   (a) Describe the randomization procedure to be used by him.

   (b) Derive an expression for the relative efficiency of this RBD with respect to a comparable completely randomized design (CRD) where blocking due to days is not done.

   (c) The data from the RBD experiment gave the following sum of squares: SS(Days) = 3.9375, SS(Error) = 5.40. Calculate the relative efficiency in (b). Do you think the blocking due to days was worthwhile? (4+8+8 = 20)

2. An experiment is to be conducted over 5 days to compare the effects of 5 catalysts on the reaction time of a process. 5 operators will carry out the experiment and in each day 5 runs of the experiment can be made.

   (a) Suggest, with reasons, a suitable design for this experiment so that operator effects and day effects may be systematically controlled.

   (b) On day 2, operator 1 did not record the observed reaction time when he was working with catalyst 5. Derive a formula for estimating this missing value x so that x will have a minimum contribution to the error sum of squares.

   (c) After estimating x as in (b), the usual analysis of the data may be performed using this estimated value in place of the missing value. If now the usual F-statistic is found to be significant, what should be your conclusion? Justify your answer. (4+8+8=20)

P.t.o.

3. The following is the layout (in usual notation) of a design when a 2^3 experiment is performed in 2 replicates, each replicate having 2 blocks. The observed response, for each treatment combination is given below:

Replicate I		Replicate II	
(1) = 0	b = 2	(1) = -2	[illegible]
a = 1	c = 1	ac = 3	b = 0
abc = 6	ab = 2	bc = 3	c = [illegible]
bc = 4.	ac = 2	ab = 2	abc = 7

(a) What is the confounding scheme used?

(b) Write down the model for analyzing this data and give the splitting of the total d.f.

(c) Obtain an estimate of main effect A and interaction effect BC.

(d) If instead of using Replicate I and II as above, the experimenter had used only Replicate II repeated twice, compare the information given by the two designs on the different effects.

(4+6+8+6 = 24)

4. (a) Explain briefly the principle of partial confounding in factorial experiments.

(b) An experiment is to be performed over 3 days to study the effect of 3 factors A,B,C, each at 3 levels. Each day, only 9 observations can be taken. (i) Suggest a suitable confounding scheme for the experiment and briefly indicate how the blocks will be formed.

(ii) Give the partition of the total d.f. in the analysis of variance. (4+8+6=18)

5. The effect of 4 cooking temperatures and 3 composition proportions on the strength of an alloy are to be studied. The experimenter decides to run the experiment for 4 days. Also, preparation a batch of alloy with a certain proportion is time-consuming, so a splitplot design seems suitable for the experiment with days as blocks, proportions as whole plot and temperature as the split-plot treatments.

(a) Describe the randomization procedure to be used.

(b) Give the model for analysis and the ANOVA table including source of variation, d.f. and E(MS).

(c) How would you test for the hypothesis that the 4 temperatures have the same effect on the strength of alloy ? (4+10+4=18)

INSTITUTE
B.Stat. (Hons.) III Year : 1994-95
Semestral-I Backpaper Examination
Sample Surveys

Date : 30.12.1994 Max.Marks:100 Time:3 Hours.

Note : Answer any FIVE Questions. Marks allotted to each question are given within parentheses.

QUESTIONS

The symbols used are as usual.

1. A simple random sample of size $n = n_1 + n_2$ with mean $\bar{y}$ is drawn from a finite population and a simple random subsample of size n_1 is drawn from it with mean $\bar{y}_1$. Show that (a) $V(\bar{y}_1 - \bar{y}_2) = S^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right]$, where $\bar{y}_2$ is is the mean of the remaining n_2 units in the sample, (b) $V(\bar{y}_1 - \bar{y}) = S^2\left[\frac{1}{n_1} - \frac{1}{n}\right]$, (c) $Cov(\bar{y}, \bar{y}_1 - \bar{y}) = 0$. Repeated sampling implies repetition of the drawing of both the sample and the subsample.

(7+7+6) = [20].

2. (a) Let $\bar{r}$ denote the mean of the ratios y_i/x_i, $i=1,2,\dots,n$ for n units drawn from a finite population by SRSWOR. Obtain an expression for the bias in $\bar{r}$ in estimating the population ratio R and hence show that an unbiased estimator of the population total Y is given by

$$\hat{Y} = \bar{r}X + \frac{n(N-1)}{(n-1)}(\bar{y} - \bar{r}\bar{x}).$$

(b) Suggest a (Midzuno, Durbin Lahiri) sampling scheme under which the regression estimator of the population mean becomes unbiased. Obtain the variance of the regression estimator under the proposed sampling scheme. Also obtain an unbiased estimator of the variance.

$(\overline{3+7} + \overline{4+3+3}) = [20]$

(a) A sampler proposes to take a stratified random sample. He expects that his field costs will be of the form $\sum c_h n_h$. His advance estimates of relevant quantities for the two strata are as follows :-

Stratum	W_h	S_h	c_h (in \$)
1	0·4	10	\$4
2	0·6	20	\$9

(i) Find the values of n_1/n and n_2/n that minimize the total field cost for a given value of $V(\bar{y}_{st})$.

(ii) Find the sample size required, under this optimum allocation, to make $V(\bar{y}_{st}) = 1$. Ignore the f.p.c.

(iii) How much will the total field cost be?

(b) After the sample in (a) is taken, the sampler finds that his field costs were actually \$2 per unit in stratum 1 and \$12 in stratum 2.

(i) How much greater is the field cost than anticipated?

(ii) If he had taken the correct field costs in advance, could he have attained $V(\bar{y}_{st}) = 1$ for the original estimated field cost in (a) above? If no, what is the minimum field cost to reduce $V(\bar{y}_{st})$ to 1?

$(\overline{2+5+3} + \overline{3+7}) = [20]$

4. (a) Show that in case of a sample of size 2 drawn according to PPSWOR sampling scheme, the variance of the Des Raj ordered estimator $\hat{Y}_D$ can be expressed as

$$\gamma(\hat{Y}) = \left(1 - \frac{1}{2}\sum_{i=1}^{N} p_i^2\right)\left\{\frac{1}{2}\sum_{i=1}^{N} p_i\left(\frac{y_i}{p_i} - Y\right)^2\right\} - \frac{1}{4}\sum_{i=1}^{N} p_i^2\left(\frac{y_i}{p_i} - Y\right)^2.$$

If a sample of size n ~~is~~ drawn by PPSWOR sampling scheme, show that Murthy's unordered estimator is unbiased and has a smaller variance than that of the Des Raj ordered estimator. Obtain an expression for the variance of Murthy's unordered estimator. Also obtain a necessary form of ~~the~~ a uniformly non-negative quadratic unbiased estimator of the variance of the Murthy's unordered estimator.

$(10 + 2+3 +3 + 2) = [20]$.

(a) Show that for a fixed sample size (n) design

(i) $\sum_{i=1}^{N} \pi_i = n$, (ii) $\sum\sum_{i \neq j = 1}^{N} \pi_{ij} = n(n-1)$

where π_i and π_{ij}'s are the first and second order inclusion probabilities respectively.

(b) Derive the expression for the Yates - Grundy's form of the variance of the Horvitz-Thompson estimator. Under what sufficient condition is it uniformly non-negative?

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

6.(a) If $\bar{y}_{sy}$ denotes the mean of a linear systematic sample of size n drawn from a population of size N with sampling interval k ($N = nk$), show that

$$V(\bar{y}_{sy}) = \frac{\sigma^2}{n}\left[1 + (n-1)\rho_c\right]$$

where ρ_c is the intra-class correlation coefficient.

so as to ensure the unbiasedness of the sample mean $\bar{y}_{sy}'$ or suggest a modified estimator which itself is unbiased under linear systematic sampling?

(ii) Explain clearly why it is not generally possible to estimate $V(\bar{y}_{sy})$ unbiasedly based on a single systematic sample.

(iii) Give a special ~~case of circular systematic sampling~~ example where it is possible to estimate unbiasedly the variance of the estimated mean based on a single circular systematic sample.

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

7. In a study of overcrowding in a large city one stratum contained 100 blocks of which 10 were chosen with probabilities proportional to estimated size with replacement. An expected over-all sampling fraction $f_0 = \frac{1}{?}$ was [illegible]. Estimate the total number of persons and average persons per room and their standard errors from the following data:-

Blocks	1	2	3	4	5	6	7	8	9	10
Rooms	60	52	58	56	62	51	72	48	71	51
Persons	115	80	82	93	105	109	130	93	109	95

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

INDIAN STATISTICAL INSTITUTE
B. STAT. (HONS.) III YEAR: 1994-95
SEMESTER - I BACKPAPER EXAMINATION
DIFFERENTIAL EQUATIONS
QUESTION PAPER

Answer any FIVE questions

Maximum Marks: 100 Time: 3 Hours.

QUESTIONS

1. ... the general solution of the ODE's

$y' + y = 3x^3 - 1$

$y + (x + x^2 y) y' = 0$ 4+4

$y''' + 3y'' + 3y' + y = x^2 e^{-x} + 2$

$y'' + y = \cot x + x$ 6+6

2. ... the ODE's

$x^2 y'' - 7xy' + 15y = 0$

$x^2 y'' + 3xy' + 4y = \cos(\ln x) + 1$ 4+6

... solution of the ODE

$xy'' - (x+1)y' + y = 0$

$y_1(x) = e^x$. Find the general solution of

$xy'' - (x+1)y' + y = x + 1$ 10

... the general solution of the ODE

$y'' + 3x^2 y' - xy = 0$

... the neighbourhood of the point $x = 0$. 10

(b) Show that the ODE

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0$$

has a polynomial solution in the neighborhood of the origin if n is a positive integer.

4 (a) Define the regular singular point of a linear ODE of order n. Show that 0, 1, ∞ are the regular singular points of the ODE

$$x(1-x)y'' - \{(a+b+1)x - c\}y' - aby = 0$$

where a, b, c are constant parameters.

(b) Find the general solution of the ODE

$$2xy'' + y' + 2y = 0$$

in the neighborhood of the origin.

(a) Let $J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m (x/2)^{2m+n}}{m!\,(m+n)!}$.

Show that

(i) $x J_n'(x) = n J_n(x) - x J_{n+1}(x)$,

(ii) $2 J_n'(x) = J_{n-1}(x) - J_{n+1}(x)$.

(b) Compute the first four successive approximations (Picard iterates) for the solution of the ODE

$$y' = 1 + xy$$

satisfying $y(0) = 1$.

6 (a) When the integral $\int_a^b F(x, y, y')\,dx$ is stationary for fixed points, you may assume that y satisfies the Euler's equation

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0 \text{ for } a \le x \le b.$$

Show that

(i) if x does not occur explicitly in F, then

$$F(y, y') - y'\frac{\partial F}{\partial y'} = \text{a constant},$$

(ii) if y does not occur explicitly in F, then

$$\frac{\partial F}{\partial y'} = \text{const.}$$ 6+4

(b) Find the form of the curve passing through two points on a plane such that the length of the curve between these two points is a minimum. 10

(a) Solve the PDE

$$2y(u-3)u_x + (2x-u)u_y = (2x-3)y$$

given that $u = 0$ on the curve $x^2 + y^2 = 2x$. 10

(b) Reduce the PDE

$$x^2 u_{xx} - y^2 u_{yy} = 0$$

to the canonical form. Hence show that the general solution is

$$u = f(xy) + x\,g\left(\frac{y}{x}\right)$$

where f and g are arbitrary functions of their respective arguments. 10

INDIAN STATISTICAL INSTITUTE
B.STAT. (HONS.) III YEAR: 1994-95
ECONOMICS-III
SEMESTRAL-I EXAMINATION

Date:23.11.94 | Maximum Marks:60 | Time:2 hrs.30 min.

Note: Answer FOUR QUESTIONS, taking Two from each group. All questions carry equal marks.

GROUP-A

1. Consider a Harris-Todaro model of rural-urban migration with endogenous demand. Analyse the effects of the following exogenous changes on urban employment, urban unemployment and the relative price between industry and agriculture:

   a. A rise in the stock of industrial capital.
   b. A wage subsidy on industrial sector.
   c. A sudden opening up of international trade. (5x3=15)

2. What is the efficiency wage hypothesis? Show in terms of a suitable model, that an initially unequal distribution of assets might lead to involuntary unemployment. (15)

3. Consider a situation where a farmer enters into an interlinked credit and output contract with a moneylender cum trader. Show that the equilibrium is Pareto optimal and that the interest rate charged is never usurious. (15)

GROUP-B

4. Evaluate critically the development of the Indian economy since Independence. (15)

5. Examine critically the various hypotheses which have been put forward for the industrial stagnation in India which started around the mid-sixties. (15)

6. Suppose you are asked to compare the level of living across the different states of India. What are the different points which are to be taken into account in this kind of comparison? Discuss some studies in this area. (15)

INDIAN STATISTICAL INSTITUTE
B.STAT.(HONS.) III YEAR : 1994-95
GEOLOGY
SEMESTRAL-I EXAMINATION

Date: 25.11.94 Maximum Marks:100 Time: 3 Hours

Note: Attempt Question No.1 and any five from the rest.

1. Fill in the blanks (any ten). Only write down one of the four choices in each blank.

(a) The lower surface of a bed can be best determined from ________ (texture/cross-bedding/fossil/mineral composition).

(b) Sedimentary structures include ________ (grain-size/facies/sorting/stromatolites).

(c) The term 'texture' includes ________ (grain-size/fold/cross-bedding/composition).

(d) Slate is a ________ (terrigenous sedimentary/metamorphic/igneous/nonclastic) rock.

(e) A rock made up of 40% by volume of angular pabbles in called ________ (breccia/mudstone/conglomerate/very coarse grained sandstone).

(f) Marble is a ________ (bed/mineral/rock/clast).

(g) The dark colour of an igneous rock is due to the presence of ________ Fe,Mg/SiO_2/K,Cl/Ca,Na).

(h) Feldspar is an important ________ (precious/economic/rock-forming/colourful) mineral.

(i) A plutonic rock is made up of ________ (fine grained/both coarse- and fine-grained/coarse grained/clastic) minerals.

(j) The oldest fossil bird has been recorded from the rocks of ________ (Triassic/Jurassic/Cretaceous/Paleozoic) age.

(l) Presence of herbivorous, four-legged animal remains in a sedimentary rock indicates that the environment of deposition of the rock was most probably ________ (marine/mixed/non-marine/non-terrestrial.

(m) The ________ (muscles/eyes/teeth/blood cells) of a dinosaur are best preserved in rocks.

p.t.o.

2. Define a sedimentary rock. Describe how sediments are transported from one place to another. What is a non-clastic rock? Give an example each of a clastic and non-clastic rock. [4+6+4+2=16

3. What do you understand by the term 'bedding', 'bed' and 'lamination'? In what way are cross-bedding and graded bedding useful? [10+ 6=16]

4. Write brief notes (any two)
   (i) Stromatolites
   (ii) Competent and incompetent rocks
   (iii) Nonclastic texture [8x2=16]

5. Name the major discontinuties present in the Earth's interior. What is the tentative chemical composition of the core? Classify the igneous rocks on the basis of silica contents. What is the difference between the silicate structure noted in the feldpar group and the mica group of minerals? [4x4=16

6. State the law of superposition. In a road-cut section, a sandstone bed is found below a limerstone bed which is overlain by soil with vegetation. Indicate the position of a large-scale gap in deposition of those three units with justification. State how post-depositional deformation (folding, etc.) can change the basic attributes of the law of superposition. If a thin layer of igneous rock penetrates a sequence of sandstone at the base and shale at the top which rock unit among the three will be considered as youngest? [4x4=16]

7. What is a fossil? Explain the fact that fossils always indicate relative ages. State the law of faunal succession. How Darwin's theory of evolution has helped to conceive the law?

8. What is a rock? What do you understand by the term metamorphism of rocks? What kinds of physical changes are noted in a metamorphosed rock ? Write down the sequence of rocks that are formed when a mudstone is subjected to metamorphism over a long period of time. [3+3+3+7 = 16]

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1994-95
Physics I
Semestral-I Examination

Date : 23.11.1994 Maximum Marks : 100 Time : 3 Hours.

Answer Group A and Group B in separate booklets.

Given : velocity of light in vacuum $c = 3\times10^{8}$ m/s.
charge of an electron $e = 1.6\times10^{-19}$ coulomb
rest mass of an electron $m_0 = 9.1\times10^{-31}$ Kg.
$1\ eV = 1.6\times10^{-19}$ J
rest mass of a proton $= 1836\ m_0$
$1\,\text{Å} = 10^{-10}$ m.

GROUP A Max.Marks : 70

1. Assignments. [10]

2. A bead slides without friction on a friction less wire in the shape of a cycloid (see figure) with equations

$x = a(\theta - \sin\theta)$
$y = a(1 + \cos\theta)$

where $0 \le \theta \le 2\pi$. Find the Lagrangian function and the equation of motion. [10]

3.(a) Classify the following (giving reasons in brief) according as it is
(i) scleronomic or rheonomic
(ii) holonomic or non-holonomic
(iii) conservative or non-conservative.
"A sphere rolling down from the top of a fixed sphere". [2+2+2 = 6]

(b) A car of rest length 5m. passes through a garage of rest length 4m. Due to the relativistic Lorentz contraction, the car is only 3m. long in the garage's rest frame. There are doors on both ends of the garage, which open automatically when the front of the car reaches them, and close automatically when the car passes them. The opening or closing of each door requires a negligible amount of time.

p.t.o.

3.(b)(i) The velocity of the car in the garage's rest frame is

(a) 0.4c (b) 0.6c (c) 0.8c (d) greater than c

(e) not determinable from the data given.

(ii) The length of the garage in the car's rest frame is

(a) 2.4m (b) 4.0m (c) 5.0m (d) 8.3m

(e) not determinable from the data given.

(iii) Which of the following statements is the best response to the question: "Was the car ever inside a closed garage ?"

(a) No, because the car is longer than the garage in all reference frames.

(b) No, because the Lorentz contraction is not a real effect.

(c) Yes, because the car is shorter than the garage in all reference frames.

(d) Yes, because the answer to the question in garage's rest frame must apply in all reference frames.

(e) There is no unique answer to the question, as the order of door openings and closings depends on the reference frames.

[3+3+2=8]

4. Find the Hamiltonian for the Lagrangian

$$L = \tfrac{1}{2} e^{-kt}(\dot{x}^2 - \mu^2 x^2).$$

Do you expect the Hamiltonian to be a conserved quantity in the case ?

[6]

5.(a) Calculate the Poisson Bracket of a cartesian momentum component with a component of the angular momentum vector.

(b) Give the wavelength shifts in the relativistic longitudinal Doppler effect for the sodium D_1 line (5896 Å) for source and observer approaching at relative velocities of 0.8c. Is the classical result a good approximation ?

[7+8=15]

6.(a) Compute the effective mass of a photon of wavelength 5000 Å.

(b) Find the energy equivalent to the rest mass of the proton.

(c) A charged π meson (rest mass = 273 m_0) at rest decays into a neutrino (zero rest mass) and a μ-meson (rest mass =207 m_0). Find the Kenetic energies of the neutrino and the μ-meson.

Contd.....

6.(c) OR

Ionization measurements show that a particular charged particle is moving with a speed given by $\beta = u/c = 0.71$, and that its radius of curvature is 0.46m. in a field of magnetic induction of 1.0 tesla. Find the mass of the particle and identify it.

[3+3+9=15]

GROUP B Max.Marks :30

Answer any two questions from the following:

[2x 15=30]

1. State the second law of thermodynamics and explain what is leads to.

   Compare the significance of the first law and the second law of thermodynamics.

   A Carnot engine converts $\frac{1}{6}$ of the heat supplied into work. If the temperature of the sink be reduced by $62^{o}C$, the efficiency is doubled. Find out the temperatures of the source and the sink.

2. Give the arguments leading to the construction of the absolute scale of temperature and fully develop the idea of this c . scale.

   A sample of a gas initially at $27^{o}C$ is compressed from 40 litres to 4 litres adiabatically and reversibly. Calculate the final temperature [= 1.4].

3. Define 'entropy' in thermodynamics. Discuss about its two main properties and state their implications.
   Using any one of the Maxwell's thermodynamic equations, if necessary, show that for one gram of a substance

$$C_p - C_v = T\left(\frac{dp}{dT}\right)_v \left(\frac{dv}{dT}\right)_p$$

   where the letters have their usual significances.

4. Write short notes on :
   (a) Conditions of reversibility
   (b) Perpetual motion of the second kind
   (c) Indicator Diagrams.

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year : 1994-95

Differential Equations

Semestral-I Examination

Date : 23.11.1994 Maximum Marks : 100 Time : 3 Hours.

Answer any FIVE questions.

1.(a) Obtain the general solutions of the ODE's

(i) $y' + e^x y = 3e^x$

(ii) $x^2y^2 - xy + x(1+x^2y)y' = 0.$ [4+4]

(b) Obtain the general solutions of the ODE's

(i) $y'' + 9y = x^2 e^{3x} + 2 \sin 3x$

(ii) $y'' + y = \sec^3 x + 1.$

[6+6]

2.(a) Solve the ODE's

(i) $x^2y'' - 7xy' + 15y = 0$

(ii) $x^4y^{IV} + 5x^3y''' + x^2y'' + 2xy' - 2y = x^2.$

[4+6]

(b) One solution of the ODE

$$x^2y'' - xy' + y = 0$$

is $y_1(x) = x$. Find the solution of the ODE

$$x^2y'' - xy' + y = x^2$$

satisfying the conditions $y(1) = 1$, $y'(1)=0$.

[10]

3.(a) Find the general solution of the ODE

$$y'' + (x-1)^2y' - (x-1)y = 0$$

in the neighbourhood of the point $x=1$.

[10]

(b) Show that $P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \{(x^2-1)^n\}$ where n is a positive integer, is a solution of the ODE

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0.$$

Also show that

$$\int_{-1}^{1} P_n(x)\, P_m(x)dx = \begin{cases} 0 & \text{for } n \neq m \\ \frac{2}{2n+1} & \text{for } n=m. \end{cases}$$

[4+6]

4.(a) Find the general solution of the ODE

$$4xy'' + y' + y = 0$$

in the neighbourhood of the origin. [10]

(b) Obtain two linearly independent solutions of the ODE

$$x^2y'' + xy' + x^2y = 0$$

which are valid near $x=0$. [10]

p.t.o.

5.(a) Let $J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m (x/2)^{2m+n}}{m!\,(m+n)!}$.

Show that

(i) $\frac{2n}{x} J_n(x) = J_{n-1}(x) + J_{n+1}(x)$

(ii) $\frac{d}{dx}(x^n J_n(x)) = x^n J_{n-1}(x)$. [5+5]

(b) What do you mean by the statement that " $f(x,y)$ satisfies a Lipschitz condition w.r.t. y in a region D" ? Show that the function f given by $f(x,y) = y^{1/2}$ does not satisfy a Lipschitz condition on S_1: $|x| \leq 1$, $0 \leq y \leq 1$, but satisfies a Lipschitz condition on S_2: $|x| \leq 1$, $\frac{1}{2} \leq y \leq 1$.

State the importance of Lipschitz condition in the existence and uniqueness of solution of the ODE

$\frac{dy}{dx} = f(x,y)$ satisfying $y(x_0)=y_0$ in a region D.

[2+2+2+4]

6.(a) If the integral $\int_a^b F(x,y,y')dx$ (whose end points are fixed) is stationary, then deduce the Euler's equation

$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0$ for $a \leq x \leq b$.

Show that this is a second order ODE in y. [8+2]

(b) A particle slides under gravity from rest along a smooth vertical curve joining two points A and B. Find the form of the curve if the time from A to B is a minimum.

[10]

7.(a) Answer either (i) or (ii)

(i) Solve the PDE

$(x-y)y^2 u_x + (y-x)x^2 u_y = (x^2+y^2)u$

given that $u = \frac{a^3}{x}$ on $y = 0$.

(ii) Define a 'weak' solution of the PDE

$u_x + u_y = 0$.

Hence verify that $u = H(x-y)$ where $H(x) = \begin{cases} 1 & \text{for } x>0 \\ 0 & \text{for } x<0 \end{cases}$

is a 'weak' solution of it. [10]

(b) Reduce the PDE

$(n-1)^2 u_{xx} - y^{2n} u_{yy} = n y^{2n-1} u_y$

to canonical form, and find its general solution.

[10]

INDIAN STATISTICAL INSTITUTE

B.Stat. (Hons.) III Year : 1994 - 95

SEMESTRAL - I EXAMINATION

Sample Surveys

Date: 21.11.1994 Maximum Marks: 100 Time: 3 hours

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

N.B.: Notations and symbols used are as usual.

1. A simple random sample of size 3 is drawn from a population of size N with replacement.

(a) Show that the probabilities that the sample contains one, two and three distinct units are respectively

$$P_1 = \frac{1}{N^2}, \quad P_2 = \frac{3(N-1)}{N^2} \quad \text{and} \quad P_3 = \frac{(N-1)(N-2)}{N^2}.$$

(b) Show that the average variance of $\bar{y}'$ is

$$V(\bar{y}') = \frac{(2N-1)(N-1)S^2}{6N^2}$$

where $\bar{y}'$ is the sample mean over the distinct units only.

(c) Hence show that $V(\bar{y}') \leq V(\bar{y})$, where $\bar{y}$ is the ordinary mean of all the observations in the sample.

(9+6+5) = [20]

2.(a) If there are two strata and if ϕ is the ratio of the actual n_1/n_2 to the Neyman optimum n_1/n_2, show that whatever may be the values of N_1, N_2, S_1 and S_2, the ratio $V_{min}(\bar{y}_{st})/V(\bar{y}_{st})$ is never less than $4\phi(1+\phi)^2$ when the fpc's are negligible.

(b) If we denote by V_{ran}, V_{prop} and V_{opt} the variances of the estimated proportions in simple random sampling, stratified random sampling with proportional and optimum allocation, show that

$$V_{ran} = V_{prop} + \frac{(1-f)}{n} \Sigma W_h (P_h - P)^2$$

$$V_{prop} = V_{opt} + \frac{\Sigma W_h(\sqrt{P_h Q_h} - \overline{\sqrt{P_h Q_h}})^2}{n}$$

where $\overline{\sqrt{P_h Q_h}} = \Sigma W_h \sqrt{P_h Q_h}$.

(8+12) = [20]

p.t.o.

3.(a) Let $\hat{R}_-$ denote the mean of the n ratios $\hat{R}_i$ obtained by removing each unit in turn from the sample, so that $\hat{R}_i = \Sigma y/\Sigma x$ over the remaining (n-1) members. Show that an unbiased estimator of the population ratio R is given by

$$\hat{R}_H = \hat{R}_- + \frac{n(N-n+1)}{N\bar{X}}(\bar{y} - \hat{R}_- \bar{x}).$$

(b) Develop a sampling scheme under which the ratio estimator of the population total becomes unbiased. Obtain the variance of the ratio estimator under the proposed sampling scheme. Also obtain a necessary form of a uniformly non-negative quadratic unbiased estimator of the variance. Mention a sufficient condition under which it is uniformly non-negative.

(10+3+3+3+1) = [20]

4. Consider the following revised version of Roychoudhury's extended PPS sampling. Select first pair with probability Q_{ij}. Let the pair selected be U_{ab}. Then delete the ath row and bth column and from the remaining pairs select second pair with probability $Q_{ij}/(1-R_{ab})$ where R_{ab} is the sum of probabilities of pairs which are in ath row and bth column. Let the pair selected at second draw be U_{cd}. Then show that

(a) the following two estimators are unbiased for the population total Y:

$$t_1 = \frac{1}{2}\left[\frac{Z_{ab}}{Q_{ab}}(1 + R_{ab}) + \frac{Z_{cd}}{Q_{cd}}(1 - R_{ab})\right];$$

$$t_2 = \frac{\dfrac{Z_{ab}\,Q_{cd}}{1-R_{ab}} + \dfrac{Z_{cd}\,Q_{ab}}{1-R_{cd}} + \dfrac{Z_{ad}\,Q_{cb}}{1-R_{ad}} + \dfrac{Z_{cb}\,Q_{ad}}{1-R_{cb}}}{\dfrac{Q_{ab}\,Q_{cd}\,(2-R_{ab}-R_{cd})}{(1-R_{ab})(1-R_{cd})} + \dfrac{Q_{ad}\,Q_{cb}\,(2-R_{ad}-R_{cb})}{(1-R_{ad})(1-R_{cb})}}$$

(b) in case of perfect linear relationship between x and y with even a non-zero intercept, the above estimators have zero variances.

(12+8) = [20]

5.(a) If π_{ij} denotes the second order inclusion probability, show that $\theta(1-\theta) + \nu(\nu-1) \le \sum\sum_{i \ne j=1}^{N} \pi_{ij} \le N(\nu-1)$

where $\nu = E[\nu(s)] = [\nu] + \theta$.

contd..... 3/.

(b) What do you mean by a πps design ? How, under certain conditions, you can modify Lahiri - Midzuno - Sen sampling scheme so as to obtain a πps scheme ? Show that the Yates - Grundy's estimator of the variance of the Horvitz - Thompson estimator is uniformly non-negative under such a πps scheme.

(10+2+3+5) = [20]

6. Suppose a sample of n f.s.u.'s is selected from a population of N f.s.u.'s according to PPSWR sampling scheme and every time the ith f.s.u. is selected, a sample of m_i s.s.u.'s is selected from it by SRSWOR. Explain how on the basis of such a sample you would estimate unbiasedly the population total. Derive the variance of the proposed estimator and also an unbiased estimator of the variance. Indicate how you would estimate unbiasedly the between and the within components of the variance.

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

7. The following table gives the area under cultivation (y) for 7 plots :

Plot No.	1	2	3	4	5	6	7
y	2.8	2.6	2.3	3.5	2.4	3.8	4.1

(y in hectares; 1 hectare = 2.471 acres)

(a) Draw a circular systematic sample of 4 villages with 2 as the sampling interval.

(b) Estimate the total cultivated area Y in the region on the basis of the sample drawn.

(c) Estimate the RSE of the estimator in (b).

(d) Compare the estimate of RSE obtained in (c) with the RSE estimated on the assumption of drawing the sample by SRSWOR and comment on the result.

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

:bcc:

INDIAN STATISTICAL INSTITUTE
B.Stat. (Hons.) III Year: 1994-95
SEMESTRAL-I EXAMINATION
Linear Statistical Models

Date: 17.11.1994 Maximum Marks: 100 Time: 3 hours

Note: Answer any FOUR questions. Marks allotted to a question are indicated in brackets [ ] at the end.

1. Consider the following linear model for six observations with the usual assumptions including normality:

$$E(y_1) = E(y_3) = E(y_5) = \theta_1 + \theta_3 + \theta_4 ;$$
$$E(y_2) = E(y_4) = E(y_6) = \theta_2 - 2\theta_3 - \theta_4 .$$

(a) Obtain a necessary and sufficient condition for estimability of $\sum_{i=1}^{4} \lambda_i \theta_i$.

(b) Get one set of LSE for $\{\theta_i\}$ $i = 1,2,3$ and hence the B.L.U.E. of $\theta_1 + \theta_2 - \theta_3$.

(c) Obtain orthogonal error functions and hence give the expression for the sum of squares due to error in terms of y_i's.

(d) Give an appropriate test statistic for testing the hypothesis

$$H_0 : \theta_1 + \theta_2 = \theta_3 .$$

(6+7+5+7) = [25]

2.(a) Prove that in the linear model $(\underset{\sim}{y}, X\underset{\sim}{\beta}, \sigma^2 I)$, β_1 is estimable if and only if its corresponding column in X is linearly independent of the remaining columns of X.

(b) Let H denote a testable hypothesis consisting of some q linearly independent functions: $C\underset{\sim}{\beta} = \underset{\sim}{d}$ under the above linear model, say Ω with normality assumption. Define

$$\mathcal{S}_\Omega = \min_{\underset{\sim}{\beta}} (\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta}) \text{ and } \mathcal{S}_\omega = \min_{\underset{\sim}{\beta} | C\underset{\sim}{\beta} = d} (\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta})$$

$w = \Omega \cap H$. Show that

$$\mathcal{S}_\omega - \mathcal{S}_\Omega = (C\hat{\underset{\sim}{\beta}} - \underset{\sim}{d})' [\mathcal{D}(C\hat{\underset{\sim}{\beta}})/\sigma^2]^{-1} (C\hat{\underset{\sim}{\beta}} - \underset{\sim}{d})$$ where $\hat{\underset{\sim}{\beta}}$ is an LSE of $\underset{\sim}{\beta}$ and $\mathcal{D}(\underset{\sim}{x})$ stands for the dispersion matrix of the random variable $\underset{\sim}{x}$. Hence or otherwise show that the test

contd..... 2/-

statistic F to test H based on the confidence ellipsoid for $C\beta$ is the same as the one based on the likelihood ratio criterion.

(5+20) = [25]

3. Consider the following full rank linear model for k sets of data (x_{ij}, y_{ij}), $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n_i$:

$$y_{ij} = \alpha_i + \beta_i (x_{ij} - \bar{x}_i) + e_{ij}$$

with the usual assumptions including normality, where $\bar{x}_i = \sum_j x_{ij}/n_i$. Suppose we want to test the hypothes

$$H_1 : \begin{array}{l} \alpha_1 = \alpha_2 = \ldots = \alpha_k \\ \beta_1 = \beta_2 = \ldots = \beta_k \end{array}$$

and if H_1 is rejected we would like to test

$$H_2 : \beta_1 = \beta_2 = \ldots = \beta_k .$$

Obtain the least square estimators for the above parameters both under the full and the restricted models and hence derive the expressions for SSE, SSH_1 and SSH_2 and finally for the two F test statistics to test H_1 and H_2.

(13+12) = [25]

4. Derive Tukey's test for non-additivity with 1 d.f. stating clearly the statistic you will use and obtaining its distribution. Also comment on its uses and possible extensions.

(20+5) = [25]

5. Suppose in a $p \times q$ $(p < q)$ two-way classified data with one observation per cell all the p observations in cells (i,i), $i = 1, 2, \ldots, p$, got lost due to some accidents. Assuming an appropriate linear model without interactions, obtain the reduced normal equations for the row effects and hence give the expressions for (i) the BLUE's of estimable functions of row and column effects in terms of adjusted/unadjusted row/column totals, and (ii) the variances of these BLUE's. Give also the analysis of variance for testing the hypotheses relative to the two factor effects, indicating clearly how the various sums of squares are to be computed.

(10+7+8) = [25]

p.t.o.

6. When and why would you recommend the analysis of covariance ? Starting with an appropriate linear model for the analysis of covariance, show how you would estimate and test the covariance parameters with the help of the corresponding analysis of variance (ignoring the covariates). Illustrate this with respect to a two-way classification model with m > 1 observations per cell and with interaction terms, showing clearly how the various tests for the factor effects and interactions are carried out under the covariance model.

(15+10) = [25]

:bcc:

INDIAN STATISTICAL INSTITUTE
B.STAT.(HONS.) III YEAR: 1994-95
BIOLOGY-I
SEMESTRAL-I EXAMINATION

Date: 16.11.94 Maximum Marks: 100 Time: 3 Hours

Note: [You may skip the detailed biochemical pathways, just give the reactants' and products' name, wherever applicable]

Q.1 and Q.2 are compulsory. Attempt any three from the rest.

1. The genetic code table specifies the relationships between codons and amino acids. How and at which stage does a cell actually use the genetic code table to manufacture a polypeptide chain from the information contained in the nucleotide sequence of a gene? [Hint: Think of the various classes of tRNA molecules]. [15]

2. Write notes on any <u>three</u>
   (a) Plant cells can withstand wider fluctuation of osmotic pressure than the animal cells.
   (b) Photosystem I and II (schematic representatian only)
   (c) Meiosis
   (d) Hormones and neurotransmitter
   (e) Types of animal tissue and their function. [25]

3. What are the acceptable distinguishing features of a 'living' object? State the underlying assumptions (such as the condition of primitive earth etc.) of the 'origin of life' and the possible stages in chemical evolution. [20]

4. Write on the major (a) components and structure of prokaryotic and eukaryotic (animal) cells; and (b) differences between plant and animal cells. [20]

5. Give an overview on Photosynthesis. [20]

6. The reactions of Oxidative Phosphorylation (O.P.) are carried out by a ''system'' consisting of a number of enzymecomplexes (''X'') coupled with a number of electron carriers (''Y'')-- Write on (a) O.P. (define),; (b) ''System'' (name); (c) ''X'' (names and number); and (d) ''Y'' (names and number).

   Please also indicate schematically the reaction which involves ''X'' and ''Y''. [20]

7. Give an account on Mitosis and its different stages. [20]

INDIAN STATISTICAL INSTITUTE

B.Stat.(Hons.) III Year : 1994-95

Statistical Inference I

Semestral- I Examination

Date : 14.11.1994 Maximum Marks : 100 Time: $3\frac{1}{2}$ Hours.

The question paper carries 120 points. Maximum you can score is 100. Answer all questions.

1.(a) Suppose that $f(x,\theta)$ is a positive density on the real line which is continuous in x for each θ and such that if (X_1,X_2) is a sample of size 2 from $f(.,\theta)$, then X_1+X_2 is sufficient for θ. Show that $f(.,\theta)$ corresponds to a one parameter exponential family.

(b) Let X be an observation from $U(\theta,\theta+1)$. Show that $X-\frac{1}{2}$ is not the best unbiased estimator by exhibiting another unbiased estimator with smaller variance.

(c) Let $X_1,\ldots,X_n$ be i.i.d samples from $N(\theta,1)$, $\theta \in R$. Show that $\bar{X}_n = n^{-1}(X_1+..+X_n)$ is minimax, with respect to squared error loss.

[10+10+10=30]

2. Suppose k independent tests are performed on the same hypothesis $H_0 : \theta = \theta_0$, based on statistics $T_1,\ldots,T_k$. Let $\alpha(T_1),\ldots,\alpha(T_k)$ be the corresponding p-values, i.e. $P_{H_0}(T_i > T_{i,obs})$, $1 \leq i \leq k$.

Let $F = -2 \sum_{1}^{k} \log \alpha(T_i)$.

(a) Justify the use of F as a test statistic for H_0.

(b) What is the null distribution of F ?

[20]

3. Let $X_i = \beta z_i + \varepsilon_i$, $1 \leq i \leq n$ be a linear regression model where $\varepsilon_1,\ldots,\varepsilon_n$ are independent $N(0,\sigma_i^2)$, $1 \leq i \leq n$. Suppose β has a $N(0,\tau^2)$ prior distribution.

(a) Obtain the Bayes estimator of β.

(b) Obtain the limiting Bayes estimator of β by letting $\tau^2 \to \infty$.

p.t.o.

3.(c) Under what conditions is the limiting estimator consistent ?

(Assume throughout that $\{z_i\}$ is a sequence of positive constants).

[15+5+10 = 30]

4. Let $X_i = \theta + \varepsilon_i$, $1 \leq i \leq m$ and $Y_i = \mu + \beta z_i + \eta_i$ $1 \leq i \leq n$ be independent observations where z_i is a given sequence of constants. Obtain the likelihood ratio test for $H_0: \theta = \mu$ assuming $\varepsilon_1, \ldots, \varepsilon_m$, $\eta_1, \ldots, \eta_n$ are iid $N(0, \sigma^2)$, ($\sigma^2 > 0$, unknown). Derive its null distribution.

[10+5 = 15]

5. The following table gives the milk yields (in lb.) of cows calving in two different seasons, winter and summer. Under suitable assumption (to be clearly stated and justified) make appropriate statistical test to determine whether the cows calving in winter, on the average yielding more milk than in summer.

Dairy farm	Winter # Cows	Winter Average yield	Summer # Cows	Summer Average yield
1	2	484	2	316
2	7	350	1	262
3	2	414	2	270
4	2	286	4	223
5	5	176	3	180
6	7	386	4	320
7	3	385	3	398
8	4	438	5	340

<u>Note</u> : More 'realistic' modelling with 'approximate' test would be given more credit than 'careless' modelling with orthodox exact tests.

[25]

INDIAN STATISTICAL INSTITUTE
B.Stat.(Hons.) III Year : 1994-95
Statistical Inference II
Semestral-II Examination

Date : 24.4.1995 Maximum Marks : 100 Time : 3 Hours.

While answering, state clearly any results that you are using. You may use any result done in class, either proved or unproved. The paper carries 120 marks and the maximum you can score is 100.

1. Suppose $X_1, X_2, \ldots X_n$ are i.i.d observations from a distribution function F. The parameter of interest is $\theta^2 = [E_F(X_1)]^2$, F is completely unknown.

(i) If $n = 1$, show that there is no unbiased estimator of θ^2.

(ii) If n=1 and we restrict to all F's which have densities, does your answer to (i) change ?

(iii) If $n \geq 2$, find the UMVUE of θ^2.

[5+5+10 =20]

2. Suppose $X_1, X_2, \ldots, X_n$, $n \geq 1$ are i.i.d observations from the location family $F(x-\theta)$, where $F(x) = \frac{1}{1+e^{-x}}$, θ is the unknown parameter. Consider the two estimators $\bar{X}_n$ = sample mean, $\tilde{X}_n$ = sample median and compare them in terms of relative efficiency. State clearly any results that you are using, to do the computations.

[20]

3. Suppose $X_1, \ldots X_n$ have the distribution F and $Y_1 \ldots Y_m$ have distribution G and it is required to test H_0: F=G. Let W = number of pairs (i,j) such that $X_i < Y_j$, $i=1\ldots n$, $j=1,\ldots m$. Set up an appropriate test on the basis W. What is the asymptotic distribution of W as $m,n \to \infty$ under H_0 ? State clearly any result that you are using. For fixed m,n, what is the power of the test at the alternative $G(x) = F(x- \quad)$, $\quad > 0$? [20]

4. Suppose $X_1, X_2, \ldots X_n, \ldots$ are iid observations empirical distribution function. Define for $\varepsilon > 0$,

$$N = \inf \{ n \geq 1 : \sup_x |F_n(x) - F(x)| < \varepsilon \}.$$

Which of the following are true ? Give reasons. State clearly any results that you are using.

a) $P(N<\infty) = 1$

b) $E(N) < \infty$

c) $E(e^{tN}) < \infty$ for t in a neighbourhood of 0.

[20]

p.t.o.

5. Let $X_1 \ldots X_n$ be i.i.d observations from a distribution F. Suppose F has density f for $f^{(r)}$ and derive its AMISE under suitable assumptions. Find the optimal estimates in your class.

[10]

6. Derive the bias of the Kernel estimator of the regression function under suitable assumptions.

[10]

7. Suppose $X_1, X_2, \ldots$ are i.i.d. $N(\mu, \sigma^2)$. Let $\bar{X}_n$ and s_n^2 devote the sample mean and variance respectively, based on n observations. Let $N = \max \; n_0, \; \frac{t_{n_0-1} \; s_{n_0}^2}{d^2} + 1$ where [x] denotes the integer part of x and t_{n_0-1} is the upper 100 $\alpha/2$ print of a t-distribution with (n_0-1) degrees of freedom. Show that

a) $P(N<\infty)=1$

b) $P(\; \mu \in (\bar{X}_N \pm d)) \geq 1-\alpha$.

[5+15=20]