INDIAN STATISTICAL INSTITUTE
Mid-semester Examination : 2009-10
B. Stat. - Second Year
Analysis III

Date : 31. 08. 2009 Maximum Score : 35 Time :3 Hours

1. This paper carries questions worth a total of 44 marks. Answer as much as you can. The maximum you can score is 35 marks.

2. You are free to use any theorem proved in the class. However, you must state any theorem that you use at least once in the answer script.

3. We shall use $\mathbb{R}$ to denote the set of all real numbers.

(1) Let $Z \subset \mathbb{R}^2$ be the union of $X-$ and $Y-$axis and $f : Z \to Z$ a continuous map.

(a) If f is, moreover, one-to-one, show that f must have a fixed point.

(b) Is the above assertion true if f is not one-to-one? Justify your answer.

[6 + 6]

(2) Let $X \subset \mathbb{R}^n$ be compact and $f : X \to X$ a map such that

$$\forall x, y \in X[x \neq y \Rightarrow |f(x) - f(y)| < |x - y|].$$

Show that f is continuous and that it has a unique fixed point.

[6]

(3) (a) Let $f : \mathbb{R} \to \mathbb{R}^n$ be a differentiable map such that

$$f' \cdot f \equiv 0$$

on $\mathbb{R}$. Show that $|f(t)|$ is a constant.

(b) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function homogeneous of degree $p > 1$. Show that

$$\nabla f(x) \cdot x = pf(x)$$

for all $x \in$ ~~U~~. $\mathbb{R}^n$

[3 + 3]

(4) Let $U \subset \mathbb{R}^n$ be covex open, $f : U \to \mathbb{R}$ a C^m-function and $x, y \in U$. Define $g(t) = f((1-t)x + ty)$ on a suitable open interval containing $[0, 1]$. Show that g is m-times differentiable

P.T.O

and

$$g^m(t) = \sum \frac{\partial^m f}{\partial_{x_{i_m}} \cdots \partial_{x_{i_m}}}|_{(1-t)x+ty}(y_{i_m} - x_{i_m}) \cdots (y_{i_1} - x_{i_1}),$$

where summation is over all $(i_1, \cdots, i_m)$ of m-tuples of integers between 1 and n.

Using this state and prove Taylor's theorem for functions of several variables.

[8]

(5) Let R be a closed and bounded rectangle in $\mathbb{R}^n$. Suppose $f, g : R \to \mathbb{R}$ be bounded functions such that f is Riemann-integrable and $\{x \in R : f(x) \neq g(x)\}$ is of measure zero.

(a) Show by an example that g need not be integrable.

(b) If g is, moreover, integrable, show that

$$\int_R f = \int_R g.$$

[3 + 3]

(6) (a) Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be a function such that

$$f(x,y) = \frac{1}{\sqrt{xy}}, \; xy \neq 0.$$

Is f integrable on $[0,1] \times [0,1]$? Justify your answer. If f is integrable, compute its integral.

(b) Find the area bounded by the ellipse

$$x^2 + 2y^2 = 1.$$

[3 + 3]

Indian Statistical Institute

Mid-Semester – Examination : 2009 – 2010

B. Stat. II yr.

C & Data Structures

Date: 2nd September, 2009 Time: 3 hrs.

Note: *This paper carries 110 marks. Answer as much as you can. You will get maximum 100 marks.*

1. Given a real valued point function **f: R → R**, whose form is not known. Write a **C** function which takes the function **f** as a parameter and integrates it numerically between a given interval **[a, b]**, using *Trapezoidal* rule. 10

2. Write an efficient **C** program to compute the number of integer co-ordinate points within or on the circle $x^2 + y^2 = r$. 10

3. Write a **C** function that will reverse a linked list while traversing it only once. At the conclusion, each should point to the node that was previously its predecessor; the head should point to the node that was formerly at the end, and the node that was formerly first should have a **NULL** link.
   Is the header of a linked list usually a static variable or a dynamic variable? 10+2

4. What is hashing? Discuss any three hash functions. What do you mean by collision? Discuss open-addressing method for collision resolution. 2+6+1+3

5. Write an algorithm to convert a fully parenthesized algebraic expression, which may contain only +, −, ×, / and ^ operators and literals of unit length i.e., a, b, c, ... etc. into a reverse polish notation. 10

6. A *Scroll* is a data structure intermediate to a *Deque* and a *Queue*. In a *Scroll* all additions to the list are at its one end, but deletions can be made at either ends. Now, suppose that data items numbered 1, 2, 3, 4, 5, 6 come in the input stream in this order. By using a *Scroll*, which of the following rearrangements can be obtained in the output order?
   a) 1 2 3 4 5 6 c) 1 5 2 4 3 6 e) 1 2 4 6 5 3
   b) 2 4 3 6 5 1 d) 4 2 1 3 5 6 f) 5 6 2 3 1 4 6 × 2

7. Write the algorithm for *Insertion-sort*. State the average and worst case time complexities of this algorithm. 8+2

8. Define the term *Stack*. What operations can be done on a *Stack*? Write algorithms for all operations on a *Stack*, assuming that a linked list is used as the Stack buffer. 2+10

9. Write the algorithm for *Fibonacci Search*. Stat the average case time complexities of this algorithm. Is it better than *Binary Search*? Justify. 10+2

10. Write an efficient algorithm to find the median of a given set of data. (Hint: Quicksort) 10

B.Stat (Hons.) Second Year

ate: 5.9.09 First semester, 2009-10

ime : 3 Hours Mid Semester Examination Full Marks : 60

Statistical Methods III

[Answer as many as you can. The maximum you can score is 60]

1. Prove or disprove

(a) Maximum likelihood estimator is always a function of the minimal sufficient statistic.

(b) If $X_1, X_2, \ldots, X_n$ are iid $U(-\theta, \theta)$, $\max|X_i|$ is the minimal sufficient statistic.

(c) If $X_1, X_2, \ldots, X_n$ are iid Poisson(λ), the joint distribution of $X_1, X_2, \ldots, X_n$ belongs to a full rank exponential family and hence it is complete.

(d) If the family of distribution $\{f_\theta;\ \theta \in \Theta\}$ is not complete, the minimal sufficient statistic based on a single observation can not be complete.

(e) If the variance of an unbiased estimator $T(X_1, X_2, \ldots, X_n)$ of θ attains the Cramer Rao lower bound, the joint distribution of $X_1, X_2, \ldots, X_n$ must belong to an exponential family.

(f) If $X_1, X_2, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ (μ and σ^2 both are unknown), $\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$ is the minimum MSE estimator for σ^2. [2×6=12]

2. (a) If $X_1, X_2, \ldots, X_n$ are i.i.d. $U(\theta_1 - \theta_2, \theta_1 + \theta_2)$, find out the maximum likelihood estimators for θ_1 and θ_2. How will you modify your estimators if it is known that $\theta_2 \geq 10$?

(b) Assuming the completeness of the minimal sufficient statistic, also show that $\frac{(n-2)(X_{(n)}+X_{(1)})}{n(X_{(n)}-X_{(1)})}$ is the UMVUE for θ_1/θ_2. [6+6=12]

3. (a) Assume that in your statistics book, the number of misprints on a page follows a Poisson distribution with parameter θ. From this book, 10 pages are selected at random and the total number of misprints on those 10 pages is found to be 20. Find out the uniformly minimum variance unbiased estimate for $\exp(-\theta)$. If it is known that there are 1,0,4,2,0,3,2,2,1,5 misprints on those 10 selected pages, how will you modify your estimate ?

(b) In this problem, find out the maximum likelihood estimate of $\exp(-\theta)$ (with proper justification) when it is known that (i) $\theta < 1.5$ (ii) $1.5 \leq \theta < 2.25$ (iii) $\theta \geq 2.25$. [6+6=12]

4. (a) Let p be probability of getting 'head' when a biased coin is tossed. Ten students of B.Stat second year program were asked to toss that coin until they get both head and tail at least once. The number of tossed required by these students were recorded as $3, 2, 5, 6, 3, 4, 2, 5, 7, 3$. If it is known that $p > 1/2$, can you use the EM algorithm to find out the maximum likelihood estimate for p ?

P.T.O

(b) A and B starts a game with this biased coin. A tosses the coin twice and he wins if he gets at atleast one head. Otherwise, B wins the game. Suppose that in a series of 100 independent games, A wins 64 times. Describe how you will use the EM algorithm to find the maximum likelihood estimate of p. Also check whether your algorithm converges to the true solution of the likelihood equation. [5+7=12]

5. (a) Check whether the following statements are true.

(i) UMVUE if exists, is unique.

(ii) MLE if exists, is unique.

(b) If $X_1, X_2, \ldots, X_n$ are iid $N(\mu, 1/2)$, find the UMVUE and MLE for $e^{-(\mu-1)^2}$ and compute their variances. [4+8=12]

6. (a) If $X_1, X_2, \ldots, X_n$ are iid $N(\mu, \sigma^2)$, find out the univariate and the multivariate Cramer Rao lower bound for the variance of an unbiased estimator of σ^2. Which bound is sharper in this case ?

(b) Give an example of a two parameter symmetric distribution, where the usual method of moments can not be used to estimate the population parameters. Give reason for your answer and estimate the population parameters using the method of quantiles. [6+6=12]

INDIAN STATISTICAL INSTITUTE
B.STAT-II (2009-2010)
Theory of Probability and its Applications - III
MID-SEMESTER EXAMINATIONS
Maximum marks: 40. Time: 3 hours.

Date: 8 Sep 2009

Note: Answer as many questions as you wish.
The whole question paper carries 47 marks.
The maximum you can score is 40.

1. (a) Suppose that the random variables $X_1, X_2, \cdots, X_n$ follow the Dirichlet $D(r_1, r_2, \cdots r_n; r_{n+1})$ – distribution. Write down the bivariate density of $(X_1 + X_2, X_3 + X_4 + \cdots + X_n)$.
(b) Let X_1, X_2, X_3, X_4 be *iid* random variables each with uniform $U(0,1)$ – distribution. Let $Y_1 < Y_2 < Y_3 < Y_4$ be the ordered X_i 's. Let $U_1 = Y_1$, $U_i = Y_i - Y_{i-1}$, $i = 2, 3, 4$. Show that (U_1, U_2, U_3, U_4) has Dirichlet $D(1,1,1,1;1)$ – distribution. Using the properties of Dirichlet distribution write down the joint density of $U_1 + U_2, U_3 + U_4$. Hence find the bivariate joint density of (Y_2, Y_4). Compute $P(Y_2 < \frac{1}{3} < Y_4)$. [2+(2+2+2+2)]

2. Let $X_1, X_2, \cdots, X_n$ be *iid* random variables with the same continuous distribution with density $f(x)$.
(a) Let $(i_1, i_2, \cdots, i_n)$ be any permutation of $(1, 2, \cdots, n)$. Show that
$$P[X_{i_1} < X_{i_2} < \cdots < X_{i_n}] = \tfrac{1}{n!}.$$
(b) Compute the following probabilities:
$P(X_1 < X_2,\ X_2 > X_4)$
$P(X_1 = X_{(k)})$, where $X_{(k)}$ is the kth order statistic based on $(X_1, X_2, \cdots, X_n)$
[1+(2+2)]

3. Let $X_1, X_2, \cdots, X_n$ be indepedent random variables with exponential $\mathcal{E}(\frac{1}{\alpha})$ – density $(\alpha > 0)$. Let $Y_1 < Y_2 < \cdots < Y_n$ be the ordered X_i 's.
(a) Write down the joint density of $Y_1, Y_2, \cdots, Y_n$.
(b) Let $U_1 = Y_1$, $U_i = Y_i - Y_{i-1}$, $i = 2, 3, \cdots, n$. Show that the random variables $U_1, U_2, \cdots, U_n$ are independent. Show that the random variable $\frac{2(n-i+1)U_i}{\alpha}$ has density $\frac{1}{2}e^{-\frac{x}{2}}$ for $x > 0$ (which is the χ^2_2 – density), for each $i = 1, 2, \cdots, n$.
[2+5]

4. (a) Define convergence in probability. Show that if $X_n \xrightarrow{P} X_0$, then given any $\epsilon > 0$, there exists a real number $M > 0$ such that $P(|X_n| \leq M) \geq 1 - \epsilon$ for all $n \geq 0$.
(b) Show that X_n converges to 0 $a.s.$ if and only if for every $\epsilon > 0$
$$P(\limsup A_n(\epsilon)) = 0$$
where $A_n(\epsilon) = \{\omega : |X_n(\omega)| > \epsilon\}$, $n = 1, 2, 3, \cdots$.
(b). Show that if $X_n, Y_n, n = 1, 2, \cdots$, are random variables defined on the same sample space and $X_n \xrightarrow{P} X$ and $E(X_n - Y_n)^2 \to 0$, then $Y_n \xrightarrow{P} X$.
(c) Let X_n, $n = 1, 2, \cdots$, be independent random variables such that
$$P(X_n = \epsilon_n) = P(X_n = -\epsilon_n) = \tfrac{1}{2} \text{ for all } n$$

P.T.O

where $\sum_1^\infty \epsilon_n^2 < \infty$. Show that $\sum_1^\infty X_n$ converges $a.s.$ State carefully any theorem that you may have to use to prove this. [(1+3)+2+2+2]

5. Let X_n, $n = 1, 2, 3, \cdots$ be independent random variables. Let $S_n = X_1 + X_2 + \cdots + X_n$, $n = 1, 2, 3, \cdots$, and μ be a real mumber.
(a) Show that, if $\frac{1}{n} S_n$ converges to μ $a.s$, then $\frac{1}{n} X_n$ converges to 0 $a.s.$
(b) Use 2nd Borel-Cantelli lemma to show that for every $\epsilon > 0$,

$$\sum_1^\infty P(|X_n| > n\epsilon) < \infty.$$

(c) Let X_n, $n = 1, 2, \cdots$, be independent random variables with

$$P(X_n = +k) = \frac{1}{2\sqrt{k}},\ P(X_n = -k) = \frac{1}{2\sqrt{k}} \text{ and } P(X = 0) = 1 - \frac{1}{\sqrt{k}}.$$

Show that $\frac{X_n}{n}$ can not converge to 0 almost surely. Hence conclude that $\frac{S_n}{n}$ doesnt converge to 0 almost surely. [4+4+4]

6. Let X_n, Y_n, $n = 1, 2, \cdots$ be random variables defined on the same sample space with

$$\sum_{n=1}^\infty (X_n \neq Y_n) < \infty.$$

Let $S_n = X_1 + X_2 + \cdots + X_n$ and $S'_n = Y_1 + Y_2 + \cdots + Y_n$.
Show that $\frac{1}{n} S_n \to 0$ almost surely if and only if $\frac{1}{n} S'_n \to 0$ almost surely. [3]

Indian Statistical Institute

Mid-Semester Examination: 2009-2010

Course Name: B. Stat II

Subject Name: Biology I

Date: September 10, 2009 *Maximum Marks:* 40; *Duration: 2.5 hrs*

All questions carry equal marks, answer any five

1. a. Why do you think that ATP is termed as the energy currency of the cell? [4]
   b. The structure of the inner membrane of the mitochondrion is distinct from its outer membrane. Justify. [4]

2. Despite a fluctuating environment cells generally maintain a constant composition. In what general ways are the synthesis of building block molecules and their assembly into macromolecules regulated in a fluctuating external environment? [8]

3. Distinguish between the following:
   a. Mitochondrial and nuclear DNA [2]
   b. Lysosome and vacuole [2]
   c. Chromatin and chromosome [2]
   d. Autotrophs and heterotrophs [2]

4. a. What are the different kinds of movements exhibited by living cells? [4]
   b. What types of cellular structures are responsible for such movement? Discuss with specific examples. [4]

5 Write the metabolic steps in which genetic defects might lead to galactosemia, albinism, alkaptonuria and phenylketonuria. [8]

6. How stearic and oleic acids, with the following chemical structures $CH_3(CH_2)_{16}COOH]$ and $[CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH]$ respectively, differ in the metabolism to generate ATP ? [8]

B Stat II
Economics I
Mid Semester Examination

Maximum Marks 40
Time 3 hours
Date: 11.9.09

Answer all questions.

1. Consider an individual with a utility function $U(x,y) = x^2 + y^2$, a money income M and prices p_x, p_y.

   (a) Find the maximum level of utility and the optimum levels of x, y when $M = 10, p_x = 2, p_y = 1$.

   (b) Derive the Marshallian demand equation for x when $M = 10, p_y = 1$.

   (c) Calculate the price and income elasticities of demand for x for various ranges of p_x.

   (d) Derive the indirect utility function, that is, utility as a function of prices and income.

   (e) Derive the Hicksian demand function for x when $U(x,y) = 100$ and $p_y = 1$.

[2+3+2+4+4 =15]

2. A consumer consumes two goods x_1, x_2 and the demand function for x_1 is given by $x_1 = 2 + \frac{M}{2p_1}$. Suppose initially M = 200, $p_1 = 5$. Now suppose that p_1 falls to 4. Decompose the change in the demand for x into *income effect* and *Slutsky substitution effect*

[10]

3. Suppose leisure is a normal good and a person receives half her income from wages and the rest from bank interest. In scenario 1, the wage rate increases by 5% while bank interests do not change. In scenario 2, the wage rate remains constant but bank interest rises by 5%. In which scenario will the person work longer hours?

[5]

P.T.O

4. A consumer has utility function

$$U(x, y) = V(x) + W(y),$$
$$V' > 0, V'' < 0, W' > 0, W'' < 0.$$

Show that both x and y are normal goods.

[10]

Indian Statistical Institute
Mid-Semester Examination: 2009-2010
Course Name: BStat II
Subject Name: Physics I

Date: September 11, 2009 **Duration: 1 hr 30 mins**
Answer as many questions as you can.
Maximum Marks: 40

1. The frictional force exerted by the medium on an oscillating system is $F_d \propto -v$ if the velocity v is small. Using this information, discuss the cases of underdamping and overdamping for damped harmonic motion for the system.

   A ball is dropped from a height h. Considering the frictional force of air for small velocity, show that the height of the ball in time t is $y(t) = h - \frac{g}{\alpha}\left[t - \frac{1}{\alpha}\left(1 - e^{-\alpha t}\right)\right]$, where g is the acceleration due to gravity and α is a constant having inverse dimension of time. [5+5]

2. State the least action principle.

   Derive Euler-Lagrange's equation for describing the motion of a particle or a system of particles. [2+8]

3. Derive Newton's law of motion from least action principle.

   Two masses m_1 and m_2 are suspended over a frictionless pulley by an inextensible string, constituting an Atwood's machine. Using Lagrangian method, show that the magnitude of acceleration of either mass is given by $a = \left(\frac{m_1 - m_2}{m_1 + m_2}\right) g$, where g is the acceleration due to gravity. [4+6]

4. Prove that both the quantities $\vec{\nabla} \times (\vec{\nabla}\phi)$ and $\vec{\nabla}.(\vec{\nabla} \times \vec{A})$ are zero for any arbitrary scalar ϕ and vector $\vec{A}$. [5+5]

5. If r is the magnitude of the position vector $\vec{r}$, show that

   (i) $\vec{\nabla}\left(\frac{1}{r}\right) = -\frac{\vec{r}}{r^3}$

   (ii) $\phi = \frac{1}{r}$ is a solution of the Laplace's equation $\vec{\nabla}^2 \phi = 0$. [5+5]

B. Stat. (Second Year) Semester Examination

Statistical Methods- III

First Semester, 2009-10

Date : 19. 11. 2009   Full Marks : 100   Time: $3\frac{1}{2}$ Hours

Answer as many as you can. The maximum you can score is 100

1. (a) Suppose that $X_1, X_2, \ldots, X_n$ are independent and $X_k \sim N(\mu, \sigma^2 k^2)$ for $k = 1, 2, \ldots, n$. If μ and σ^2 both are unknown, find out minimal sufficient statistic for $\theta = (\mu, \sigma^2)$ and check whether it is complete.

   (b) Find out the best linear unbiased estimator for μ. Is it also the UMVUE ? Justify your answer.

   (c) Is it possible to find the UMVUE for σ^2 ? If possible, find it out. Otherwise, give rigorous justification for your answer. [4+4+4=12]

2. (a) In order to estimate $\theta = \int_a^b g(x)dx$ $(g(x) \geq 0 \; \forall \; x \in [a, b])$, A computes $g_0 = \max_{x \in [a,b]} g(x)$ (assume that $g_0 < \infty$) and generates n observations from the bivariate uniform distribution over the rectangle $[a, b] \times [0, g_0]$. Then, the population proportion $\theta/[(b-a)g_0]$ is estimated using its sample analog, and the estimate of θ is obtained from that. B adopts another approach for estimating θ. He generates observations $x_1, x_2, \ldots, x_n$ from $U(a, b)$ and then uses $(b-a)\sum_{1=1}^{n} g(x_i)/n$ as an estimate for θ. Which of these two approaches will you recommend ?

   (b) If $X_1, X_2, \ldots, X_n$ are i.i.d. $P(\theta)$, check whether the UMVUE of $g(\theta) = e^{-3\theta}$ attains the Cramer Rao Lower Bound ? [6+6=12]

3. (a) Consider a bi-allelic locus with two alleles A and a. Also assume that the distribution of a phenotype X depends on the underlying genotype (AA, Aa or aa) in the following manner. $X \mid AA \sim N(\mu_1, \sigma_1^2)$, $X \mid Aa \sim N(\mu_2, \sigma_2^2)$ and $X \mid aa \sim N(\mu_3, \sigma_3^2)$, where $\mu_1 > \mu_2 > \mu_3$. Suppose that a sample of size n is taken from a population having allele frequencies p and q $(p+q = 1)$, respectively for A and a, and the phenotype measurements on those n samples are observed as $x_1, x_2, \ldots, x_n$. Describe in detail how you will find the maximum likelihood estimates of all these parameters ?

   (b) Instead of two alleles, if the locus has k (> 2) alleles with allele frequencies $p_1, p_2, \ldots, p_k$, show that $\sum_{i=1}^{k} p_i^2 \leq \max_{1 \leq i \leq k} p_i$. [9+3=12]

P. T. O

(a) Let x_{1i} and x_{2i} be the marks obtained by the i-th student ($i = 1, 2, \ldots, n$) in the first and the second paper of B.Stat admission test. Assume that the student get selected if $w_1 x_{1i} + w_2 x_{2i} > w_0$, where w_0, w_1 and w_2 are some positive real numbers. Define $y_i = 1$ if the i-th student is selected and $y_i = 0$ otherwise. On this data set, if we perform logistic regression of Y on X, show that the maximum likelihood estimate of the model parameters will not exist.

(b) Suppose that we have n independent observations from a univariate distribution with p.d.f. $f(x) \propto (1 + (x - \theta)^2)^{-1}$, where $\theta \in R$ is an unknown parameter. Can you write down the Fisher's scoring algorithm to find the MLE for θ ? [6+6=12]

5. Prove or disprove the following :-

(a) MLE is always consistent.

(b) "$E(T_n - \theta)^2 \to 0$ as $n \to \infty$" is a necessary condition for T to be a consistent estimate for θ [Here $T_n = T(X_1, X_2, \ldots, X_n)$ is a statistic based on n observations].

(c) If $X_1, X_2, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ (μ and σ both are unknown), $\left[\frac{S^2}{\chi^2_{0.975,n-1}}, \frac{S^2}{\chi^2_{0.025,n-1}}\right]$ gives the shortest 95% confidence interval for σ^2, where $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2$.

(d) If ϕ_1 and ϕ_2 are two level α tests for testing $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$, any convex combination of ϕ_1 and ϕ_2 is also a level α test. [3+3+3+3=12]

6. (a) Assume that we are interested in testing $H_0 : f(x) = \frac{1}{\pi}(1 + x^2)^{-1}$ vs. $H_1 = 0.5e^{-|x|}$ based on a single observation from a univariate distribution f. Construct the MP test of size α and show that the power of this test is bigger than $1 - \exp\{-tan(\alpha)\}$. Will that still remain the MP if we want to test $H_0 : f(x) = \frac{1}{\pi}(1 + x^2)^{-1}$ vs. $H_1 : f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$?

(b) Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. $U(\theta, \theta + |\theta|)$, and we want to test $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$. Is it possible to construct a UMP test for this problem ? If possible, find the power function of the test. If not, give the reason for your answer. [9+3=12]

7. (a) Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. $U(-\theta, \theta)$. Construct a UMP test for testing $H_0 : \theta \leq 1$ vs. $H_1 : \theta > 1$. Is this UMP test unique ?. If it is so, prove the uniqueness. Otherwise, give an alternative UMP test for this problem.

(b) Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\theta, 1)$. Is it possible to construct a UMP test for testing $H_0 : |\theta| \leq 1$ vs. $H_1 : |\theta| > 1$? If possible, construct the test. Otherwise, justify your answer. [8+4=12]

8. Instead of reporting the individual marks of different students, the results of the midsem and the semester exams of an institute are summarized as below.

Marks (in %)	0-34	35-44	45-59	60-79	80-100	Total
No. of students (midsem)	9	26	33	24	8	100
Semester (semester)	11	24	27	26	12	100

It is also known that the average and standard deviation of the marks in the midsem exam were 53.1 and 18.9, respectively, whereas those for the final exams were 54.8 and 20.1.

(a) If we do not assume any parametric structure for the distribution of marks in the midsem and the semester exam, how will you test the equality of the two distributions ?

(b) If the marks distributions are assumed to be normal, can you test the equality of variances of these two distributions ?

(c) Based on your finding in (b) and using the normality, can you suggest an alternative test for the equality of the two distributions ? [6+3+3=12]

9. (a) Assume that the life time an electric bulb of a company follows an exponential distribution with mean θ. The company claims that the average life time of the bulbs is at least 1000 hours. In order to verify this claim, 100 bulbs were switched on for 1000 hours and after that period, only 40 of them were found to be in the functioning state. Propose an appropriate statistical method based on that information to test the validity of their claim.

(b) If you know the life times of the 60 bulbs (which stopped working within 1000 hours), how will you use that additional information to modify your test procedure ? [4+8=12]

10. (a) If $\mathbf{x}$ is the u-th $(-1 < u < 1)$ spatial quantile of a continuous univariate distribution F, then show that $2F(x) = (1 + u)$.

(b) A sensor received k messages from different sources, and the direction of those sources from the sensor were reported as $\theta_1, \theta_2, \ldots, \theta_k$ (measured clockwise from north), respectively. How will you find the mean and the median direction ?

(c) Assume that f is a bivariate unimodal density function with elliptic probability contours. If $\boldsymbol{\mu}$ and Σ are the mean vector and dispersion matrix of this distribution, find out the area of the smallest ellipse E with $P_f(\mathbf{X} \in E) \geq 0.75$. If you know that f is bivariate normal, how will you modify your answer ? [2+4+6=12]

INDIAN STATISTICAL INSTITUTE
First Semestral Examination : (2009-2010)
B. Stat. - Second Year
Analysis III

Date : 24. 11. 2009 | Maximum Score : 100 | Time : 4 Hours

1. This paper carries questions worth a total of 120 marks. Answer as much as you can. The maximum you can score is 100 marks.

2. Unless otherwise stated, $\mathbb{R}$ will denote the space of all real numbers.

3. You are free to use any theorem proved in the class. However, you must state a theorem used by you at least once in the answer-script.

(1) Let $X \subset \mathbb{R}^n$ be a compact set and $f : X \to X$ a map such that

$$\forall x, y \in X[|f(x) - f(y)| = |x - y|].$$

Show that f is onto X.

[14]

(2) Let $U \subset \mathbb{R}^n$ be an open set, R a compact rectangle contained in U and $f \in C^1(U)$. Show that there is a $M > 0$ such that

$$\forall x, y \in R[|f(x) - f(y)| < n^2 M|x - y|].$$

[10]

(3) (a) State the inverse function theorem.
(b) State the implicit function theorem.
(c) Use inverse function theorem to prove implicit function theorem.

[(4 + 4) + 16]

(4) (a) Find and classify the extreme values (if any) of the following function:

$$f(x, y) = y^2 - x^3.$$

(b) Find the point on the line of intersection of two planes

$$a_0 + a_1x + a_2y + a_3z = 0$$

and

$$b_0 + b_1x + b_2y + b_3z = 0$$

nearest to the origin.

[8 + 8]

P. T. O.

(5) Let V be a vector space, T a k-linear and S a ℓ-linear functional on V such that $Alt(S) = 0$. Show that $Alt(T \otimes S) = 0$

[16]

(6) (a) Define divergence ($div(F)$) and curl ($curl(F)$) of a vector field F on an open set in $\mathbb{R}^3$.

(b) Let F be a vector field on a star-convex open subset U of $\mathbb{R}^3$. Show the following:

(i) $curl(F) = 0$ if and only if there is a smooth $f : U \to R$ such that $F = \nabla f$.

(ii) $div(F) = 0$ if and only if there is a vector field G on U such that $F = curl(G)$.

[(2 + 2) + 10 + 10]

(7) Let $\gamma : [0, 1]^2 \to [0, 1]^2$ be the identity singular cube and $P, Q \in C^\infty([0, 1]^2)$. (Without invoking Stokes' theorem) prove the following:

$$\int_{\partial\gamma} Pdx + Qdy = \int_\gamma \left(\frac{\partial Q(x, y)}{\partial x} - \frac{\partial P(x, y)}{\partial y}\right)dx \wedge dy.$$

[16]

INDIAN STATISTICAL INSTITUTE
B.STAT-II (2009-2010)
Probability - III
Semestral-I Examination
Maximum marks: 60. Time: 3 hours.

Date : 27 Nov, 2009.

Note: Answer as many questions as you wish.
The whole question paper carries 70 marks.
The maximum you can score is 60.

1. $X_1, X_2, \cdots,$ are *iid* random variables with exponential $\mathcal{E}(\alpha)$ – densitty: $\alpha e^{-\alpha x} I_{(x>0)}$ $(\alpha > 0)$. Let $S_0 = 0$ and $S_n = \sum_1^n X_i, n \geq 1$. Let $t > 0$ be a fixed real number; let us define a positive integer-valued random variable(depending on t), as follows :

$$N = k \text{ if and only if } S_{k-1} < t \leq S_k\,,\ k = 1, 2, \cdots .$$

(a) Find $P(N = k, X_k \leq a)$, for $k = 1, 2, \cdots$ and $a > 0$.

(b) Show that the random variable X_N has the density :

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \alpha^2 x e^{-\alpha x} & \text{for } 0 < x \leq t \\ \alpha(1 + \alpha t)e^{-\alpha x} & \text{for } x > t \end{cases}$$

[4+6]

2. Let $(X_1, X_2, \cdots, X_r)$ follow $r-$ variate Normal with $E(X_i) = \mu, Var(X_i) = \sigma^2$ and $Cov(X_i, X_j) = \rho\sigma^2$, $i \neq j$, $i, j = 1, 2, \cdots, r$, where $-\frac{1}{r-1} < \rho < 1$.

Let $\overline{X} = \frac{1}{r}\sum_1^r X_i$, $S^2 = \sum_1^r (X_i - \overline{X})^2$.

(a) Show that $\overline{X}$ follows $N(\mu, \frac{[1+(r-1)\rho]}{r}\sigma^2)$.

(b) Show that $\frac{S^2}{(1-\rho)\sigma^2}$ follows χ^2_{r-1}.

(c) Show that $\overline{X}$ and S^2 are independent.

[2+4+4]

3. Define convergence in probability . Show that if X_n converges in probability to 0 then there is a subsequence $\{X_{n_k}\}$ of $\{X_n\}$ such that X_{n_k} converges to 0 $a.s.$

[10]

4. (a) Let $X_n, n \geq 1, X$ be random variables defined on the same sample space satisfying $\sum P(|X_n - X| \geq \epsilon) < \infty$ for every $\epsilon > 0$, then show that $X_n \to X$, $a.s.$

(b) Let $Y_n, n \geq 1,$ be independent random variables with uniformly bounded fourth central moments, i.e. for some real number $M > 0$,

$$E(|Y_n - E(Y_n)|^4) \leq M, \text{ for all } n \geq 1.$$

Let $S_n = Y_1 + Y_2 + \cdots + Y_n, n \geq 1$. Show that $E\left(\left|\frac{S_n - E(S_n)}{n}\right|^4\right) < \infty$. Using this and the conclusion of the part (a), show that $\left|\frac{S_n - E(S_n)}{n}\right| \to 0,\ a.s.$

[4+6]

(Please turn overleaf)

5. (a) Define convergence in distribution . Show that if (X_n, Y_n) are random variables defined on the same sample space for $n = 1, 2, \cdots$, such that $X_n \xrightarrow{\mathcal{D}} X$ and $Y_n \xrightarrow{P} a$, then $X_n Y_n \xrightarrow{\mathcal{D}} aX$.

(b) Let $X_n, n \geq 1$, be *iid* random variables with finite second moments, $E(X_1) = \mu$, and $Var(X_1) = \sigma^2$. Let $f : \mathbb{R} \to \mathbb{R}$, be a function such that f is differentiable at μ and $f'(\mu) \neq 0$. Show that

$$\frac{\sqrt{n}(f(\bar{X}_n) - f(\mu))}{f'(\mu)\sigma} \xrightarrow{\mathcal{D}} N(0,1).$$

[7+3]

6. (a)Prove that if X is a random variable with $E(|X|) < \infty$, then its characteristic function (c.f.) is differentiable and

$$\phi'(t) = E(iXe^{itX}).$$

(b) Let X be a random variable having the triangular density $f(x) = (1 - |x|)I_{(|x|<1)}$

(i) Show that the c.f. ϕ of X is given by: $\phi(t) = 2\frac{(1-Cos\, t)}{t^2}$.

(ii) Show that $g(x) = \frac{1}{\pi}\frac{(1-Cos\, x)}{x^2}$, $-\infty < x < +\infty$, is a density . Also verify that if Y is a random variable with density $g(x)$, then the c.f. $\psi(t)$ of Y is given by $\psi(t) = (1 - |t|)I_{(|t|<1)}$. (Hint:Use the inversion formula for Densities).

[5+(2+5)]

7. (a) Let X_n, $n \geq 1$, be independent random variables, with $P(X_n = 1) = p_n$ and $P(X_n = 0) = (1 - p_n)$, $n \geq 1$, $(0 < p_n < 1)$. What conditions on p_n are equivalent to the condition $X_n \to 0$ $a.s.$?

(b) Let X_n, $n \geq 1$, be independent random variables. Is it possible that $P\{X_n \to 0\} = \frac{1}{3}$? Give reasons.

[2+6]

INDIAN STATISTICAL INSTITUTE

Semesteral-I Examination : 2009 – 2010

B. Stat. II yr.

C & Data Structures

Date: 30th November, 2009 | Time: 3½ hrs.

Note: *This paper carries 120 marks. Answer as much as you can. You will get maximum 100 marks.*

1. Suppose **A** and **B** respectively are two sorted lists of n_1 and n_2 numbers. Write an algorithm (using binary search) for finding the median of the above $n_1 + n_2$ numbers, without merging the lists. 12

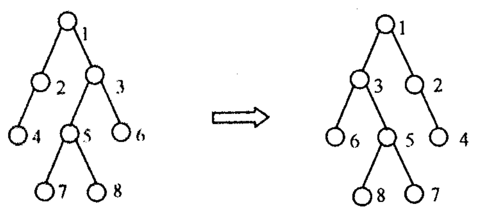
2. Write an algorithm to find the maximum and the minimum from a list of n distinct numbers with $(1.5n - 2)$ comparisons. 12

3. What is *maxheap*? Write the algorithm of *heapsort* while mentioning all the steps in heap generation. 12

4. Define the term *Binary Tree*. For a non-empty *Binary Tree* T, prove that $n_0 = n_2 + 1$, where n_0 is the number of leaf nodes and n_2 is the number of nodes with two children. Write an algorithm to find the total number of nodes in a *Binary Tree*. 2 + 4 + 6

5. Make a new version of *binary search* that uses no divisions. Use only additions to construct an auxiliary array of power of 2 (as *Fibonacci* numbers are created in *Fibonacci Search*), which helps you to write this new version. 12

6. Suppose that you are given two sequences ABCDEFGHI and BCAEDGHFI that correspond to the **preorder** and **inorder** traversals of a *Binary Tree*. Reconstruct the *Binary Tree*.
   Write an algorithm that will interchange all left and right subtrees in a *Binary Tree*. (See the example in the following figure). 6 + 6



7. Write an algorithm to test whether a given binary tree is a *binary search tree*. 12

8. Write a C program that uses two stacks. If two stacks are used separately, one might overflow while the other has sufficient unused space. Use these two stacks as a common buffer in such a way that there will be no overflow until all spaces of this common buffer are used. Declare a new data type *DoubleStack* and write programs for four operations PushA, PopA, PushB, and PopB to handle two stacks. 12

9. Write an algorithm to find the number of nodes and the height of a *binary tree*. 12

10. Write an algorithm to convert an infix algebraic expression into reverse polish notation. The expression may contain only +, –, ×, / and ^ operators and literals of unit length. 12

B Stat Second Year
Economics I
First Semester Examination (2009-10)

Maximum Marks: 60
Time: 3 hours
Date: 3.12.09

*Date.Answer **any four** questions. Each question carries 15 marks.*

1.(a) N friends went to a restaurant to have dinner. They decided that each will have a meal of her choice and the total cost will be equally split between them. The payoff of the ith friend is given by $\Pi_i = \sqrt{p_i} - \frac{1}{N}\sum p_i$ where p_i is the price of the meal consumed by her. Show that the payoff of each person is strictly lower in a Nash equilibrium than in an arrangement where each friend pays for her meal separately.

(b) Consider the following game played by two players:

	L	M	R
U	5,5	0,0	0,6
M	0,0	4,4	0,1
D	6,0	1,0	1,1

where player 1 chooses a row and player 2 chooses a column *simultaneously*. For this game find the set of *undominated* strategies. [10 – 5]

2. There are two coffee shops on a street. In each shop the cost of producing a cup of coffee is c. There are N consumers and N is large. A fraction λ of the consumers is informed about the price prevailing in each shop and chooses the one which charges the lower price; and if the shops charge the same price an informed consumer chooses a shop with equal probability. The remaining $(1 - \lambda)$ fraction, who is not informed about the prices, randomly goes to a shop with probability half-half. Finally, each consumer, whether informed or uninformed, is willing to pay a maximum R for a cup of coffee where $R > c$.

(a) Write down the profit (payoff) of each shop as a function of the prices chosen by the two shops.
(b) Determine the Nash equilibrium of the game for $\lambda = 0$ and for $\lambda = 1$.
(c) Prove that for $0 < \lambda < 1$, a pure strategy Nash equilibrium does not exist.

[3+4+8=15]

3. A firm produces an output (x) using a single factor of production labour (L) through a production function $x = L^c$, $0 < c < 1$. The wage rate is assumed to be unity and is given to the firm.

P.T.O

(a) Find total cost, average cost and marginal cost. Find the profit maximizing level of output when output price is 2 and c = ½.

(b) Suppose there are n price taking firms in the market each having a production function as given above. If market demand is given by $x^d = p^{-a}$, find the equilibrium price.

[8 + 7]

4. Consider a two-good-two-factor general equilibrium model. Show that

(a) A rise in the relative price of a good unambiguously increases the real income of the factor which is intensively used in the production of that good and unambiguously reduces the real income of the other factor.

(b) At constant commodity prices, a rise in the endowment of one of the factors, the endowment of the other factor remaining constant, raises the output level of the commodity which intensively uses the factor whose endowment has increased and reduces the output level of the other commodity.

[7 + 8]

5. Mr. A's yearly budget for his car is Rs. 100000, which he spends completely on petrol (P) and all other expenses for his car (M). M is measured in rupees, so you can take price of M to be Re 1. When price of P is 40 per liter, Mr. A buys 1000 liters per year. The price of P rises to Rs. 50 per liter and to offset the harm to Mr. A government gives him a cash transfer of Rs. 10000 per year.

(a) Write down Mr. A's yearly budget equation under the 'price rise plus transfer' situation.

(b) What will happen to his consumption of P after the 'price rise plus transfer' situation? Will he be better off or worse off?

(c) Suppose instead of the cash transfer, the government gives a certain amount of petrol free of cost to Mr. A such that he enjoys the same utility as under the cash transfer scheme. Assuming that the government has to buy P at Rs. 50 per liter, will this cost the government more, the same or less as compared with the cash subsidy scheme?

[4+4+7]

Indian Statistical Institute
First Semester Examination : 2009-10
Course Name : BStat II
Subject Name : Physics I

Date : 03/12/2009 **Maximum Marks : 90** **Duration : 3 hrs**

Note : Answer as many questions as you can.

1. (a) Prove that the conservation of energy holds good for a 3-dimensional system.
   (b) A pendulum bob is oscillating in a medium which exerts a frictional force $F_f \propto v$, where v is the velocity of the bob. If an external harmonic force of the form $F_d \propto e^{i\omega_0 t}$ acts on the bob, analyze the motion of the bob as a forced harmonic motion.
   (c) A planet revolves round the sun under the central force $F(r) \propto \frac{1}{r^2}$. Using the plane polar coordinate (r, θ), find out the equation of motion for the planet and show that the angular velocity of the planet $\dot{\theta} \propto \frac{1}{r^2}$.

   [3+7+5 =15]

2. (a) Derive Euler-Lagrange's equation of motion for a classical system.
   (b) Using the least action principle prove that the shortest distance between two points on a plane is a straight line.
   (c) A solid homogeneous cylinder of radius r rolls without slipping inside a stationary large cylinder of radius R. Show that the period of oscillation of the rolling cylinder about its equilibrium position is given by $T = \left[\frac{6\pi^2(R-r)}{g}\right]^{1/2}$.

   [7+3+5 =15]

3. (a) Using the principle of small oscillations, find out the normal frequencies of oscillation for a CO_2 molecule, considering this as a classical system.
   (b) Hence find out the displacement vector for each separate frequencies and analyze the motion of the carbon and oxygen atoms in each case.

   [7+8 =15]

4. (a) Derive the relativistic velocity addition formula. Can you explain from here that the speed of light in vacuum is the maximum attainable speed of a signal?
   (b) Find out the expression for relativistic kinetic energy and hence show that it reduces to the usual Newtonian expression for small velocity limit.
   (c) Two objects A and B travel with velocities $\frac{4}{5}c$ and $\frac{3}{5}c$ respectively (with respect to a stationary observer sitting on the earth) along the same straight line in the same direction. How fast should another object C travel between them, so that it feels that both A and B are approaching C at the same speed? Also, how much is this speed of A (or B) as measured by C?

   [5+5+5 =15]

Please turn over

5. (a) Using Stokes' theorem, prove that $\vec{\nabla} \times \vec{E} = 0$, where $\vec{E}$ is the electric field due to a distribution of charges.

(b) Show that the work done due to a continuous charge distribution can be expressed as $W = \frac{\epsilon_0}{2} \int E^2 d\tau$, where the integration is taken over the entire space.

(c) Prove that the electric potential $V \propto \frac{1}{r}$ is a solution of the Laplace's equation $\vec{\nabla}^2 V = 0$.

[5+5+5 =15]

6. (a) Show that for any two vectors $\vec{A}_1$ and $\vec{A}_2$, the following relations hold good:
(i) $\vec{\nabla}.(\vec{A}_1 \times \vec{A}_2) = \vec{A}_2.(\vec{\nabla} \times \vec{A}_1) - \vec{A}_1.(\vec{\nabla} \times \vec{A}_2)$
(ii) $\vec{\nabla} \times (\vec{\nabla} \times \vec{A}_1) = \vec{\nabla}(\vec{\nabla}.\vec{A}_1) - \vec{\nabla}^2 \vec{A}_1$.

(b) Using Biot-Savart's law show that the divergence of magnetic field for magnetostatics is zero. Can you explain its physical significance?

(c) Derive Poisson's equation for magnetic field $\vec{\nabla}^2 \vec{A} = -\mu_0 \vec{J}$, where $\vec{A}$ and $\vec{J}$ are magnetic potential and volume current density respectively.
In answering (b) and (c) you can take help of (a), if required.

[6+4+5 =15]

7. (a) Is the vanishing of the curl of electric field still valid in presence of a variable magnetic field? If yes, why? If not, find out the corresponding expression.

(b) Show that the electric field in electrodynamics is corrected to $\vec{E} = -\vec{\nabla} V - \frac{\partial \vec{A}}{\partial t}$, where V and $\vec{A}$ are scalar and vector potentials respectively.

(c) Using Lorentz gauge show that Maxwell's equations in electrodynamics can be expressed in a convenient form in terms of potentials.

[4+5+6 =15]

First Semestral Examination (B. Stat-II, Biology I, 2009), Date: 4.12.2009

Answer any five; All questions carry equal marks; Full marks = 50; Time = 2.5 hours
Answer should be brief

1. What is the chemical composition of genetic code and why the code is a triplet? How many different protein sequences, of 15 amino acids long, are possible with 20 different amino acids? Will the number of possible DNA sequences, coding all these proteins, be more than or equal to that of possible protein sequences? Why? [2+2+2+4]

2. What is meant by allele, trait, genotype and phenotype taking the example of sickle cell anemia? Although the individuals suffering from sickle cell disease die at early ages, the disease gene is not eliminated from a population, why? [10]

3. (a) A normal woman, whose father had hemophilia, marries a man. What is the chance that their second child will have hemophilia? If it is assumed that both the man and woman are heterozygous for albinism then what will be the chance that the child will suffer from both hemophilia and albinism? (Hemophilia and albinism are X-linked and autosomal recessive diseases respectively) [8]

   (b) Why are the lengths of the same gene different in bacteria and human? [2]

4. (a) How many different DNA sequences are possible with 6 codons: ATG, TAA, AGG, CGT, GGG, TAC such that (a) the sequences start with the codon ATG, end with the codon TAA and (b) there will not be five consecutive "G"s within the sequences. [5]

   (b) In a family, the husband and the wife have "A" and "AB" blood groups respectively: calculate the probability that their child's blood group will be (a) "A", (b) "B", (c) "AB". [5]

5. a) The biomolecules of living organisms are ordered into a hierarchy of increasing molecular complexity. Depict this with appropriate examples. [6]

   b). Describe the intracellular pathway most likely taken by histone proteins from synthesis to their destination. [4]

6. a). What features distinguish a prokaryotic cell from a eukaryotic cell? [6]

   b). One theory suggests that mitochondria arose during evolution from the invasion of small oxygen-using cells into large non-oxygen using cells. What evidence supports this theory? [2]

P. T. O

c). What role did cyanobacteria (i.e. blue green algae) have in the endosymbiotic hypothesis for the origin of eukaryotic cells? [2]

7. (a) Will DNA isolated from different tissues of an individual be identical? Give reasons. [3]

(b) Why may the DNA sequences in the genomes from two brothers born of the same parents not be identical? [3]

(c) In a population of 250 individuals the frequency of "A/A", "A/a" and "a/a" genotypes are 80, 100 and 70 respectively. Calculate the allele frequencies. [4]

INDIAN STATISTICAL INSTITUTE

B. Stat II: 2009 – 2010

Mid-semester Examination

Economic Statistics and Official Statistics

Duration: 3 hours | Maximum marks: 100 | 22 February 2010

Economic Statistics and Official Statistics

*(Answer question no. 1 and any **three** from the rest of the questions)*

1. Suppose the distribution of income of persons in a community follows Pareto(500,2.4). Find the ratio of shares of income of the top 25 percent people to that of bottom 25 percent people. Also derive Elteto-Frigyes measures. Prove that these measures will remain unchanged if the distribution is truncated from below at Q_1, the first quartile?
[8+12+5=25]

2. (a) State and prove the moment distribution property of lognormal distribution.
(b) Prove that

$$E|X_1 - X_2|/(2M) = 1 - 2\int_0^1 F_1 dF,$$

where X_1 and X_2 are independent and identically distributed positive valued random variable with common mean M, distribution function F and first moment distribution F_1. In particular show that this is equal to $2\Phi(\sigma/\sqrt{2}) - 1$ when the distribution is Lognormal(μ,σ^2).
[8+(10+7)=25]

3. Describe Positive and Normative Measures of Inequalities. Write down the desirable properties of a measure of inequality. Examine Coefficient of Variation in the light of these properties.
[10+5+10=25]

4. What do you mean by concentration in business and industry? Write down the criteria for a good measure of concentration in business and industry. Show how these criteria are satisfied by (i) Herfindahl – Hirschman index and (ii) Hall and Tideman index. How is the measure of concentration different from a measure of inequality? [1+3+9×2+3=25]

5. Write short notes on any two of the following:
(i) Problems with data.
(ii) Law of Proportionate Effect.
(iii) Three parameter lognormal distribution.
(iv) Properties of Lorenz Curve.
[12½×2=25]

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2009-2010
Course Name: B. Stat.-II yr.
Subject Name: Biology II

Date: 24.2.10 Maximum Marks: 50 marks Duration: 3 hours

Part A

Short questions (For many questions, there are no right or wrong answers since experiments are not yet conclusive worldwide; therefore the point of view put forward by you will be assessed for logical structure based existing knowledge). Please answer all questions.

1. This is a true story taken from Robert Weinberg's lecture at MIT (http://ocw.mit.edu). Jim Springer and Jim Lewis raised 80 miles apart were identical twins. They had all genes similar. At the age of 39 years, they discovered themselves through a chance meeting. They had all recorded phenotypes similar except one or two. For example, they both married women named Linda. They were both divorced and in their second marriages both of them married women named Betty. They both vacationed on the same half mile of beach in Florida. If I say, well that is an extreme case. Do you agree with me? If `not', can you find out a significant difference in the above fact and say `hey' it is statistically still conceivable and the fact is embedded in the description of similarity itself. Please give an explanation if you say `yes'. (4)
2. Who is the first recorded developmental biologist in recorded history? This person did an elegant experiment. He / she opened a fertilized chick egg on each successive day of its 3-week incubation and saw a thin band of cells to give an entire bird. If you redo the experiment today 30 x 10^6 times in an incubator – record the stages in a film everyday for 3 weeks and make 30 x 10^6 black and white movies. A machine analyzes each frame of the each movie and matches the pattern of development. What do you expect to see among the following observations- (4)
   i. There is only one pattern on all movies and the are same in all respects.
   ii. In 0.000001% cases, there are differences in 1 out of 30 x 10^6 movies that you made in early stages of development (First week of development).
   iii. In 0.00001% of cases, there are differences in 10 out of the 30 x 10^6 movies that you made in mid stages of development (Second week of development).
   iv. In 0.0001% of the cases, there are differences in 100 of the 30 x 10^6 movies that you made in last stage of development (Third stage of development).

   Will you ever see the first option as a biologist because that means there are no variations, so theoretically no evolution? If you see rest of three options in the results - then what it means?
3. Continuing with the question number 2, development of an organism (as taught by Prof. R. L. Brahmachary on 10 February 2010) occurs on templates and the blue print of the body is dependent on those templates. Let us say that we have 3 x 10^9 of such templates and for changing a frame in the movie one needs to make at least 3 changes (mutations) of such templates. Then when you see 0.000001%, 0.00001% and 0.0001% changes in the frames- then how many changes will you see in the template in each of the cases stated above (in this question)? From the knowledge of Statistics, why do you think that changes are more frequent in the last week of the development? (4)
4. What is your explanation- if you see reverse- i.e. 0.0001% changes in the frames in the first week, in the second week- 0.00001% and in the third week- 0.000001%? (Note: Mark the changes from above. Number of cases is absent.) Is it due to the nature of the frames of the movie that frames that have been added later cannot be edited (let us say they are linked or there is a strong linkage) - they are so essential for the story-if cut, the movie will be meaningless. The last week of development therefore seems to be plastic in nature? What is your idea? (4)
5. When you see changes in frames of 30 x 10^6 black and white movies- that means that you have broken linkage equilibrium. Now suppose, you have broken the linkages and there is a linkage disequilibrium, then what do you expect from the laws of thermodynamics? Thermodynamics tells us

 P.T.O.

that in order to grow- we need to lower the entropy level or degree of chaos. Let us assume that breaking of linkage allows more chaos- then would there be difference if breakage of linkages of templates takes place early in development versus later in development? (4)

Part-B

(Answer five questions)

1. Write major differences between rice and cow embryo / zygote development. Define plant and animal models. Mention about one plant and one animal model. What are the approximate sizes of their genomes? (6)
2. Write about few experimental studies undertaken to study growth and development of plants and animals and give sketchy details. (6)
3. Biodiversity within a given genus and species arises from Intra-specific incompatibility and Inter-specific incompatibility. Define these two terms. How do they help in speciation? What are the major challenges of embryogenesis? (6)
4. What are the general questions of developmental biology? Define embryology. What are homologous and analogous structures in development? (6)
5. Define isometric and allometric growth. Write about mathematical modeling of growth in general with example. (6)
6. Write in detail about mathematical model of organismal growth and pattern formation in plants and animals. (6)
7. Define plant breeding. Write about aims and objectives of breeding better quality crops. What are sexual groups for plant breeding? (6)
8. Write about the similarities and differences in breeding methods for self and cross pollinated plants. (6)
9. What are three major classes of signaling events during the development of animal zygote? What is a morphogen gradient? Define properties of stem cell with examples. (6)
10. What is *Agrobacterium tumifaciens*? Outline the procedure for making a transgenic plant with text and sketches. (6)

INDIAN STATISTICAL INSTITUTE

Mid –SemesterExamination :2010
Course:B-Stat II Year 2010
Subject: Economics II (Macroeconomics)

Date- 25.2.10

Duration: 150 minutes

Maximum Marks 40

Answer all questions

1. In a given period only two firms existed in the economy. Data regarding their activities and those of the government in the given period are given below:

All figures are in crores of rupees

(i)Firm 1:-

Receipts from		Expenditure on		
Sales to		(i) Purchases from		
Households	700	Labour from DomesticHouseholds		500
		Labour from foreigners	20	
Government	100	Intermediate inputs from abroad	500	
Foreigners	500	Machinery from abroad	1000	
		Firm 2	50	
Firm 2	50	(ii) Depreciation	50	

(ii)Firm 2 :-

Receipts from		Expenditure on		
Sales to		(i) Purchases from		
Households	1000	Labour from Domestic Households		1500
Government	100	Intermediate inputs from abroad	500	

Foreigners	1200	Machinery from abroad	2000
		Firm 1	50
Firm 1	50	(ii) Depreciation	100

(iii)**Subsidy received, taxes and business transfer payments made by the two firms:-**

Subsidy	200	Indirect taxes	100
		Corporate Profit Tax	20
		Business transfer	50

No inventory investment took place in either firm. Neither firm produced houses. Neither firm purchased capital goods from domestic sources. Net foreign transfer from abroad is zero.

(iv)Government Administration:-

Expenditure		Receipts	
Government Purchase from firms	200	Indirect taxes	100
Government wages	20	Corporate Profit Tax	20
Transfer	10	Personal taxes	10
Subsidy	200		

Find out

(a) NDP, using all the three methods specifying the value of each component of final expenditure, each component of factor income ,and value added of every production unit

(b) NNP and NI

(c) aggregate National Saving

(13+4+2=19)

2.Suppose , in an economy in a given year the foreign exchange stock of the central bank declined by Rs.15,000crore to accommodate the balance of payment deficit at the existing exchange rate.The net inflow of foreign loan to the domestic economy is Rs.35,000 crore. In addition foreigners purchased shares of domestic company worth Rs.15,000crore.Domestic economic agents did not purchase shares of any foreign firm. Net investment in the domestic sector is worth Rs.120,000crore, difference between NNP and private disposable income is Rs.20,000crore, ~~business transfer is zero~~, net foreign transfer is ~~also~~ zero, Government expenditure is Rs.40,000crore, household consumption is Rs.100,000crore.

Compute the values of official reserve settlement balance, capital account balance, current account balance, and private disposable income.

(1+1+1+5=8)

3. State whether the following statements are true or false with reasons:

(i) In the identity $C+I+G+X-M \equiv \text{GDP}$, M does not include imported intermediate inputs.

(ii) In one case a firm was using up completely all the inputs it bought from other firms in production. In another case it was using up completely only half of the inputs it bought from other firms. In the latter case both GDP and aggregate national saving will be higher.

(iii) To calculate the profit of a firm the value of durable inputs it used must be subtracted.

(iv) An increase in interest payment by the government means an increase in aggregate national saving.

(4.5+4.5+4.5+4.5=18)

B-STAT (II) II Year (Semester II 2009-2010)) Mid Sem. PHYSICS II

Total Marks: 50; Maximum marks that can be scored: 40 NO. of Students: 02

Group A: Quantum Mechanics

All questions carry equal marks (**7**). Answer any **five** questions.

1. How did the quantum hypothesis of Planck solve the Blackbody radiation problem? Derive the Planck Blackbody radiation law. 3+4

2. (a) Explain Photoelectric effect in the light of quantum theory.
(b) The work function of zinc is 3.6 ev. What 1s the maximum energy of the photoelectrons ejected by ultraviolet light of wavelength 3000 A? 1 $A = 10^{-8} cm$. 3+4

3. (a) Considering circular orbits, derive the expression for energy levels of the bound electron in Hydrogen atom in the Bohr model.
(b) How does the behavior of the energy levels for large quantum numbers agree with the Correspondence Principle? 4+3

4. (a) What is Wilson-Sommerfeld quantization rule?
(b) Use it to derive the energy levels of the rigid rotor? Explain the selection rule $\Delta n = \pm 1$. 2+5

5. (a) Calculate the discrete energy levels of a particle in a (one dimensional) box. Why s it difficult to observe the discreteness of energy levels for a macroscopic system?
(b) Find the energy levels of the quantum harmonic oscillator. 3+4

6. Derive the expression for the frequency of the scattered photon undergoing Compton scattering. 7

7. Show that the Fourier transform of a Gaussian function is also Gaussian function. Compute the product of he widths of the Gaussian and its Fourier transform. 7

Group B: Thermodynamics

Answer any one question.

(1) (i) Show that for a thermodynamic system,

$$C_P = C_V + [P + (\partial U/\partial V)_T](\partial V/\partial T)_P$$

(ii) Show that for an ideal gas, $C_p - C_v = R$,

where C_P, C_V and R have their usual meaning.

(iii) An ideal monoatomic gas is allowed to expand slowly until its pressure decreases exactly half its original value. Calculate the extent of change of volume if

(a) the process is isothermal,

(b) the process is adiabatic.

For this gas, $\gamma = \frac{5}{3}$ 4 + 4 + 7

(2) (i) Show that the work done in an adiabatic expansion of an ideal gas is given by

$$W = -\frac{1}{1-\gamma}(P_2V_2 - P_1V_1)$$

where P_1, V_1 and P_2, V_2 are initial and final pressure and volume respectively.

(ii) Consider a gas in a container obeying Van der Waals gas equation

$$(P + \frac{a}{v^2})(V - b) = nRT$$

The initial volume is V and isothermally it is compressed to one third of its volume. Find the work done by the gas and justifies its sign. 7 + 8

INDIAN STATISTICAL INSTITUTE
Mid-semester Examination : 2009-10
B. Stat. - Second Year
Analysis II

Date : 01. 03. 2010 Maximum Score : 35 Time :3 Hours

1. Answer any three questions from Group—A, any three from Group—B and the question in Group—C.

2. You are free to use any theorem proved in the class. However, you must state any theorem that you use at least once in the answer script.

3. Each question in Groups A and B carries 4 marks. The question in Group C carries the mark shown against it.

Group—A

(1) Show that every finitely generated subgroup of $(\mathbb{Q}, +)$ is cyclic.

(2) Let G and H be finite groups. Assume that there is an epimorphism $\varphi : G \to H$. Show that the order of H divides the order of G.

(3) Let $\sigma = (i_1, \ldots, i_k) \in S_n$ be a cycle. We say that the length of σ, denoted by $|\sigma|$, is k. Let

$$\sigma = \sigma_1 \cdots \sigma_m \in S_n$$

be a decomposition of σ into disjoint cycles $\sigma, \ldots, \sigma_m$ of lengths $p_1, \ldots, p_m$ respectively. Show that the order of σ equals the least common multiple of $p_1, \ldots, p_m$.

(4) Let R be the subgroup of all rotations in the dihedral group D_{2n} of order $2n$ and H a subgroup of D_{2n} containing R. Show that either $H = R$ or $H = D_{2n}$.

Group—B

(1) Let R be an integral domain (not necessarily having 1), Q its quotient field and $\varphi : R \to Q$ the canonical imbedding of R into Q. Suppose $\mathbb{F}$ is a field and $\psi : R \to \mathbb{F}$ a homomorphism. Show that there is a unique homomorphism $\eta : Q \to \mathbb{F}$ such that $\psi = \eta \circ \varphi$.

(2) Let R be an integral domain with 1. Suppose there is a $x \neq 0$ in R such that $mx = 0$ for some non-zero integer m. Show that

there is a unique positive prime p such that $px = 0$ for every $x \in R$.

(3) Show that every finite integral domain (not necessarily having a 1) is a field.

(4) Show that the quadratic ring of integers $\mathbb{Q}[\sqrt{-5}]$ is not a Euclidean domain.

Group—C

(1) Let R be a PID. Show the following.

(a) Suppose

$$p_1 \cdots p_k = q_1 \cdots q_l,$$

p_i's and q_j's irreducibles in R. Show that $k = l$ and each p_i is an associate of some q_j and each q_j is an associate of some p_i.

(b) Show that R satisfies the **ascending chain condition** (in short, **a.c.c**). i.e., there is no strictly increasing infinite sequence of ideals in R.

(c) Show that every PID is a UFD.

[4 + 4 + 10]

STATISTICAL METHODS IV
B Stat 2nd Year (2009-10)
Mid Semester Examination

Date 03.03.2010

Time 2 hours

Total Marks 40

1. Consider the usual linear regression model:

$$y_i = \alpha + \beta x_i + e_i; \ i = 1, 2, \ldots, n$$

where e_i's are i.i.d. $N(0, \sigma^2)$. Suppose $\hat{\alpha}$ and $\hat{\beta}$ are the least squares estimates of α and β respectively and $Q = \sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta} x_i)^2$.

(a) Obtain the joint distribution of $(\hat{\alpha}, \hat{\beta})$.

(b) Show that $\bar{y}, \hat{\beta}$ and Q are independently distributed and Q/σ^2 follows a chi-square distribution with $(n-2)$ d.f.

(c) Using (a) and (b), obtain a $100(1-\alpha)\%$ confidence interval for β.

(7+8+5=20)

2. Suppose $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ is a random vector with density f given by:

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} c & \text{if } x'x \le r^2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $c = \dfrac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}} r^n}$

(b) If $Y = \underset{\sim}{X}' \underset{\sim}{X}$, compute $E(Y)$ (5+5=10)

3. At a particular gene, an individual can have one of three profiles. The probabilities of these three profiles are $p^2, 2pq$ and q^2 respectively where $o < p < 1$. An individual is affected with a disease if and only if he / she has the third genetic profile. In a random sample of N individuals, it was found that M of them were unaffected. Find the asymptotic distribution of the m.l.e. of p based on the above data. 10

7	26.0	5
8	27.0	4
9	24.8	7
10	21.4	4
11	23.9	3
12	24.1	5
13	27.0	4
14	26.8	6
15	26.4	2
16	27.0	5
17	28.4	3
18	25.4	7
19	26.2	6
20	27.0	5

Obtain the X-Bar - R chart for controlling the volume of soft drink bottles.

Q.3. 10 + 5 = 15

i) Define: Producer's Risk, Consumer's Risk & a Process.
ii) What is a Cp index? Using the data of Q3(iii) obtain the Cp of the filling process and comment about the process.

Q.4. 15

A paint manufacturer produces both interior and exterior paints from two raw materials M1 & M2. The followiing table provides the basic data for its production planning:

Tons of rawmaterial per ton of

	Exterior paint	Interior Paint	Max. daily availability (ton)
Rawmaterial M1	6	4	24
Rawmaterial M2	1	2	6
Profit/ton (Rs.1000/-)	5	4	

P.T.

Formulate this production planning problem as an L.P problem with the objective of maximising total profit. How many basic feasible solutions can exist for this problem.

Solve this problem using Simplex Algorithm.

INDIAN STATISTICAL INSTITUTE
B.STAT-II (2009-2010)
Probability - III
Back paper examination
Maximum marks: 100. Time: 3 hours.

Date : . 22.4.10

Note: Answer all questions.

1. Let the random variables (X, Y) have a joint density given by :

$$f(x,y) = \frac{1}{2\pi}\frac{1}{\sqrt{(1+x^2+y^2)^3}}, \qquad -\infty < x, y < \infty$$

(a) Show that the marginal densities of X is : $\frac{1}{\pi}\frac{1}{1+x^2}, \quad -\infty < x < +\infty.$

(b) Make the polar transformation $(X, Y) \mapsto (R, \Theta)$ where

$$X = R\,Cos\Theta \text{ and } Y = R\,Sin\Theta,\ 0 < R < \infty,\ 0 < \Theta < 2\pi.$$

Show that R and Θ are independent random variables, where R has the density $g(r) = r(\sqrt{(1+r^2)^{-3}})$, $r > 0$ and Θ is uniformly distributed on $(0, 2\pi)$.

[5+5]

2. $X_1, X_2, \cdots, X_n$ are *iid* random variables with a common continuous distribution with density $f(x)$ and distribution function $F(x)$. Let $U_1 < U_2 < \cdots < U_n$ be the order statistic based on $(X_1, X_2, \cdots, X_n)$.

(a) What is the joint density of $(F(U_1), F(U_2), \cdots, F(U_n))$?

(b) Show that $(F(U_1), F(U_2) - F(U_1), F(U_3) - F(U_2), \cdots, F(U_n) - F(U_{n-1}))$ has a Dirichlet distribution. What are the parameters of this distribution?

(c) Find the density of $\frac{F(U_n)-F(U_2)}{F(U_n)-F(U_1)}$.

[3+5+7]

3. Define convergence in distribution. Show that if X_n converges in probability to a random variable X, then X_n converges in distribution to X.

[3+12]

4. (a) Show that X_n converges almost surely to 0 if and only if for every $\epsilon > 0$

$$P(\limsup A_n(\epsilon)) = 0$$

where $A_n(\epsilon) = [|X_n| > \epsilon]$.

(b) Let for each n, X_n be a $0-1$ valued random variable defined on the same sample space. Show that $X_n \to 0$ *a.s.* if $\sum_1^\infty P(X_n = 1) < \infty$. Show that the converse also holds if X_n, $n = 1, 2, \cdots$ are mutually independent.

[6+(2+2)]

5. Let X_n, Y_n, $n = 1, 2, \cdots$ be random variables such that $\sum_{n=1}^\infty P(X_n \neq Y_n) < \infty$.

(a) Show that $P(\liminf A_n) = 1$, where $A_n = \{\omega : X_n(\omega) = Y_n(\omega)\}$.

(b) Let $S_n = X_1 + X_2 + \cdots + X_n$ and $S_n' = Y_1 + Y_2 + \cdots + Y_n$.

Use (a) to show that $\frac{1}{n}S_n \to 0$ almost surely if and only if $\frac{1}{n}S_n' \to 0$ almost surely.

[5+5]

(Please turn overleaf)

6. (a) Let $X_1, X_2, \cdots, X_n$ be independent random variables with $E(X_r) = 0$ and $E(X_r^2) = \sigma_r^2 < \infty$ for $r = 1, 2, \cdots, n$.
Show that

$$P\left(\max_{1 \le k \le n} |S_k| \ge \epsilon \right) \le \frac{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}{\epsilon^2}$$

where $S_k = X_1 + X_2 + \cdots + X_k$, $k = 1, 2, \cdots, n$.

(b) Show that if X_n, $n \ge 1$, are independent random variables with $E(X_n) = 0$ and $\sum_{n=1}^{\infty} E(X_n^2) < \infty$, then $S_n = \sum_{i=1}^{n} X_i$ converges $a.s.$

[8+12]

7. Show that the characteristic function of a random variable having the density $\frac{1}{2}e^{-|x|}$, $-\infty < x < \infty$, is $\phi(t) = \frac{1}{1+t^2}$. Using this and invoking the inversion formula for densities find the characteristic function of the Cauchy density $\frac{1}{\pi}\frac{1}{1+x^2}$.

[4+6]

8. Let $X_1, X_2, \cdots$ be *iid* random variables with $X_n \sim \chi_n^2$, (Chi-square distribution with d.f. $= n$),

(a) Show that $\frac{\sqrt{n}(\frac{X_n}{n} - 1)}{\sqrt{2}} \xrightarrow{\mathcal{D}} N(0, 1)$.

(b) Show that $\sqrt{2n}(\sqrt{\frac{X_n}{n}} - 1) \xrightarrow{\mathcal{D}} N(0, 1)$.

[3+7]

INDIAN STATISTICAL INSTITUTE

B. Stat II: 2009 – 2010
Second Semester Examination
Economic Statistics and Official Statistics

Duration: 3 hours | Maximum marks: 100 | Date: 29.04.2010

*[Use separate answer scripts for Group A and Group B. Answer question no. 1 and any **three** of the rest of the questions in Group A and answer all questions in Group B]*

<u>Group A: Economic Statistics</u>

1. Suppose in a population each person is put in one of two groups according to whether the income of the person is below or above a certain critical level. The mean incomes, Lorenz Ratios and the numbers of persons in both the groups are known. Find the overall Lorenz ratio. [20]

2. Discuss the main considerations in defining and calculating poverty rates from household level data. [20]

3. (a) Define and compare fixed-based versus chain-based index numbers. When do the two types of indices become same?

   (b) Write down the five axioms of the economic-theoretic approach to Price Index Numbers. Prove that none of these is superfluous or impossible given the other four axioms.

   (c) Define and discuss the bounds of cost of living index numbers. [5+9+6=20]

4. Describe in detail the economic and statistical criteria for choosing an Engel Curve. [20]

5. Define Engel elasticity of demand. How does it help in classifying the commodities? Discuss the different methods of estimating it in a constant elasticity Engel curve. [20]

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

   (i) Universality of Pareto Law.

   (ii) A stochastic model leading to Lognormal distribution.

   (iii) Elasticity of Substitution.

   (iv) Cobb Douglas Production Function. [10×2=20]

PTO

Group B: Official Statistics

1. (a) Write down the ten fundamental principles of Official Statistics. [10]

Or

(b) Write down the Millennium Development Goals along with the targets envisaged in each Goal. [10]

2. (a) Describe the structure and functions of Ministry of Statistics and Programme Implementation (MoS&PI). [10]

Or

(b) Fill in the gaps in the following sentences: [10]

(i) The United Nations Statistical Commission was established in the year ...

(ii) The concerned Minister of State in the MoS&PI is ...

(iii) The Census Act was first enacted in India in the year ...

(iv) The first census in India was conducted in the year ...

(v) The Office of the Registrar General is under the Ministry of ...

(vi) A Main worker must work for at least ... months.

(vii) The Perinatal Mortality Rate is the sum of ... deaths and ... deaths per 1000 births.

(viii) The Index of Industrial Productions (IIP) in India is confined to ..., ... and ... sectors only.

(ix) The number of items in the calculation of IIP with base 1993-94 is ... which is clubbed into ...

(x) The Formulae for the Index of Industrial Production is ...

Indian Statistical Institute

B. Stat – II: 2009-10

Semester Examination

Statistical Methods IV

Date: 03.05.2010 Duration: 3 hours

Marks: 75

(The question paper carries 100 marks. Maximum you can score is 75)

1. Let each of the variables $X_1, X_2, \ldots \ldots, X_{2p}$ has a common mean and variance. The correlation coefficient between any pair of them is also the same and positive. Calculate the correlation coefficient ρ_{YZ} between $Y = X_1 + X_2 + \ldots \ldots + X_p$ and $Z = X_{p+1} + X_{p+2} + \ldots \ldots + X_{2p}$. Also find ρ_{YZ} when $p \to \infty$. (6)

2. (a) What do you mean by a variance stabilization transformation of a statistic? Under appropriate assumptions, derive the asymptotic distribution of $Z = tanh^{-1} r$, where r is the sample correlation coefficient based on a random sample of size n from a bivariate normal population.

   (b) Let there be 4 independent bivariate normal populations with correlation coefficients ρ_1, ρ_2, ρ_3 and ρ_4, and let n_i be the size of the sample drawn from the i-th population, $i = 1, \ldots, 4$. Develop a testing procedure to test $H_0 : \rho_1 = \rho_2 = \rho_3 = \rho_4$. (3+6+7=16)

3. Let $X_{(1)}, X_{(2)}, \ldots \ldots, X_{(n)}$ be the order statistics based on a sample drawn from an absolutely continuous distribution function $F(x)$. Define $Y_{(i)} = F(X_{(i)}); i = 1, \ldots, n$ and let ξ_p be the population p-th quantile of X and $\hat{\xi}_p = X_{(r)}$ for some $r,\ 1 \le r \le n$, be the corresponding sample quantile, where $r = [(n+1)p]$.

   (a) Noting the fact that $F(X) \sim U(0,1)$, show that $\hat{\xi}_p$ follows asymptotic normal distribution with mean ξ_p and variance $\frac{p(1-p)}{n\{f(\xi_p)\}^2}$.

   (b) Use the above asymptotic distribution to test $H_0 : \theta = 0$ against $H_1 : \theta \neq 0$ when p.d.f. of X is given by $f(x) = \frac{1}{\pi} \cdot \frac{1}{1+(x-\theta)^2}$, $-\infty < x < \infty$. (10+6=16)

4. A DNA sequence comprises of nucleotides A, T, G, and C; each occurring independently with probabilities p_1, p_2, p_3, and p_4, respectively, where $\sum_{i=1}^{4} p_i = 1$. Given a DNA sequence comprising n nucleotides, obtain the Fisher's information matrix for the vector $(p_1, p_2, p_3)'$. (8)

5. In a genetic study on Type 2 diabetes, it was found that among 150 Type 2 diabetes patients, the frequencies of the genotypes AA, Aa, and aa were 73, 64, and 13, respectively, while among 190 normal individuals, the frequencies of the three genotypes were 48, 42, and 100, respectively. Do the data provide evidence that the genotypic distributions of Type 2 diabetes patients and normal individuals are the same?

(8)

6. Suppose $X \sim N_p(0, \sigma^2 I)$. If A is a symmetric idempotent matrix, obtain the distribution of $X'AX$, and that of $X'AX / X'X$.

(5+4)

7. Consider a random sample of size n from a bivariate normal distribution with parameters $(\mu, \mu, \sigma^2, \sigma^2, \rho)$. Obtain the m.l.e of μ, σ^2, and ρ.

(8)

8. Using random observations from $U(0,1)$, explain how you would generate a pair of observations (X,Y) such that $(X,Y) \sim N_2(2, 3, 0.81, 1.21, 0.7)$. Using this pair, thus drawn from the above-mentioned bivariate normal distribution, explain how you would generate (i) two observations U and V, where $U \sim N(5,5)$ and $V \sim N(3,3)$, and (ii) one observation from a Cauchy distribution.

(6+3+3)

9. It has been postulated that the variance of the end-semester scores in a subject is twice that of the mid-semester scores. A student, who suspects that this ratio is in reality less than what is postulated, collects data on the mid-semester and end-semester scores of 4 randomly chosen students as follows: (13,29), (18,31), (18,36) and (19,33). Is the student's suspicion justified? State all your assumptions clearly.

(7)

10. The Hardy-Weinberg Equilibrium law implies that at any locus with alleles A_1, A_2, and A_3, the probability distribution of the genotypes is as follows: $P(A_iA_i) = p_i^2$ for all i, $P(A_iA_j) = 2p_i p_j$ for all $i \neq j$, where $\sum_i p_i = 1$, $P(A_i) = p_i$, $i = 1,2,3$. Data are collected on 146 patients suffering from Alcoholic Cirrhosis. The genotype frequencies at a marker acid phosphatase (ACP) are as follows:

Genotype	Frequency
A_1A_1	5
A_1A_2	55
A_1A_3	15
A_2A_2	65
A_2A_3	5
A_3A_3	1

Do the above data provide evidence that the marker ACP is in Hardy-Weinberg Equilibrium?

(10)

INDIAN STATISTICAL INSTITUTE

B.Stat. II Year Course, 2009-10
Second Semestral Examination
Subject: Demography

Date: 06.05.2010

Duration: 2 hours

Answer as much as you can. Maximum marks you can score is (50)

1. Show that $m_x = \mu_{x+1/2}$, where μ_x is the force of mortality and m_x is the age specific death rate at age x. (4)

2. (a) Show that the average age at death of those persons who die between age x and age y is $x + \{T_x - T_y - (y-x)l_y\} \big/ (l_x - l_y)$, where l_x and T_x are the number of persons at exact age x and number of years lived by the cohort after attaining age x respectively. (5)

(b) What are the chances that a son just born to a mother aged 25 years and a father aged 30 years will be alive for 15 years but will be

(i) orphaned only by his father

and (ii) orphaned by at least one of his parents. (3 + 3)

3. Describe the estimation procedure of Net Migration for inter-censal period. (4)

4. (a) Write whether the following statements are true or false : (8)

(i) Crude Birth Rate(CBR) is not very sensitive to small fertility change.

(ii) Although CBR is affected by the age composition of population and level of fertility but is not affected by the age pattern of fertility.

(iii) The ratio of total number of yearly births to the total female population in the reproductive period is called general fertility rate.

(iv) General Fertility Rate (GFR) is a more acceptable measure of fertility level because it controls the variations in age composition within the reproductive age range.

(v) Distribution of fertility levels in the child-bearing ages is best revealed by the computation of Age Specific Fertility Rate(ASFR).

(vi) ASFRs are widely affected by the variations in population composition.

(vii) Computation of Total Fertility Rate(TFR) of a hypothetical cohort accounts attrition due to death occurred within the reproductive period.

(viii) TFR is also affected by the age structure of women under study.

(b) In a community and within a given period the following information is obtained.

Age group	Female population	Live births	Female survival rates
15 - 19	79865	8813	0.90503
20 - 24	63315	17620	0.90048
25 - 29	51860	14585	0.89472
30 - 34	44440	10235	0.88799
35 - 39	38795	7569	0.88009
40 - 44	32250	2760	0.87013
45 - 49	26720	381	0.85705

Calculate Gross Reproduction Rate(GRR) and Net Reproduction Rate (NRR). Assume that the sex ratio at birth is 105 males per 100 females.

(6)

5. (a) Describe how Verhulst (1840) derived Logistic model for population growth. State assumptions clearly.

(10)

6. State the assumptions for a population to be stable.
Derive Lokta and Dublin's integral equation for stable population analysis.
Hence find the value of growth rate from this model. (12)

INDIAN STATISTICAL INSTITUTE
Semestral Examination : 2009-10
B. Stat. - Second Year
Elements of Algebraic Structures

Date : 11. 05. 2010 | Maximum Score : 100 | Time : 4 Hours

1. Answer all the questions.

2. You are free to use any theorem proved in the class. However, you must state any theorem that you use at least once in the answer script.

3. The paper carries a total of 110 marks. The maximum score on each question is shown against them.

(1) (a) State and prove the class equation for finite groups.

(b) Let G be a finite group, p a prime number and $k \geq 1$. If $p^k || G|$, show that G has a subgroup of order p^k.

(c) If p is a prime number, show that every group of order p^2 is abelian.

[10 + 10 + 8]

(2) (a) Let $m, n > 1$. Show that $\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic if and only if $(m, n) = 1$.

(b) Show that a group of order 9 is isomorphic to either $\mathbb{Z}_9$ or to $\mathbb{Z}_3 \times \mathbb{Z}_3$.

(c) Show that a group of order 99 is isomorphic to one of $\mathbb{Z}_{99}$, $\mathbb{Z}_9 \times \mathbb{Z}_{11}$ and $\mathbb{Z}_3 \times \mathbb{Z}_{33}$.

[6 + 8 + 10]

(3) Show that every PID is a UFD. [10]

(4) (a) Let $\mathbb{F}$ be a field. Show that $f(x) \in \mathbb{F}[x]$ has a multiple root if and only if $f(x)$ and its derivative $f'(x)$ are relatively prime in $\mathbb{F}[x]$.

(b) Determine if $8x^3 - 6x - 1 \in \mathbb{Q}[x]$ has a multiple root or not.

(c) Determine if the polynomial

$$2x^2y - 3xy^2z + 4z^5$$

is irreducible in $\mathbb{Z}[x, y, z]$ or not.

[6 + 6 + 6]

(5) (a) If $\mathbb{F}$ is a finite field, show that there is an integer $n > 1$ such that

$$\mathbb{F} = \{a^n : a \in \mathbb{F}\}.$$

P. T. O.

(b) Let a be an algebraic number and r a rational number. Is a^r an algebraic number? Justify your answer.

[6 + 6]

(6) (a) Let $\mathbb{F}$ be a field and $p(x) \in \mathbb{F}[x]$ be irreducible in $\mathbb{F}[x]$. Show that all the roots of $p(x)$ in its splitting field are of the same multiplicity.

(b) Let $\mathbb{K}$ be an algebraically closed field and $\mathbb{F}$ a subfield of $\mathbb{K}$ such that $\mathbb{F}$ is not algebraically closed and for every algebraically closed field $\mathbb{L}$ with $\mathbb{F} \subset \mathbb{L} \subset \mathbb{K}$ either $\mathbb{L} = \mathbb{F}$ or $\mathbb{L} = \mathbb{K}$. Show that $\mathbb{K}$ is an algebraic extension of $\mathbb{F}$.

[8 + 10]

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2009-2010
Course Name: B. Stat.-II yr.
Subject Name: Biology II

Date: 13.5.10 Maximum Marks: 50 marks Duration: 3 hours

Part A

(This question is compulsory)

1. A prey population follows the laws of logistics growth. The predator population consumes the prey population following Holling type II functional response and the death rate of predator is density dependent.
   a. Write down a basic mathematical model portraying the above system.
   b. Find out the biologically feasible equilibrium.
   c. Find out the condition(s) for which both prey and predator will persist.
   d. Give the biological justification(s) of your results. (4+2+6+3)

Part B

(Answer 7 questions)

1. Draw the major differences between rice and cow embryo / zygote development. Define plant and animal models. Name one typical plant and animal model each. What are the approximate sizes of their genomes? (5)
2. Define Agricultural biodiversity. What is the scope of agricultural biodiversity? Write short notes on the following – (a) Agro-ecosystems vs. natural ecosystems, (b) Human dependency on diversity (c) Genetic erosion in agricultural and livestock biodiversity. (5)
3. Biodiversity within a given genus and species arises from Intra-specific incompatibility and Inter-specific incompatibility. Define these two terms. How do they help in speciation? What are the major challenges of embryogenesis? (5)
4. Define isometric and allometric growth. Discuss mathematical modeling of growth in general with example. (5)
5. Describe Turing's reaction-diffusion model and its role in pattern development in animals and plants. (5)
6. Define plant breeding. What are the aims and objectives of breeding better quality crops? What are sexual groups for plant breeding? (5)
7. What are the three major classes of signaling events during the development of animal zygote? What is a morphogen gradient? Define properties of stem cell with examples. (5)
8. What are the common mechanisms of plant disease resistance. Write a short note on plant breeding for disease resistance with two examples. (5)
9. Define plant nutrition. Write about the processes important for plant nutrition. Write about the functions of five (each type) macro- and micronutrients in plant nutrition. (5)
10. Define Plant disease resistance and tolerance. Describe the plant disease triangle concept with a diagram. Write a short note on plant immune system and plant defense signal transduction. Give two examples of plant disease resistant varieties. (5)

Indian Statistical Institute
Second Semester Examination (2010-2011)
Course – B Stat II Year
Subject: Economics II (Macroeconomics)
Maximum marks –60

Answer all questions. Date: 14.5.10 **Duration 3 hours**

1 **(i)** Consider a simple Keynesian model for a closed economy with government. Suppose $\left(\delta C / \delta YD\right) = 0.8$, $\left(\delta I / \delta y\right) = 0.3$. Can you design a tax function that will ensure stability of equilibrium in this model? Is this tax function unique? (Assume transfer payment $R = 0$. All the symbols have usual meanings).

(ii) A Simple Keynesian model for a closed economy with government is given by the following set of equations,

$C = 160 + 0.8YD, YD = y - T + R$

$I = 120 + 0.3y$

$G = 30$

$R = 0$

$T = 20 + 0.5y$

(a) Compute the equilibrium level of y.

(b) Consider a different situation where total tax collection, T, is given by the above function until it reaches the ceiling denoted $\overline{T} = 400$. Once this level is reached it is kept at that level for all higher levels of y.

What implication will this have for the aggregate demand function and the equilibrium of the model?

(c) How does your answer to part (b) change if $\overline{T} = 620$ instead? (8+16=24)

2 **(i)** Show that the supply of broad money is equal to high powered money (H.P.M) plus commercial bank credit.

Suppose, in an economy $H.P.M = 1000$, currency deposit ratio $= 1/3$, cash reserve ratio $= 1/2$. Compute the amount of broad money. Depict the consolidated balance sheet of commercial banks.

Now suppose, following an increase in $H.P.M$ by 500 units broad money is found to increase by 700 units only due to inadequate demand for commercial bank credit. Compute the increase in broad money, following a further increase in $H.P.M$ by 500 units from this situation.

(ii) Examine the impact of an autonomous shift in the import function in the Mundell Fleming model in a flexible exchange rate regime under perfect capital mobility. (16+10 =26)

P.T.O

3. Suppose in an IS-LM model interest rate (r) adjusts instantaneously to correct any disequilibrium in the money market, while the change in the production level (y) that corrects any disequilibrium in the commodity market takes a longer time.

Now consider an increase in real balance supply by one unit.

(i) In an IS-LM diagram, trace the path along which the system moves from the initial to the new equilibrium, following this change.

(ii) Suppose r is found to drop by $1/62.5$ units before one can observe any change in y following the above said change in money supply. This change also causes the LM curve to shift horizontally by 4 units.

In the same IS-LM model an increase in government expenditure (G) is found to increase the level of y by 1000 units. Compute the associated change in r.

(4+12=16)

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination : (2009-2010)

B.Stat. II Year

Physics II

Group A

Date: 14.5.10 Maximum Marks 30 Duration 1.30 hour

Answer any two questions

1. i) Find the efficiency of reversible Carnot's engine operating between temperature T_1 and T_2 where $T_1 > T_2$. Argue why the efficiency of any other reversible or irreversible engine can not go beyond that of reversible Carnot's engine.
ii) Let in one case the temperature T_1 is increased by 10 percent, keeping T_2 fixed. In another case T_2 is increased by 10 percent keeping T_1 fixed. In which case the efficiency is higher?
iii) A heat engine operating between 1000^0C and 25^0C produces a quantity of work which is entirely used to run a refrigerating machine operating between 0^0C and 25^0C. Calculate the ratio of the heat absorbed by the engine to that absorbed by the refrigerating machine.

$7 + 3 + 5$

2. a) State the second law of thermodynamics and using it derive the following relation

$$\sum \frac{Q_i}{T_i} \leq 0$$

where the system undergoes a cyclic transformation during which it exchanges (absorbs or receives) heat energy with a set of n sources T_1, T_2,....T_n. The amount of heat exchanged between the system and these sources are Q_1,Q_2,.....Q_n respectively. Q's are positive when system absorbs it and otherwise negative.
b) Show that if the above cyclic process is reversible, then

$$\sum \frac{Q_i}{T_i} = 0$$

c) A system having constant heat capacity C_p and a temperature T_i is put into contact with a reservoir at temperature T_f. Equilibrium between the system and the reservoir is established at constant pressure. Determine the entropy change for the system alone as well as for the system and reservoir combined.

$6 + 4 + 5$

 P.T.O

3. a) For a thermodynamic system, derive the following equation;

$$(\partial U/\partial V)_T = T(\partial p/\partial T)_v - p\ ;$$

all the symbols having their usual meaning. Show that for a perfect gas, the internal energy U does not depend on volume.

b) For a liquid-vapour system, derive the following equation;

$$\frac{dp}{dT} = \frac{\lambda}{T(v_2 - v_1)}$$

where, λ is the latent heat for vaporization, v_1 and v_2 are volume per unit mass for the liquid and the vapor respectively and other symbols have their usual meaning. Using this equation explain why the boiling point of water decreases with the decrease of pressure.

c) Derive the following distribution law;

$$n_i = N\frac{e^{-\beta\epsilon_i}}{\sum_i e^{-\beta\epsilon_i}}$$

for a system of N particles. The total energy of the system is sum of energies of individual particles, and n_i is the no. of particles having energy ϵ_i and β is a positive constant.

4 + 6 + 5

Indian Statistical Institute

B. Stat – II, 2009-10

Back Paper

Statistical Methods IV

ate: 05·08·10 Duration: 3 hours

Marks: 100

(a) Show that the sampling distribution of the sample correlation coefficient in samples from a bivariate normal population depends only on the population correlation coefficient and on no other parameters.

(b) The joint score distribution in Mathematics and Statistics is a bivariate normal with known means but all other parameters unknown. In a random sample of size 300, it is found that only for 120 pairs of observations, both the x-value and y-values exceed their respective means. From this observed data suggest an estimate of ρ, the population correlation coefficient. (10+15=25)

(a) Obtain a general expression for the variance stabilizing transformation of a statistic T whose mean is θ and variance $g(\theta)$.

(b) Deduce the Z-transformation of the sample correlation coefficient, assuming samples are drawn from a bivariate normal population.

(c) Three independent samples from a bivariate normal population yield the following:

Sample size:	20	30	35
Sample corr. coeff.:	0.65	0.59	0.63

Suggest a combined estimate for the population correlation coefficient using Z-transformation. (7+9+9=25)

(a) Consider the usual multiple linear regression model with normal homoscedastic errors. Show that the least square estimators of the regression coefficients are actually their best linear unbiased estimators.

(b) Suppose $Y \sim N_p(0,\sigma^2 I)$. State and prove a necessary and sufficient condition for the quadratic form $Y'AY$ to have the chi-square distribution. (10+15=25)

. (a) Define frequency chi-square statistic for testing goodness of fit and derive its sampling distribution for large samples.

(b) A statistical package that generates random numbers is to be tested for its randomness. The package is used to generate 100 single digit integers between 0 and 9, and the frequencies of the integers are summarized below.

Integer observed:	0	1	2	3	4	5	6	7	8	9
Frequency:	7	10	11	8	9	8	10	6	14	17

At 1% level of significance, test whether there is any reason to suspect the unbiasedness (or randomness) of the package. (15+10=25)