INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2016-17 (First Semester)
Bachelor of Statistics (B. Stat.) II Year
Probability Theory III

Teacher: Parthanil Roy

Date: 29/08/2016 **Maximum Marks: 40** **Duration: 10:30 am - 1:00 pm**

Note:

- **Please write your roll number on top of your answer paper.**
- **Show all your works and write explanations when needed. Maximum you can score is 40 marks.**
- **You may use any fact proved in the class but do not forget to quote the appropriate result. However, if you use a result that was given as an exercise in the class, then you have to provide a *complete proof* of that result based on whatever has been covered in the class. This, of course, does not apply to the results for which proofs will be given in future (i.e., either later in this semester or in a future course).**
- **In this examination, you are *not* allowed to use class notes, books, homework solutions, list of theorems, formulas etc. You are also requested to switch off your cell phone and keep them inside your bag. Failing to follow the examination guidelines, copying in the examination, rowdyism or some other breach of discipline or unlawful/unethical behavior, etc. are regarded as unsatisfactory conduct. Any student caught cheating or violating examination rules will get a zero in this examination.**

State whether each of the following statements is true or false. If it is true, give a detailed proof. On the other hand, if it is false, produce a counter-example with full justification. Each question is worth 10 marks. If you both prove and disprove a statement, you will get a zero in that problem.

1. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$ distribution and B is a symmetric matrix. Define the random vector $\mathbf{U} = (X_1 - \mu, X_2 - \mu, \ldots, X_n - \mu)^T$. If $P(\mathbf{U}^T B \mathbf{U} < 0) = 1$, then all the eigen values of B are real and non-positive.

2. $X_n \xrightarrow{P} X$ if and only if every subsequence of $\{X_n\}$ has a further subsequence that converges to X almost surely.

3. If $X_n \xrightarrow{P} X$ and $g : \mathbb{R} \to \mathbb{R}$ is a continuous function, then $g(X_n) \xrightarrow{P} g(X)$.

4. If X and Y are two independent random variables such that X has a continuous cumulative distribution function, then $P(X = Y) = 0$.

INDIAN STATISTICAL INSTITUTE

Mid-Sem Examination, 1^{st} Semester, 2016-17

Statistical Methods III B.Stat 2^{nd} Year

Date: August 30, 2016 Time: 2 hours

This is an open notes examination. The paper carries 30 marks. Answer all questions.

1. Suppose X is an observation from a Geometric distribution with parameter p. Explain whether there exists an unbiased estimator of p^2 based on X. [5]

2. Suppose $(x_1, y_1), (x_2, y_2), \ldots, (x_{10}, y_{10})$ represents a set of bivariate observations on (X, Y) such that $x_1 = x_2 = x_3 \neq x_4 = x_5 = \ldots = x_{10}$. Obtain necessary and sufficient conditions such that the least squares regression line of Y on X is identical to the least absolute deviation line. [8]

3. Suppose $(0.75, 0.34), (0.35, 0.4), (1.58, 0.63), (0.42, -0.91)$ are independent observations from a bivariate normal distribution with parameters $(0, 0, 1, 1, \rho)$. Explain how you would use the Fisher's Scoring Method to obtain the m.l.e. of ρ. Show all computational steps clearly with one cycle of iteration. [10]

4. Suppose $X_1, X_2, \ldots, X_5$ is a random sample from a Poisson distribution with mean λ. Show that the variance of the sample variance is greater than 0.128λ. [7]

INDIAN STATISTICAL INSTITUTE

MID-TERM EXAMINATION (2016–17)

B. STAT. II YEAR, FIRST SEMESTER

ANALYSIS III

Date : 31.08.2016 Maximum Marks : 80 Time : $2\frac{1}{2}$ hours

The question paper carries 85 marks. Maximum you can score is 80. Precisely justify all your steps. Carefully state all the results you are using.

1. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Let $x_0 \in [a, b]$.

   (a) Define $\omega_f(x_0)$, the oscillation of f at x_0. [5]

   (b) Prove that f is continuous at x_0 if and only if $\omega_f(x_0) = 0$. [10]

2. Let $A \subseteq \mathbb{R}^n$ be nonempty. Define

$$d_A(x) = \inf\{\|x - a\| : a \in A\}, \qquad x \in \mathbb{R}^n.$$

   (a) Show that d_A is uniformly continuous. [5]

   (b) Show that $d_A(x) = 0$ if and only of $x \in \bar{A}$, the closure of A. [10]

   (c) Show that if $A \subseteq \mathbb{R}^n$ is connected, then either A is a singleton or uncountable. [10]

   (d) Show that $A \subseteq \mathbb{R}^n$ is compact if and only if every continuous $f : A \to \mathbb{R}$ is bounded above. [10]

3. Let $A \subseteq \mathbb{R}^n$ be nonempty. Define an equivalence relation on A as follows:

   $x \sim y$ if there is a connected subset $C \subseteq A$ such that $x, y \in C$.

   Let $x_0 \in A$. Let $C_{x_0} = \{y \in A : y \sim x_0\}$. Show that

   (a) C_{x_0} is connected. [5]

   (b) If $C_{x_0} \subseteq C \subseteq A$ and C is connected, then $C = C_{x_0}$. [3]

[PTO]

4. Let A be a $n \times n$ real symmetric matrix. Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as

$$f(x) = Ax \cdot x$$

Show that f is differentiable on $\mathbb{R}^n$ and find ∇f, the gradient of f. [7]

5. A helix is a curve in $\mathbb{R}^3$ parametrized by $r(t) = (a\cos t, a\sin t, bt)$, $t \in \mathbb{R}$. Show that the curvature of a helix is constant. [10]

6. Find all the local maxima, local minima and saddle points of the function [10]

$$f(x, y) = 2x^3 + 2y^3 - 9x^2 + 3y^2 - 12y$$

Indian Statistical Institute

B.Stat (Hons) II Year: Microeconomics

Mid-Semester Examination

Date: 01.09.2016 Maximum Marks: 40 Time: 3 hrs

1. Show that the Lexicographic preference is not continuous. What is the implication of this result? [4]
2. Derive the expression for the indirect money metric utility function for the utility function $U(x,y) = x^{1/2}y^{1/2}$ and interpret. [6]
3. Show that, in general, the consumer will prefer an income tax over a commodity tax where both the schemes give identical revenue for the government. [5]
4. A consumer consumes only two goods x and y. The price of good x in the local market is p and that in a distant market is q, where $q < p$. However, to go to the distant market, the consumer has to incur a fixed cost c. Suppose that the price of good y is unity in both markets. The consumer's income is M and $M > c$.

(i) Draw the budget set.

(ii) Let x^* be the optimal consumption of good x. If the consumer has the usual smooth downward sloping and convex indifference curves, then find out a relation involving q, p, c and x^*. [3+4]

5. A consumer consumes electricity E and other goods O. The price of other goods is unity. To consume electricity the consumer has to pay a rental charge R and a per unit price p. However, p increases with the quantity of electricity consumed according to the relation $p = \frac{1}{2}E$. The utility function of the consumer is $U = E + O$ and his income is $M > R$.

(i) Draw the budget line of the consumer.

(ii) If $R = 0$ and $M = 1$, find the optimum consumption bundle.

(iii)Find the maximum R that the electricity company can extract from the consumer. [3+3+3]

6. A consumer has the utility function $U(x,y) = 2x + 32y - 3y^2$ in an economy with two goods x and y. Suppose she has income 20.

(i) If the prices of x and y are each 1, what is her optimal consumption bundle?

(ii) Suppose the price of y increases to 4, all else remaining the same. Which consumption bundle does she choose now?

(iii)How much extra income must the consumer be given to compensate her for the increase in price of y ? [3+3+4]

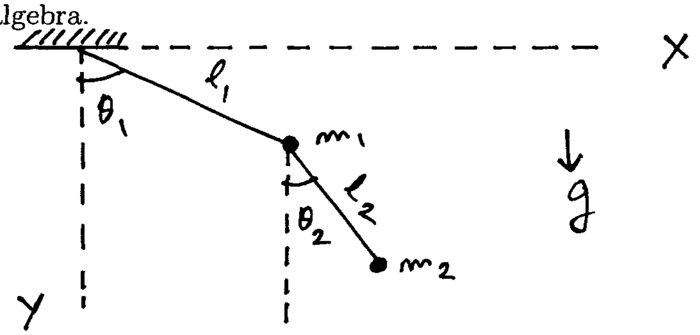
7. Class Test [9]

B.STAT. II YEAR, MID-SEMESTRAL EXAM, PHYSICS I
DATE: 01.09.2016
FULL MARKS = 40. ANSWER ANY FIVE QUESTIONS. EACH QUESTION CARRIES 8 MARKS.

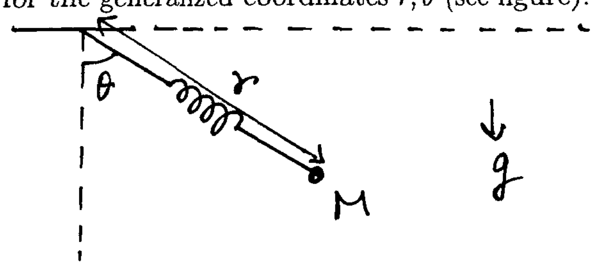
Time: 2 hrs.

1) (a) Derive the Lagrangian equation of motion from the principle of virtual work.
(b) Consider $L' = L + \frac{dF(q_i,t)}{dt}$, where $F(q_i, t)$ is an arbitrary function of the generalized coordinates q_i and t, and L' and L are two Lagrangians. Show that L' and L give rise to same equations of motion.
5+3=8

2) Consider a double pendulum as given in the figure below with masses m_1 and m_2 that are connected by a string of fixed length l_2. The mass m_1 is attached to a fixed overhead support by a string of fixed length l_1. The whole system is placed in constant downward gravitational field with g the acceleration due to gravity. The motion is restricted to the $X - Y$ plane.
(a) Write down the Cartesian to generalized coordinate (θ_1, θ_2 as defined in the figure) transformation equations.
(a) Set up Lagrangian of the system in terms of generalized coordinates θ_1, θ_2.
There is no need to simplify the algebra.
4+4=8



3) Construct the Lagrangian of a single mass M that is attached to a fixed overhead support by a spring of spring constant k and equilibrium length L and is placed in constant downward gravitational field with g the acceleration due to gravity. The motion is restricted to the $X - Y$ plane.
(a) Set up the Lagrangian in terms of r, θ.
(b) Derive the equations of motion for the generalized coordinates r, θ (see figure).
5+3=8



4) (a) Derive the Lagrangian equation of motion from Least Action Principle.
(b) Consider a total time derivative term $\frac{dG}{dt}$ added to the Lagrangian. What are the boundary conditions that need be imposed on G so that G will not contribute in the Lagrangian equations of motion.
5+3=8

5) Consider two bodies having masses m and M with the force between them being gravitational in nature.

(a) Write down the Lagrangian in terms of generalized coordinates, (coordinate of the center of mass and the relative coordinate of one mass, M, relative to the other mass, m,).

(b) Define reduced mass and its significance.

(c) Find the cyclic coordinate and corresponding conserved quantity.

(d) Write down the equations of motion.

3+1+2+2=8

6) Consider the following Lagrangian,

$$L = \frac{A}{2}\dot{q}_1^2 + \frac{B}{2}q_1\dot{q}_2^2 + C\dot{q}_1 q_2^2 + Dq_1 q_2 \dot{q}_3 + E\dot{q}_1,$$

where q_1, q_2, q_3 are coordinates and $A, B, C.D, E$ are constants.

a) Find the cyclic coordinate and corresponding conserved quantity.

b) Write down the equations of motion.

c) Find the conjugate momenta.

d) Construct the Hamiltonian in terms of the coordinates and the conjugate momenta.

1+3+2+2=8

7) Consider a closed electrical circuit with an inductance L and a capacitance C connected in series. Let the current in the circuit be I. Treat the charge Q moving in the circuit as a generalized coordinate.

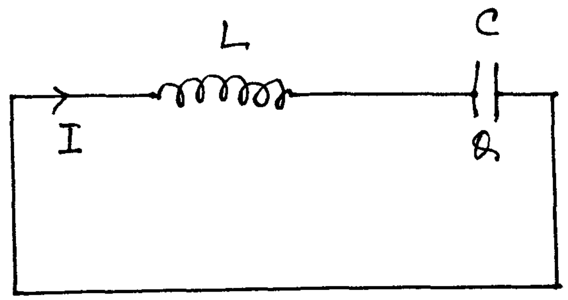
a) Set up the Lagrangian for the system.

b) Derive the equation of motion for Q.

c) How will the solutions look like?

Hint: Current (I) in the circuit is the rate of flow of charge (Q). Use expressions of energy for L with I current passing through and energy for C with Q stored charge.

4+1+2+2=8



Indian Statistical Institute

Mid-Semestral Examination: 2016-17

Course Name: B. Stat II, **Subject name:** Molecular Biology

Date: **01/09/2016** ***Maximum Marks:*** **40;** ***Duration: 2.0 hrs***

All questions carry equal marks, answer any four

1) a). How many ATPs would be generated when one molecule of acetyl-CoA is completely oxidized in TCA cycle? Mention the specific reactions/steps which are responsible for generation of NADH and $FADH_2$ in TCA cycle. [7]

b. During physical activity we may feel muscle pain: what could be the reason in terms of glucose metabolism? [3]

2). a). Explain why a protein "X" might become positively charge at pH:7.0 whereas another protein "Y" might become negatively charge at pH:7.0 ? [4]

b). How a mixture of two proteins (mol.wt. 8,000 kD and 15,000 kD, respectively) could be separated from each other without denaturation? [6]

3). How stearic and oleic acids, with the following chemical formula $CH_3(CH_2)_{16}COOH]$ and $[CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH]$ respectively, differ in oxidation steps to generate ATP ? [10]

4). How one will determine whether a protein molecule of mol. wt. 42,000 kD consists of one or two polypeptide chains? If it consists of two chains then how to determine whether chains are identical or different? [10]

5(a) Distinguish between DNA and RNA with respect to their chemical and physical properties and functions. [5]

(b) Why single "origin of replication" is sufficient for replication of a bacterial DNA but multiple "origin of replication" is needed for replication of human DNA? [5]

INDIAN STATISTICAL INSTITUTE

First Mid-Semestral Examination: 2016-17

Subject Name : **Elements of Algebraic Structure** Date: 05/09/16
Course Name : B.Stat. II yr. Maximum Score: 35 Duration: 2 Hours 30 min

Note: Attempt all questions. Marks are given in brackets. Total score is 40. **State results** clearly which you want to use. Use **separate page** for each question.

Problem 1. Let D be a commutative ring **possibly without identity** which satisfies no zero divisor property and there is an Euclidean nonnegative function d such that division property holds (for all $a, b \neq 0 \in D$ there exists q such that either $a = bq$ or $d(a - bq) < d(b)$). Show that D is an Euclidean domain. Identify the set of all unit elements in terms of d function. [3+3=6]

Problem 2. Let $R = \mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{3}i, a, b \in \mathbb{Z}\}$ be a ring under usual complex addition and multiplication. Show that 2 is irreducible but not prime. [4+4 =8]

Problem 3. Let G be a group and $x, y \in G$ such that the smallest group generated by x and y is G. Suppose we know that $o(x) = n$, $o(y) = o(xy) = 2$ where $n \geq 2$ is an integer. Find $|G|$ with an explanation. Moreover, for every i, compute $o(x^i y)$. [4 + 3 =7]

Problem 4. Let G be a group and $S = \{aba^{-1}b^{-1} : a, b \in G\}$. Let C be the smallest normal subgroup of G containing the set S. Prove that G/C is abelian. Moreover, show that any subgroup of G containing C is normal in G. [3+4=7]

Problem 5. Show that the multiplicative group of non-zero elements of a finite field is cyclic. [6]

Problem 6. Construct an example of domain which is not a factorization domain. [6]

INDIAN STATISTICAL INSTITUTE

Supplementary Mid-Semester Examination: 2016-17 (First Semester)
Bachelor of Statistics (B. Stat.) II Year
Probability Theory III

Teacher: Parthanil Roy

Date: 04/11/2016 **Maximum Marks: 40** **Duration: 10:30 am - 1:00 pm**

Note:

- **Please write your roll number on top of your answer paper.**
- **Show all your works and write explanations when needed. Maximum you can score is 40 marks.**
- **You may use any fact proved in the class but do not forget to quote the appropriate result. However, if you use a result that was given as an exercise in the class, then you have to provide a *complete proof* of that result based on whatever has been covered in the class. This, of course, does not apply to the results for which proofs will be given in future (i.e., either later in this semester or in a future course).**
- ***In this examination, you are not* allowed to use class notes, books, homework solutions, list of theorems, formulas etc. You are also requested to switch off your cell phone and keep them inside your bag. Failing to follow the examination guidelines, copying in the examination, rowdyism or some other breach of discipline or unlawful/unethical behavior, etc. are regarded as unsatisfactory conduct. Any student caught cheating or violating examination rules will get a zero in this examination.**

1. State whether each of the following statements is true or false. If it is true, give a detailed proof. On the other hand, if it is false, produce a counter-example with full justification. Each question is worth 10 marks. If you both prove and disprove a statement, you will get a zero in that problem.

   (a) If F_n is a sequence of distribution functions that converges in distribution to Φ (the standard normal distribution function), then $F_n(x_n) \to \Phi(x)$ whenever $x_n \to x$.

   (b) $X_n \xrightarrow{P} X$ if and only if $E\big(\min\{|X_n - X|, 1000\}\big) \to 0$.

2. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$ distribution. Find, with full justification, the probability density function of
$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (X_i - \mu)^2},$$
where $\bar{X}$ is the sample mean based on the random sample. [10]

3. Suppose two types of passengers (domestic and international) arrive independently in an airport following Poisson processes with rates 30 and 20 per hour, respectively. If it is given that exactly 75 passengers arrived in an hour in that airport, find the conditional probability density function of the length of the busy period for the security personnel during that hour. (Note that *length of the busy period* is defined as the time difference between the last and first arrivals.) [10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2016-17 (First Semester)
Bachelor of Statistics (B. Stat.) II Year
Probability Theory III

Teacher: Parthanil Roy

Date: 07/11/2016 **Maximum Marks: 60** **Duration: 10:30 am - 2:00 pm**

Note:

- **Please write your roll number on top of your answer paper.**
- **Show all your works and write explanations when needed. Maximum you can score is 60 marks.**
- **You may use any fact proved in the class but do not forget to quote the appropriate result. However, if you use a result that was given as an exercise in the class, then you have to provide a *complete proof* of that result based on whatever has been covered in the class. This, of course, does not apply to the results whose proofs will be given in a future course.**
- **This is an open note examination. You are allowed to use your own hand written notes (such as class notes, your homework solutions, list of theorems, formulas etc). Please note that no printed or photocopied material is allowed. In particular you are not allowed to use books, photocopied class notes etc. You are also requested to switch off your cell phone and keep them inside your bag. Failing to follow the examination guidelines, copying in the examination, rowdyism or some other breach of discipline or unlawful/unethical behavior, etc. are regarded as unsatisfactory conduct. Any student caught cheating or violating examination rules will get a zero in this examination.**

1. Buses arrive at a certain bus stop according to a Poisson process with rate $\lambda > 0$. If you take the bus from that stop, then it takes a deterministic time r, measured from the time at which you enter the bus, to arrive home. If you walk from the bus stop, then it takes a deterministic time w to arrive home. Suppose that your policy when arriving at the bus stop is to wait up to a deterministic time s, and if a bus has not yet arrived by that time then you walk home. Compute the expected time from when you arrive at the bus stop until you reach home. [8]

2. Suppose that $X_1, X_2, \dots$ are independent and identically distributed random variables following standard normal distribution. Compute the almost sure limit of the following sequence of random variables:-

   (a) $\frac{1}{n}\sum_{i=1}^{n} X_i^4$. [4]

   (b) $\min\{X_1^2, X_2^2, \dots, X_n^2\}$. [6]

3. Suppose U and V are two independent and identically distributed random variables following a uniform distribution on $(-1, 1)$. Define $X = \min(U, V)$ and $Y := \max(U, V)$. Compute the conditional distribution of U given Y and the conditional distribution of X given Y. [7 + 7 = 14]

[P. T. O]

4. Suppose that $X_1, X_2, \ldots$ are independent and identically distributed random variables having characteristic function $\phi(t) = \exp\{-|t|^{1.9}\}$, $t \in \mathbb{R}$. Find (with justification) whether the following sequence of random variables are tight:-

(a) $\displaystyle \frac{1}{n^{1/2}} \sum_{i=1}^{n} X_i$. [5]

(b) $\displaystyle \frac{1}{n^{5/9}} \sum_{i=1}^{n} X_i$. [5]

5. Fix $k \in \mathbb{N}$. For a sequence of random vectors $\tilde{\mathbf{X}}_n = \left(X_n^{(1)}, X_n^{(2)}, \ldots, X_n^{(k)}\right)$, $n \geq 1$, and another random vector $\tilde{\mathbf{X}} = \left(X^{(1)}, X^{(2)}, \ldots, X^{(k)}\right)$, we write $\tilde{\mathbf{X}}_n \xrightarrow{P} \tilde{\mathbf{X}}$ (convergence in probability) if for all $\epsilon > 0$,

$$P\left(\|\tilde{\mathbf{X}}_n - \tilde{\mathbf{X}}\| > \epsilon\right) \to 0.$$

as $n \to \infty$. Here $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^k$. State whether each of the following statements is true or false. If it is true, give a detailed proof. On the other hand, if it is false, produce a counter-example with full justification. If you both prove and disprove a statement, you will get a zero in that problem.

(a) $\tilde{\mathbf{X}}_n \xrightarrow{P} \tilde{\mathbf{X}}$ if for all $i = 1, 2, \ldots, k$, $X_n^{(i)} \xrightarrow{P} X^{(i)}$. [6]

(b) $\tilde{\mathbf{X}}_n \xrightarrow{P} \tilde{\mathbf{X}}$ only if for all $i = 1, 2, \ldots, k$, $X_n^{(i)} \xrightarrow{P} X^{(i)}$. [6]

(c) If $\tilde{\mathbf{X}}_n \xrightarrow{P} \tilde{\mathbf{X}}$, then $\tilde{\mathbf{X}}_n \xrightarrow{d} \tilde{\mathbf{X}}$. [6]

INDIAN STATISTICAL INSTITUTE

Final Examination, 1^{st} Semester, 2016-17

Statistical Methods III, B.Stat 2^{nd} Year

Date: November 9, 2016 Time: 3 hours

The paper carries 55 marks. Answer all questions.

1. Consider the usual regression model:

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

where, x_is are fixed and e_is are i.i.d. $N(0, \sigma^2)$.

(a) Compute the Fisher's information matrix for the vector $(\alpha, \beta, \sigma^2)$ and hence show that the least squares estimator of β is the UMVUE for β.

(b) If y_n is missing, explain how you would use the EM algorithm to obtain the maximum likelihood estimates of α, β and σ^2. Show all computational steps clearly. [10 + 10]

2. Suppose T is an unbiased estimator of a parameter θ. Consider estimators of the form cT to estimate θ where c is a real number. Show that the estimator in the above class that minimizes the mean squared error satisfies the condition $0 < c < 1$. [5]

3. Suppose $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ represents the coordinates of a set of independent bivariate observations that are distributed uniformly in the region between two concentric circles with centre (0,0) and radii of the inner and the outer circles being r and r_0, respectively, where r_0 is known.

(a) Examine whether the maximum likelihood estimator of r is consistent for r.

(b) Consider a test for $H_0 : r \geq r_1$ vs $H_1 : r < r_1$ that rejects H_0 iff the maximum likelihood estimator of r is sufficiently small. If the size of the above test is α, draw a rough sketch of the power function of the test. Determine the minimum sample size required such that the power of the test is at least β at $r = r_2$ $(< r_1)$.

[4 + 6]

4. Suppose fasting glucose levels are distributed as normal for both obese ($BMI \geq 25$) and non-obese ($BMI < 25$) individuals, but possibly with different parameters for the two groups. In a study, 12 obese people and 8 non-obese people were randomly sampled from a population and their fasting glucose levels measured. The maximum likelihood estimates of the mean and the standard deviation of fasting glucose levels in the first group were found to be 140.8 pmol/L and 13.2 pmol/L, respectively; while those in the second group were found to be 101.7 pmol/L and 12.5 pmol/L, respectively. Construct a 95% equal tail confidence interval for the difference between the mean fasting glucose levels for obese and non-obese individuals in the population. State your assumptions clearly with suitable justifications. Based on the confidence interval obtained, can you infer at level 0.01 whether the mean fasting glucose levels in the two groups are same or not? [10]

5. Suppose the number of flight delays on any specific day of a week is distributed as Poisson and varies independently across different days. A reporter, who suspects that the mean number of flight delays during any of the weekdays (Monday to Friday) is less than that during any of the days in the weekend (Saturday or Sunday), observes that there were a total of 8 flight delays during the weekdays and a total of 5 delays during the weekend of a particular week. Do the data validate the reporter's suspicion? Show all computations clearly. [10]

INDIAN STATISTICAL INSTITUTE

Final Examination, 1^{st} Semester, 2016-17

Statistical Methods III, B.Stat 2^{nd} Year

Date: November 9, 2016 Time: 3 hours

The paper carries 55 marks. Answer all questions.

1. Consider the usual regression model:

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

where, x_is are fixed and e_is are i.i.d. $N(0, \sigma^2)$.

(a) Compute the Fisher's information matrix for the vector $(\alpha, \beta, \sigma^2)$ and hence show that the least squares estimator of β is the UMVUE for β.

(b) If y_n is missing, explain how you would use the EM algorithm to obtain the maximum likelihood estimates of α, β and σ^2. Show all computational steps clearly. [10 + 10]

2. Suppose T is an unbiased estimator of a parameter θ. Consider estimators of the form cT to estimate θ where c is a real number. Show that the estimator in the above class that minimizes the mean squared error satisfies the condition $0 < c < 1$. [5]

3. Suppose $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ represents the coordinates of a set of independent bivariate observations that are distributed uniformly in the region between two concentric circles with centre (0,0) and radii of the inner and the outer circles being r and r_0, respectively, where r_0 is known.

(a) Examine whether the maximum likelihood estimator of r is consistent for r.

(b) Consider a test for $H_0 : r \geq r_1$ vs $H_1 : r < r_1$ that rejects H_0 iff the maximum likelihood estimator of r is sufficiently small. If the size of the above test is α, draw a rough sketch of the power function of the test. Determine the minimum sample size required such that the power of the test is at least β at $r = r_2$ $(< r_1)$.

[4 + 6]

4. Suppose fasting glucose levels are distributed as normal for both obese ($BMI \geq 25$) and non-obese ($BMI < 25$) individuals, but possibly with different parameters for the two groups. In a study, 12 obese people and 8 non-obese people were randomly sampled from a population and their fasting glucose levels measured. The maximum likelihood estimates of the mean and the standard deviation of fasting glucose levels in the first group were found to be 140.8 pmol/L and 13.2 pmol/L, respectively; while those in the second group were found to be 101.7 pmol/L and 12.5 pmol/L, respectively. Construct a 95% equal tail confidence interval for the difference between the mean fasting glucose levels for obese and non-obese individuals in the population. State your assumptions clearly with suitable justifications. Based on the confidence interval obtained, can you infer at level 0.01 whether the mean fasting glucose levels in the two groups are same or not? [10]

5. Suppose the number of flight delays on any specific day of a week is distributed as Poisson and varies independently across different days. A reporter, who suspects that the mean number of flight delays during any of the weekdays (Monday to Friday) is less than that during any of the days in the weekend (Saturday or Sunday), observes that there were a total of 8 flight delays during the weekdays and a total of 5 delays during the weekend of a particular week. Do the data validate the reporter's suspicion? Show all computations clearly. [10]

Indian Statistical Institute

11.11.2016

Semester Examination (B. Stat-II, Molecular Biology, Year-2016)

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

1. (a) Describe the features of genetic code. **[5]**

(b) In a couple, husband is albino and wife has an albino sibling. What is the risk of their baby to be an albino? (Albinism is an autosomal recessive disease and parents of the wife are normal). **[5]**

2. (a) Identify the alleles, genotypes and traits in individuals having A, B and AB blood groups? **[5]**

(b) Explain why two brothers, from the same parents, are not genetically identical. What is the chance that they would be genetically identical (monozygotic twins to be excluded)? **[5]**

3. If four babies are born on a given day: what are the chances that (a) number of boys and girls will be equal; (b) all four will be girls and (c) at least one baby will be girl? What combination of boys and girls is most likely to occur? **[2.5x4]**

4. A man with X-linked color blindness marries a woman with no history of color blindness in her family. The daughter of this couple marries a normal man and their daughter also marries a normal man. What is the chance that the last couple will have a child with color blindness? If the last couple already had a child with color blindness, what is the chance that their next child will be color blind? Answer the questions with a pedigree. **[10]**

5. (a) Why the number of nucleotides of the same gene is different in bacteria and human? [5]

(b) Distinguish between transcription and translation. [5]

6. (a) If father and son both are color blind, is it likely that the son inherited the trait from his father? Show possible genotypes with a pedigree diagram. **[5]**

(b) How many different DNA sequences of 18 bases are possible with 6 codons: ATG, TAA, AGG, GGG, CGT and TAC; provided there will not be five consecutive "G"s within the sequences and each codon is used at least once to construct the DNA. **[5]**

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2016-2017

B.Stat. II Year

Physics I

Date : 11-11-2016 Maximum Marks: 60 Duration : 3 hours

1. Answer *any two* questions :

i. (a) Derive Hamilton's equations of motion from the least action principle by varying the action.
(Hint: use the relation $L = p_i\dot{q}_i - H$ *where* q_i, p_i *are canonically conjugate coordinates and momenta respectively and* $L(q_i, \dot{q}_i), H(q_i, p_i)$ *are Lagrangian and Hamiltonian respectively.)*

(b) For the Hamiltonian of a one-dimensional harmonic oscillator :

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2,$$

(m, k are constants) derive Hamilton's equations of motion.

[2 + 3 = 5]

ii. Consider a particle of mass m and charge e in an electromagnetic field $\{\phi(x),\ A_i(x)\}$ with Lagrangian

$$L = \frac{m\dot{x}_i\dot{x}_i}{2} - e\phi + eA_i\dot{x}_i.$$

(a) Write down the Hamiltonian of the system.

(b) Find the equations of motion using either Lagrangian or Hamiltonian formalism.

[1 + 4 = 5]

iii. (a) Consider a free particle in three dimensions. Write down its Lagrangian and Hamiltonian.

(b) Using the definition of angular momentum vector $\vec{l} = \vec{r} \times \vec{p}$ where $\vec{r}$, $\vec{p}$ are Cartesian coordinates and momenta, calculate the Poisson bracket $\{l_x,\ l_y\}$ and identify the result.

[2 + 3 = 5]

2. Answer the following short questions :

i. Which of the following vector fields could describe an electric field and why :

(a) $\vec{E}(\vec{r}) = x_1\hat{\varepsilon}_1 - x_2\hat{\varepsilon}_2$

(b) $\vec{E}(\vec{r}) = x_2\hat{\varepsilon}_1 + x_1\hat{\varepsilon}_2$

(c) $\vec{E}(\vec{r}) = x_1\hat{\varepsilon}_1 - x_2\hat{\varepsilon}_1$

Hint : Note that $\vec{E} = -\vec{\nabla}V$. [$3 \times 3 = 9$]

ii. State whether the following statements are *True* or *False* :

(a) A metal sphere is placed in a non-uniform electric field. Its surface may not be an equipotential one.

(b) The electric field inside a metal cylinder placed in an non-uniform external electric field always vanishes.

(c) A polar molecule placed in an external electric field aligns itself such that its dipole moment points opposite to the external field.

(d) A neutral rectangular slab, made of dielectric material, is placed in an external electric field. The surface charge density of such an object must always be zero.

(e) The potential due to the charge configuration, as shown in Fig.[1], falls off as $\frac{1}{r^3}$ far away from the origin (the origin is assumed to be at the center of the square).

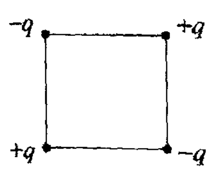


Figure 1

(f) A charged particle moving in a uniform magnetic field must maintain a constant kinetic energy (the only force acting on it is the magnetic force).

[$1.5 \times 6 = 9$]

3. Answer the following questions :

i. Show the following (use $\epsilon_{ijk}, \delta_{ij}$ notation):

(a) $(\vec{A} \times \vec{B}).(\vec{C} \times \vec{D}) = (\vec{A}.\vec{C})(\vec{B}.\vec{D}) - (\vec{A}.\vec{D})(\vec{B}.\vec{C})$

(b) $\vec{\nabla}.(\vec{A} \times \vec{B}) = \vec{B}.\vec{\nabla} \times \vec{A} - \vec{A}.\vec{\nabla} \times \vec{B}$

(c) $\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla}.\vec{A}) - \nabla^2\vec{A}$

Hint : $\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$.

[$3 \times 3 = 9$]

ii. Show that $\int_{\mathcal{V}} \vec{B}.(\vec{\nabla} \times \vec{A})\, d\tau = \oint_S (\vec{A} \times \vec{B}).\vec{da} + \int_{\mathcal{V}} \vec{A}.(\vec{\nabla} \times \vec{B})\, d\tau$. Symbols have their usual significance.

[3]

iii. Consider a vector field $v = x_1x_2\hat{e}_1 + 2x_2x_3\hat{e}_2 + 3x_1x_3\hat{e}_3$. Estimate $\int_S(\vec{\nabla} \times \vec{v}).\vec{da}$ on the shaded surface as shown in Fig.[2].

[5]

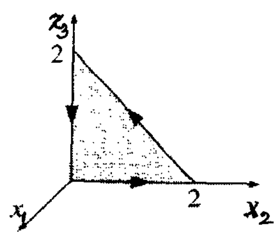


Figure 2

Answer *any three* **from question nos. 4, 5, 6 and 7.**

4. Consider a uniformly charged sphere of radius R and charge density ρ ($\rho > 0$).

i. *Use Gauss's law* to find the electric field inside the sphere.

ii. Suppose another sphere with the same radius, but with an uniform charge density $-\rho$, is placed such that these two spheres partially overlap. Let the distance between the two centers be $d = |\vec{\mathbf{d}}|$ as shown in Fig.[3]. Show that the electric field in the region of overlap is constant and find its magnitude and direction.

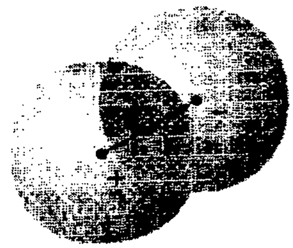


Figure 3

[2+3 =5]

5. A spherical conductor of radius R consists of two hollow spherical cavities of radii a and b as shown in Fig. 4. At the center of these cavities two point charges q_a and q_b have been placed respectively.

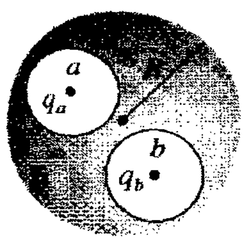


Figure 4

i. Find out the surface charges σ_a, σ_b and σ_R.

ii. What is the field outside the conductor at a distance r?

iii. What is the field within each cavity?

iv. What is the force on each of the charges q_a and q_b?

v. If a third charge q_c were placed near the conducting sphere, which of the above answers will change?

[$1 \times 5 = 5$]

6. Consider three point charges placed as follows [Fig. 5]: a charge q is placed at (0,0,a), and two point charges, $-q$ each, are placed at (0,a,0) and (0,$-a$,0) respectively. Cartesian co-ordinate system has been used to specify these co-ordinates.

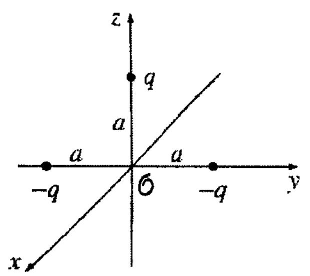


Figure 5

i. Find out an approximate formula for the potential valid at a point P far ($|\vec{r}_P| \gg a$) from the origin $\mathcal{O}$.

ii. Find the approximate electric field at the point P.

Use spherical co-ordinates to express your answer and only keep the two lowest orders in the multipole expansion i.e. the monopole and the dipole contributions. Note that :

$$\vec{\nabla} = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}.$$

[$2 + 3 = 5$]

7. A point charge q of mass m is released from rest at a distance d above an infinite grounded conducting plane. How long will it take for the charge to hit the plane?

[5]

Indian Statistical Institute

B.Stat. II Year: Microeconomics

Semester Examination

Date: 11.11.2016 | Maximum You Can Score: 60 | Time: 3 hrs

1. Consider a monopolist operating an amusement park with a constant marginal cost equal to 6. He faces two types of visitors, with different willingness to pay, given by the inverse demand functions $P_1 = 24 - Q_1$ and $P_2 = 12 - 0.5Q_2$ if they enter the park but the monopolist must charge one single price P for them. However, he sets an entry fee f. Find out the optimal f and P. Find the average prices paid by each type of visitors. [4+4+2]

2. Consider a pure exchange economy with two consumers 1 and 2 with two goods X and Y. The initial endowments of the consumers are respectively $w_1 = (1,0)$ and $w_2 = (0,1)$. They have identical utility function $U(x, y) = xy$.
   (a) Describe the economy in an Edgeworth box diagram.
   (b) Write down the equations for the Pareto efficient allocations. Identify them in the Edgeworth box. Is the initial endowment point Pareto efficient? Explain.
   (c) Calculate the competitive equilibrium indicating the final consumptions and the equilibrium prices. [3+3+4]

3. Define the Arrow-Pratt measure of absolute risk aversion and show that it is related to the premium that a risk averse individual is willing to pay to avoid a fair gamble. [2+5]

4. The return on a bond fund has expected value 10% and standard deviation 15%, while the return on a stock fund has expected value 20% and standard deviation 25%. The correlation between the returns is 0.50. Suppose that an investor's utility functional is of the form $U = \mu - \frac{1}{100}\sigma^2$. Determine the investor's optimal allocation between stocks and bonds assuming short selling without margin is possible (that is, the allocated proportion is any real number). [10]

5. Two persons have equal access to a common resource of size 10,000 units. They live for two periods. In period 1, each can demand to withdraw resources c_1 and c_2 respectively which are non-negative. If $c_1 + c_2 > 10{,}000$, each gets 5000. If $c_1 + c_2 < 10{,}000$, they can share the remaining resource in period 2 equally. Both have identical utility function $U(x) = \log x$ in each period and future utility is not discounted. Find out the Nash equilibrium of the resources withdrawn. What happens if they decide to collude to optimise their aggregate utility? Explain the difference between the solutions obtained in the two situations in economic terms. [3+3+2]

6. Consider the pay-off matrix of a two person zero sum game. Find the equilibrium, if any.

$$\begin{pmatrix} 1 & 2 & -2 & 2 \\ 3 & 1 & 2 & 3 \\ -1 & 3 & 2 & 1 \\ -2 & 2 & 0 & -3 \end{pmatrix}$$

[5]

7. Class Test and Project Presentation [20]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2016-17

Subject Name : **Elements of Algebraic Structure** Date: 15//11/16

Course Name : B.Stat. II yr. Maximum Score: 50 Duration: 3 Hours

Note: Attempt all questions. Marks are given in brackets. Total score is 55. Clearly **state results** which you want to use. Use **separate page** for each question.

Problem 1. Let $F \leq E_1, E_2 \leq E$ be all field extensions of **finite** dimensions. Show that $E_1E_2 := \{\sum_{i=1}^{n} a_ib_i : a_i \in E_1, b_i \in E_2, n \geq 1\}$ is a subfield of E. [6]

Problem 2. (i) Show that $x^3 - 3$ is irreducible over $\mathbb{Q}$ (the field of rational numbers). (ii) Is $3^{\frac{1}{2}}$ an element of $\mathbb{Q}(3^{\frac{1}{3}})$, $\mathbb{Q}(3^{\frac{1}{2}} + 3^{\frac{1}{3}})$ or both? Justify your answer. [3+(3+4)=10]

Problem 3. Find all non-isomorphic groups of size 10. [8]

Problem 4. Let ρ be an automorphism on G (i.e. it is an isomorphism from G to itself). Prove that $C(\rho(a)) = \{\rho(x) : x \in C(a)\}$ where $C(a)$ is the centralizer of the element a in the group G. [5]

Problem 5. Let G be a group in which for some integer $n > 1$, $(ab)^n = a^nb^n$ for all $a, b \in G$. Show that (a) $G^n = \{x^n : x \in G\}$ is a normal subgroup of G and (b) $G^{n-1} = \{x^{n-1} : x \in G\}$ is a normal subgroup of G. [5+5=10]

Problem 6. Let R be the ring of all 2×2 matrices of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ where a and b are real numbers. Prove that R is isomorphic (as a ring) to $\mathbb{C}$, the field of complex numbers. [6]

Problem 7. Let D be a UFD such that for every pair of elements $a, b \in D$, there exists $x, y \in D$ such that $ax + by = d$ where d is the GCD of a and b. Show that D is PID if and only if every ideal of D is finitely generated. [8+2 = 10]

INDIAN STATISTICAL IN[redacted]

FIRST SEMESTRAL EXAMINATIO[redacted]7)

B. STAT. II YEAR

ANALYSIS III

Date : 18.11.2016 Maximum Marks : 100 [redacted] Time : $3\frac{1}{2}$ hours

The question paper carries 115 marks. Maximu[redacted]u can score is 100. Precisely justify all your steps. Carefully [redacted] all the results you are using.

1. Let $A \subseteq \mathbb{R}^n$ be a non-empty open set. Find all possible values of the dimension of the linear span of A. [10]

2. Decide whether for any of the following vector fields on $\mathbb{R}^2$, there exists a scalar field φ on $\mathbb{R}^2$ such that $f = \nabla\varphi$. If it is, find such a φ : [10]

   (a) $f(x,y) = (3x^2y, x^3y)$

   (b) $f(x,y) = (3x^2y, x^3)$

3. Find the extreme values of $f(x,y) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \leq 4$. [10]

4. Solve the following differential equations by either showing that it is exact, or, by finding a suitable integrating factor: [10]

   (a) $(x + 2y)dx + (2x + y)dy = 0$

   (b) $ydx + 2xdy = 0$.

5. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable. Show that f cannot be 1-1. [Hint : The function $g(x,y) = (f(x,y), y)$ has nonzero Jacobian at some point.] [15]

PTO

6. Let C be a smooth simple closed curve in the first quadrant $\{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$. Show that [7]

$$\oint_C \ln x \sin y \, dy = \oint_C \frac{\cos y}{x} dx.$$

7. Use Green's Theorem to evaluate

$$\oint_C (y^4 - 2y)dx - (6x - 4xy^3)dy$$

where C is the boundary of the rectangle with vertices $(0, 0)$, $(6, 0)$, $(0, 4)$, $(6, 4)$ traversed counterclockwise. [8]

8. Use a double integral to determine the volume of the region formed by the intersection of the two cylinders $x^2 + y^2 = 4$ and $x^2 + z^2 = 4$. [20]

9. Determine the surface area of the portion of the plane $2x + 3y + 6z = 9$ that is inside the cylinder $x^2 + y^2 = 7$. [10]

10. Use "spherical coordinates" to evaluate

$$\iiint_S \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right) dxdydz,$$

where S is the solid bounded by the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. [15]

INDIAN STATISTICAL INSTITUTE

Backpaper Examination: 2016-17 (First Semester)
Bachelor of Statistics (B. Stat.) II Year
Probability Theory III

Teacher: Parthanil Roy

Date: 26|12|16 **Full Marks: 100** **Duration: 3 hours**

Note:

- **Please write your roll number on top of your answer paper.**
- **Show all your works and write explanations when needed. Maximum you can score is 45 marks.**
- **You may use any fact proved in the class but do not forget to quote the appropriate result. However, if you use a result that was given as an exercise in the class, then you have to provide a *complete proof* of that result based on whatever has been covered in the class. This, of course, *does not* apply to the results whose proofs will be given in a future course.**
- **This is an open note examination. You are allowed to use your own hand written notes (such as class notes, your homework solutions, list of theorems, formulas etc). Please note that no printed or photocopied material is allowed. In particular you are not allowed to use books, photocopied class notes etc. You are also requested to switch off your cell phone and keep them inside your bag. Failing to follow the examination guidelines, copying in the examination, rowdyism or some other breach of discipline or unlawful/unethical behavior, etc. are regarded as unsatisfactory conduct. Any student caught cheating or violating examination rules will get a zero in this examination.**

1. Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables with characteristic function $\varphi(t) = \exp\left(-|t|^{0.87}\right)$. Write down the distribution function of $S_n = \sum_{i=1}^n X_i$ in terms of the distribution function of X_1. [6]

2. Suppose $\{X_n\}_{n\geq 1}$ is a sequence of random variables converging to X in probability and $E|X_n| \leq 32$ for all $n \geq 1$. State (with proper justification) whether the following statements are true or false. If it is true, you have to give a complete proof of the statement, and if it is false, you have to disprove it with full justification.

   (a) X is integrable and $E|X| \leq 32$. [6]

   (b) $E(X_n) \to E(X)$ as $n \to \infty$. [6]

3. Let $\{X_n\}$ be an i.i.d. sequence with finite mean μ and finite variance $\sigma^2 > 0$. Let $\bar{X}_n = \frac{1}{n}\sum_1^n X_i$ and $S_n^2 = \frac{1}{n-1}\sum_1^n (X_i - \bar{X}_n)^2$ be the sample mean and sample variance, respectively. Show that $\sqrt{n}(\bar{X}_n - \mu)/S_n$ converges weakly and find the limiting distribution. [12]

[P. T. O]

4. Suppose $\mathbf{X} \sim N(\mathbf{0}, I_k)$, A is a symmetric idempotent $k \times k$ matrix, and $\mathbf{b} \in \mathbb{R}^k$ is a nonzero column vector such that $A\mathbf{b} = \mathbf{0}$. State (with proper justification) whether the following statement is true or false (if it is true, you have to give a complete proof of the statement, and if it is false, you have to disprove it with full justification: $\mathbf{X}^T A\mathbf{X}$ and $\mathbf{b}^T\mathbf{X}$ are independent random variables. [15]

5. Suppose that X is a random variable with finite k^{th} moment and $\varphi(t)$ is its characteristic function. Show that $\varphi(t)$ is k times differentiable and express the k^{th} moment of X in terms of the k^{th} derivative of $\varphi(t)$. [15]

6. Suppose U and V are two independent and identically distributed random variables following an exponential distribution with parameter 1. Compute the conditional distribution of U given $Z := \max(U, V)$. [15]

7. (a) Let ϕ denote the probability density function of standard normal distribution and Φ denote its cumulative distribution function. Prove that $\lim_{x \to \infty} \frac{1 - \Phi(x)}{\phi(x)/x} = 1$. [5]

   (b) If $\{X_n\}$ is a sequence of independent and identically distributed standard normal random variables, show that $P\left[\limsup_{n \to \infty} \frac{|X_n|}{\sqrt{\log n}} = \sqrt{2}\right] = 1$. [10]

8. Suppose two types of passengers arrive in an auto rickshaw stand with unlimited supply of auto rickshaws. The first type of passengers will wait patiently till *four passengers* (= capacity of an auto rickshaw) arrive and the second type will simply reserve the auto rickshaw and go away immediately. Assume that these two types of passengers arrive independently according to Poisson processes with rates 10 and 5 per hour, respectively. Priorities are given to second type of passengers even if a few passengers of the first type are waiting in the queue. However, while each of the first type of passengers pay Rs. 10 for the trip, the second type of passengers are charged Rs. 60. Assume also that no time is lost in passengers getting into the auto rickshaw, the driver takes the money from the passenger(s) and departs immediately without wasting any time. Given that exactly 6 passengers arrive during 9 : 00 am - 9 : 30 am, compute the expected total earning of auto rickshaw drivers in that time-span. [10]

INDIAN STATISTICAL INSTITUTE

FIRST SEMESTER BACKPAPER EXAMINATION (2016–17)

B. STAT. II YEAR

ANALYSIS III

Date : 26.12.2016 Maximum Marks : 100 Time : 3 hours

Precisely justify all your steps. Carefully state all the results you are using.

1. Show that a nonempty connected subset of $\mathbb{R}^n$ is either a singleton or uncountable. [10]

2. Find and classify the critical points (if any) of the function $f(x,y) = y^2 - x^3$. [10]

3. Solve the following differential equations by either showing that it is exact, or, by finding a suitable integrating factor: [10]

   (a) $(x+2y)dx + (2x+y)dy = 0$

   (b) $ydx + 2xdy = 0$.

4. Let S be an open connected subset of $\mathbb{R}^2$. Let $f, g : S \to \mathbb{R}$ be continuously differentiable functions. Show that
$$\oint_C f\nabla g \cdot d\alpha = -\oint_C g\nabla f \cdot d\alpha$$
for every piecewise smooth Jordan curve C in S. [10]

5. Identify the space of all $n \times n$ matrices with $\mathbb{R}^{n^2}$. Show that the map $A \mapsto A^2$ is invertible on some open set containing the identity matrix. [15]

6. Let $A \subseteq \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}^n$ is continuously differentiable 1-1 function such that $\det f'(x) \neq 0$ for all $x \in A$. Show that $f(A)$ is an open set and $f^{-1} : f(A) \to A$ is differentiable. Show also that $f(B)$ is open for any open set $B \subseteq A$. [10]

PTO

7. Let $S = \mathbb{R}^2 \setminus \{(0,0)\}$ and let $f = (f_1, f_2)$ be a vector field defined on S by the equation

$$f(x,y) = \frac{-y}{x^2+y^2}\mathbf{i} + \frac{x}{x^2+y^2}\mathbf{j}.$$

(a) Show that $D_1 f_2 = D_2 f_1$ everywhere on S but f is not a gradient on S. [5]

(b) Show that f is a gradient on the set [10]

$$T = \mathbb{R}^2 \setminus \{(x,y) : y = 0, x \leq 0\}.$$

8. Sketch the region

$$S = \{(x,y) : x^2 \leq y \leq 2, -1 \leq x \leq 1\}$$

and express the double integral $\iint_S f(x,y)dxdy$ as an iterated integral in polar co-ordinates. [8]

9. Evaluate

$$\iiint_S \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) dxdydz,$$

where S is the solid bounded by the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$ [12]

INDIAN STATISTICAL INSTITUTE

Backpaper Examination, 1^{st} Semester, 2016-17

Statistical Methods III, B.Stat 2^{nd} Year

Date: 28. 12. 2016 Time: 3 hours

The paper carries 100 marks. Answer all questions.

1. Suppose $(2.3, 1.7, 3.2)$ are three independent observations from the density $f(x) = \frac{1}{2\lambda} exp\{-|x-\theta|/\lambda\}; \quad -\infty < x < \infty$ where, θ and λ are unknown parameters. Obtain the maximum likelihood estimates of θ and λ based on tha above data. Examine whether the maximum likelihood estimator of θ is unbiased for θ. [20]

2. Suppose $X_1, X_2, \ldots, X_{10}$ is a random sample from the density $f(x) = \beta/x^2, x \geq \beta$ and $\widehat{\beta}$ is the maximum likelihood estimator of β. Examine whether $\widehat{\beta}$ is a consistent estimator of β. Consider estimators of the form $c\widehat{\beta}$ to estimate β, where c is a real number. Which estimator in the above class has the minimum mean squared error? [15]

3. Assume that the number of times an individual has defaulted in his credit card payment is distributed as Poisson with mean λ. A survey is carried out among n randomly selected individuals (irrespective of whether they possess credit cards or not) as to their default history. Of course, those who do not possess credit cards would answer the survey question as zero. Thus, the probability that the recorded number of default cases is zero would be higher than that under a Poisson distribution. (*This is known as a zero-inflated Poisson distribution*). Given the frequency distribution of the recorded number of default cases, describe an EM algorithm to estimate λ and the proportion, p, of individuals in the population possessing credit cards. Show all the computational steps clearly. [15]

4. Suppose data are available on two binary variables X and Y for n unrelated individuals. If one uses logistic regression to model the conditional distribution of Y given X, obtain the maximum likelihood

estimators of the parameters of the logistic model. If one uses linear regression to model the dependence of Y on X, determine the LAD regression line. [15]

5. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from an exponential distribution with mean θ. Consider a test for $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ which rejects H_0 if and only if $\overline{X} > k$. If the size of the test is α, determine the value of k in terms of the quantiles of a chi-squares distribution. Draw a rough sketch of the power function of the above test. If n is large, show that the power of the above test at $\theta = \theta_1 (> \theta_0)$ can be approximated by $\Phi\{\frac{\sqrt{n}(\theta_1 - \theta_0) - \theta_0 z_\alpha}{\theta_1}\}$ where Φ and z_α are the c.d.f. and the (1-α)th quantile, respectively, of a standard normal variate. [15]

6. (a) It is claimed that the weight of the salt (in grams) produced in a chemical reaction is at least doubled by increasing the temperature from 70^oC to 80^oC. An investigator who suspects that the claim is false carries out the chemical reaction five times independently at both the temperatures and obtains the following results:

Experiment	Weight at 70^oC	Weight at 80^oC
1	238	457
2	264	501
3	253	512
4	227	460
5	259	506

Based on the above data, do you think that the investigator's suspicion is validated?

(b) Suppose one wants to estimate the proportion of a population that supports a particular legislation using a 95% equal tail asymptotic confidence interval. What is the minimum sample size required such that the error in estimation is at most 0.03? If the confidence interval obtained from a survey is (0.71,0.78), would you conclude at level 0.05 that at least 75% of the population supports the legislation? [10 + 10]

Indian Statistical Institute

(Back Paper), Semester Examination: 2016

Molecular Biology B. Stat. II

30|12|2016

Full marks: 50, Answer any Five, Duration: 2hr 30min

(Each question carries equal marks)

1. (a) Polymerase chain reaction and cell culture experiments could be used for synthesis of huge amount of DNA. What are the distinctive features between these two experiments? [5]

(b) Why the number of nucleotides of the same gene is different in bacteria and human? [5]

2. A man, who is color blind and possesses "O" blood group, has children with a woman who has normal color vision and "AB" blood group. The woman's father had color blindness. X-linked and autosomal genes determine color blindness and blood group, respectively. (a) What are the genotypes of the man and the woman? (b) What proportion of the children will have color blindness and type "B" blood group? (c) What proportion of their children will be color blind and have type "AB" blood group? [3+3+4]

3. (a) What is Okazaki fragments in DNA replication? What are sense and anti-sense strands in a double-stranded DNA? [3+2]

(b) Distinguish between transcription and translation. [5]

4. Can a reason be ascribed why genetic codon consists of three nucleotides instead of two or one nucleotide? In what sense and to what extent is the genetic code degenerate, comma-less and universal? [5+5]

5. How did Gregor Mendel arrive at the conclusion from his experiments that an allele (at a locus with two alleles) may be dominant over the other? In the blood group system (A, B, O and AB), how many alleles are present? Give examples of dominant, recessive and co-dominant genotypes in blood group system. [5+5]

6. If both the parents are carriers of sickle cell anemia, a recessive disease caused by mutation in one gene, then find out the chances that (a) all five children will be normal (b) four children will be normal and one will be affected by disease (c) at least three children will be affected by disease and (d) the first child will be a normal girl. [2.5 x 4]

INDIAN STATISTICAL INSTITUTE
Back Paper Examination: 2016-17

Subject: **Elements of Algebraic Structure** Course: B.Stat. II yr.
Maximum Score: 100 Date: 30th Dec. Duration : 3 Hours

Note: Attempt all questions. Marks are given in brackets. State the results clearly you use. Use separate page for each question. Notation are followed as taught in classes.

Problem 1. Let $F \leq E_1 \leq E$ and $F \leq E_2 \leq E$ be field extensions such that $[E_1 : F] = m$ and $[E_2 : F] = n$. Let E' denote the smallest subfield of E containing E_1 and E_2.
(i) Prove that $[E' : F] \leq mn$.
(ii) If m and n are relatively prime then show that $[E' : F] = mn$. [12+6 = 18]

Problem 2. Show that $\mathbb{Q}(i + \sqrt{2})$ is the splitting field of $x^4 - x^2 - 2$ over $\mathbb{Q}$. [15]

Problem 3. Let $\mathbb{R}[x, y]$ be the set of all bivariate polynomials. Show that it is not an Euclidean domain. [12]

Problem 4. Find out all ring homomorphisms $f : \mathbb{Z} \to \mathbb{Z}_n$. [10]

Problem 5. For every positive even integer $2n$, find a non-cyclic finite group of size $2n$ for which converse of the Lagrange's theorem holds. You need to show that your example is not a cyclic group. [15]

Problem 6. Let G be a group of size p^2 where p is a prime. Show that G is either isomorphic to $\mathbb{Z}_{p^2}$ or to $\mathbb{Z}_p \times \mathbb{Z}_p$. [15]

Problem 7. Let G be a cyclic group of size n. Show that for every divisor d of n, there exists exactly one subgroup of order d. Hence or otherwise show that $\sum_{d|n} \phi(d) = n$ where $\phi(d)$ denotes the Euler's totient function. [10+5 = 15]

Indian Statistical Institute

Mid-semester of Second Semester Examination: 2016-17

Course Name: BSTAT II

Subject Name: Economic and Official Statistics

Date: 20.02.17 **Maximum Marks: 30** **Duration : 1 hour**

Note: No marks will be awarded for answering any part of a question.

1. Explain how the Cost of Living Index number is calculated ? 6

2. Given the Average Wholesale Prices of four groups of commodities from the year 2012 to 2016, Compute the chain base index number. 6

Commodity	2012	2013	2014	2015	2016
X1	2	3	4	2	7
X2	3	6	9	4	3
X3	4	12	20	8	16
X4	5	7	18	11	22

3. Write Short Notes on the following: a) Millennium Development Goal and performance of India b) Methodology of crop estimation undertaken by Government of India and the states of India? C) Components of All India Weather Summary and forecast bulletin. 6*3=18

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination 2016-17

Course Name: B.Stat. (Hons.) 2nd Year

Subject: Economic and Official Statistics and Demography

Date: February 20, 2017 Maximum Marks: 40 Time: 1.5 hours

(Use two separate answer booklets for Group A and Group B.)

Group: B (Demography)

(Answer as much as you can.)

1. Write short notes on the following.

   a) Population Momentum

   b) Demographic dividend

   c) U.N. Joint Score

   d) Age heaping and age shifting (2+3+4+3 = 12)

2. a) Explain the logic behind the construction of the Whipple Index.

   b) Construct and explain the table for calculation of the Myers' Index. (4+6 = 10)

3. Given that the quality of age data in most developing nations is poor, suggest a method of adjustment for the age distribution of a developing nation, which is done by comparing with the age distribution of a suitable stable population, assuming it as a standard. State the necessary assumption(s). (10)

4. a) Write mathematical expressions for Sex Ratio Score along with explanations of the used symbols.

   b) State the formula, used in Carrier-Farrag Ratio Method, along with all necessary explanations. (4+4 = 8)

5. State the main components of the Sample Registration System (SRS). Elaborate on the sample design and the sample size in the SRS. (3+4+3 = 10)

= END =

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

B. Stat. - II Year (Semester - II)

Discrete Mathematics

Date : 21·02·17 Maximum Marks : 60 Duration : 3:00 Hours

Note : You may answer any part of any question, but maximum you can score is 60.

1. Recall that $R_k(3) = R_k(3,3,\dots 3)$ is the minimum integer n such that any k-coloring of edge set of K_n contains a monochromatic K_3. Prove or disprove that $R_k(3) \leq 3(k!)$ for any integer $k \geq 1$. [15]

2. Let $G(V,E)$ be a graph. $|V| = n$ and G does not contain any 4-cycle. Prove that $E(G) \leq \frac{n}{4}(1+\sqrt{4n-3})$. [13]

3. Intersection number of a given graph G is represented by $\omega(G)$. Prove or disprove the following statements.

   a) The number of cliques of G does not exceed $\omega(G)$.

   b) The number of cliques of G is not less than $\omega(G)$.

   [6+6=12]

4. Let $G(V,E)$ be an undirected graph having degree sequence $d_1 \geq d_2 \geq \dots \geq d_n$. Suppose $|V| = n$ and $|E| = m$. Let $t_i(G)$ be the number of triples of $V(G)$ having i edges in between each triples where $0 \leq i \leq 3$. Similarly, $t_i(\bar{G})$ be the number of triples of $V(G)$ having $0 \leq i \leq 3$ edges in $\bar{G}$. Prove that $t_0(G) + t_3(G) = \binom{n}{3} - m(n-2) + \sum_{i=1}^{n} \binom{d_i}{2}$. Hence or otherwise prove that $t_3(\bar{G}) + t_3(G) = \binom{n}{3} - m(n-2) + \sum_{i=1}^{n} \binom{d_i}{2}$ [13+5=18]

5. Let $d(u,v)$ represent the length of the shortest path between two vertices u and v that is the distance between two vertices u and v. Consider a tree T having n number of vertices. Observe that $T-u$ form one or many disjoint components and vertex u has exactly one neighbor in each such components. Prove that $\sum_{v\in V} d(u,v) \leq \binom{n}{2}$. [Hints: You may use induction on size of tree.] [15]

INTRODUCTION TO MARKOV CHAINS
B. STAT. IIND YEAR SEMESTER 2
INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination
Time: 2 Hours Full Marks: 35
Date: February 22, 2017

This is an OPEN NOTE examination. You may use any handwritten notes that you have brought with you, but you are not allowed to share them. Also no notes in printed, photocopied, cyclostyled or any other form (including electronic storage) is allowed.
Calculators and any other computing devices are strictly not allowed.

1. A company desires to operate s identical machines. Each working machine fails within a week independently of each other with probability p. At the end of the week, orders for new machines are placed to replenish the failed machines to make up for the total number of machines as s. The ordered machines take a further week to be delivered. Let X_n be the number of working machines at the beginning of the n-th week. Show $\{X_n\}$ is a Markov chain and write down its transition matrix. [2+2=4]

2. For a Markov chain, define

$$h_{ij}^{(n)} = \mathrm{P}_i[X_m \neq i, 1 \leq m < n, X_n = j], \quad n \geq 1 \qquad \text{and} \qquad h_{ij}^{*} = \sum_{1}^{\infty} h_{ij}^{(n)}.$$

Show that, for $i \neq j$ and $n \geq 1$,

$$p_{ij}^{(n)} = \sum_{m=0}^{n-1} p_{ii}^{(m)} h_{ij}^{(n-m)}.$$

Also show that $i \rightsquigarrow j$ iff $h_{ij}^{*} > 0$. [3+3=6]

3. For a simple random walk on $\mathbb{Z}$, where the particle moves right with probability p and left with probability $1-p$,for $p \in [0,1]$, show that $P_{ii}(s) = (1 - 4p(1-p)s^2)^{-1/2}$.
~~underline~~Hence show that the walk is recurrent iff $p = 1/2$.
In the recurrent case, obtain $F_ii(s)$ and hence show that the walk is null recurrent.
Hence show that SSRW in dimension 2 is null recurrent. [4+3+2+2+5=16]

4. Indicate whether the following statements are True or False, with appropriate justifications in each case: [3× 3 =9]
   (a) If $\{X_n\}$ is a Markov chain,then so is $\{X_{2n}\}$.
   (b) It is not possible to have a Markov chain on two states, one of which is transient and the other being recurrent.
   (c) Any sequence of random variables $\{X_n\}$, where X_n follows Bernoulli distribution with parameter 1/2, must be a Markov chain.

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2016-17

Course Name : B. STAT. II Year (Second Semester)

Subject Name : Physics II

Date : 23/02/2017 **Maximum Marks : 40** **Duration : $2\frac{1}{2}$ hours**

Use separate answer sheet for each group

Group A

Total Marks: 20

Answer any Four: All questions carry equal marks

1. Which of these magnetic field can exist ? Determine the current density that created the valid fields. Determine the corresponding vector potential.

   (a) $\mathbf{B}(\mathbf{r}) = e^{-y^2}\hat{x}$ (cartesian coordinate system)

   (b) $\mathbf{B}(\mathbf{r}) = \sin(kr)$ (cylindrical coordinate system)

2. (a) Determine the vector potential of a finite segment of straight wire, carrying a current I.

   (b) Calculate the force per unit length by two infinitely long straight conductors on each other.The conductors are arranged parallel to each other at a distance d.

3. Given the magnetic vector potential at point (r, θ, ϕ) is

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{m\sin\theta}{r^2}\hat{\phi},$$

P.T.O

where the symbols have their usual meanings. Find the magnetic field of this dipole of moment $\mathbf{m}$.

4. An infinitely long circular cylinder carries a magnetization $\mathbf{M} = ks\hat{\mathbf{z}}$ parallel to its axis, where k is a constant and s is the distance from the axis. There is no free current anywhere. Find the magnetic filed inside and outside the cylinder.

5. (a) Explain the hysteresis loop

   (b) Show that in a linear homogeneous material volume bound current density is proportional to the free current density.

Group B

Total Marks: 20

Answer any four questions: All questions carry equal marks

1. During a process a system receives 30 kJ of heat from a reservoir and does 60 kJ of work. Is it possible to reach initial state by an adiabatic process?

2. 1 kg of gas enclosed in an isolated box of volume v_1, temperature T_1 and pressure p_1 is allowed to expand freely till volume increases to $v_2 = 2v_1$. Determine the change in entropy. (R=287 kJ/kg K).

3. Show that for a Maxwellian gas

$$\bar{c}\overline{\left(\frac{1}{c}\right)} = \frac{4}{\pi},$$

where c denotes molecular speed.

4. Defining volume expansivity (α) and isothermal compressibility (β_T) as $\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p$ and $\beta_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T$ respectively, show that $C_p - C_v = \frac{\alpha^2 TV}{\beta_T}$, the symbols having their usual meaning.

5. Show that for a Van der Waal gas

$$\left(\frac{\partial C_v}{\partial V}\right)_T = 0$$

where the symbols have their usual meaning.

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination 2016 – 2017
B. Stat. II Year
AGRICULTURAL SCIENCE

Date : 23/2/17 **Maximum Marks : 30** **Duration :Three hours**

(Attempt any three questions)
(Number of copies of the question paper required : 15)

1. Define drought. What are the different types of drought? Write any two of the drought indices to measure the extent of drought. What is contingent cropping strategy to cope the drought situation.

2+4+2+2

2. Name different meteorological variables that are related to crop production. Write the names of apparatus used to measure those variables with their unit. Why the ambient humidity is considered as "RELATIVE HUMIDITY"?

4+4+2

3. Write short notes on the following
a) PET b) Onset of monsoon c) cessation of monsoon
d) Microclimate e) Cup counter anemometer.

2 X 5

4. Calculate the amount of Potential Evapotranspiration in mm from the following data:

a) Water added at 0830 hrs. IST on 21 June 2016 to bring the water level to the reference point is 32 cm. Rainfall for last 24 hrs. is nil.

b) Water removed at 0830 hrs. IST of 25July 2016 is 10.3 mm. The rainfall recorded at 0830 hrs. IST of 25 July 2016 is 4.8 mm

c) The water added at 0830 hrs. IST on 5 July 2016 is 32 cm. The rainfall recorded at 0830 on 5 July 2016 is 1.6 mm.

d) The area of a rice crop field is 1.25 hectare and the rainfall recorded in a day is 40 mm. Calculate the volume of water in kilo litres received by the Rice crop.

2.5 X 4

Indian Statistical Institute
Mid Semester Examination 2017
Course Name: B Stat Second Year
Subject Name: Economics II (Macroeconomics)

Date of Examination: 23 /02/2017 **Maximum Marks – 40** **Duration: 2 Hours**

Answer all questions

1. Explain the concepts of personal saving, private saving , national saving (NS) and, hence, derive the identity involving NS on the one hand and investment and current account balance on the other. Is this investment net or gross? [10]

2. Suppose in an economy in a given year the central bank had to sell foreign exchange worth Rs.20000 crore from its stock to hold the exchange rate at the target level. Net inflow of foreign loan to the domestic economy was Rs.18000 crore. In addition, foreigners purchased shares of domestic companies worth Rs. 8000 crore. Domestic residents purchased land abroad of Rs.200 crore. Difference between NNP and private disposable income in the domestic economy in the given period was Rs.20020 crore, public sector enterprises' profit was Rs.20 crore and government expenditure on goods and services (G) was Rs.20540 crore. The economy in the given period produced machinery equipment and construction worth Rs.60,000 crore and the firms could sell all the consumption goods they produced. All the raw materials produced in the economy were fully used up as intermediate inputs. All imports consisted of machinery and equipment worth Rs.10000 crore. Depreciation in the domestic economy was Rs.500 crore. Exports consisted of consumer goods only. Aggregate consumption expenditure of the domestic economy was Rs.10,000 crore. From the given data, compute government saving, official reserve settlement balance, capital account balance, current account balance, net investment and private disposable income of the economy in the given period. [15]

3. The following are data for a hypothetical economy for the year1993 in billions of rupees.

Net rental income of persons	24.1
Depreciation	669.1
Wages and salaries of employees	3780.4
Personal consumption expenditure	4378.2
Sales and excise taxes	525.3
Business transfer payments	28.7

P.T.O

Gross investment	882.0
Exports of goods and services	659.1
Subsidies paid by the government	9.0
Government purchases of goods and services	1148.4
Imports of goods and services	724.3
Net interest	399.5
Proprietors' income	441.6
Corporate profits	485.8
Net factor income from the rest of the world	5.7

a. Compute GDP using the spending approach.

b. Compute net domestic product.

c. Compute national income in two ways, and compute the statistical discrepancy

[15]

INDIAN STATISTICAL INSTITUTE

Mid-Sem Examination, 2^{nd} Semester, 2016-17

Statistical Methods IV B.Stat 2^{nd} Year

Date: February 24, 2017 Time: 2 hours

This is an open notes examination. The paper carries 25 marks. Answer all questions.

1. Suppose $\mathbf{X} = (X_1, X_2, X_3)$ is distributed uniformly within a sphere with radius r.

   (a) Compute the expectation of $\mathbf{X}'\mathbf{X}$

   (b) Obtain the regression of ${X_1}^2$ on X_2 and X_3 [4+4]

2. Consider the usual regression model:

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

   where, x_is are fixed and e_is are i.i.d. N(0,σ^2). Show that the likelihood ratio test for $H_0 : \beta = \beta_0$ vs $H_1 : \beta \neq \beta_0$ rejects H_0 if and only if $\dfrac{|\widehat{\beta} - \beta_0|}{\sqrt{\Delta^2 - \widehat{\beta}^2}}$ is large, where $\widehat{\beta}$ is the least squares estimator of β and $\Delta^2 = \dfrac{\sum_{i=1}^n (y_i - \overline{y})^2}{\sum_{i=1}^n (x_i - \overline{x})^2}$. [6]

3. Suppose $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}$ is a random sample from $N_p(\mu, \sigma^2 I)$ such that the sample mean vector and the sample variance-covariance matrix are $\overline{\mathbf{X}}$ and S, respectively. What is the distribution of $\dfrac{\overline{\mathbf{X}}'\overline{\mathbf{X}}}{trace(S)}$? [5]

4. Suppose $\{(X_1, Y_1), (X_2, Y_2) \ldots, (X_n, Y_n)\}$ is a random sample from a bivariate normal distribution with parameters $(\mu, -\mu, \sigma^2, \sigma^2, \rho)$. Obtain the maximum likelihood estimators of μ, σ^2 and ρ. [Hint: Consider two linear combinations of X_i and Y_i that are uncorrelated] [6]

INDIAN STATISTICAL INSTITTUE, KOLKATA

Second Semestral Examination 2016-17

B. Stat. (Hons.) II Year

Subject: Economic and Official Statistics and Demography

Date: 24.04.17 Maximum Marks 100 Duration: 4 hours

*(**Instructions**: This question paper has two Groups, A and B. Answer Group A and Group B on separate answer booklets. **Each Group carries a maximum of 50 marks**. Marks allotted to each question are given within parentheses. **The total duration is 4 hours**. Standard notations are followed.)*

Group A: Demography and Economic Statistics

(Answer as many as you can.)

1. Write short notes on the following. All used symbols must be clearly defined.

a) Gross Reproduction Rate and Net Reproduction Rate;

b) Important properties of Lorenz curve;

c) Relation between Lorenz Ratio and Gini Mean Difference.

[4 + 4 + 4 = 12]

2. Answer the following.

i) The equation

$$q_x = 2m_x / (2 + m_x)$$

shows how the m-type and q-type mortality rates are related to one another. Derive a similar equation for the more general case with age groups of width n years.

ii) The standardized death rate for town A was 1.23 when the population of town B was used as the standard. What does this tell you about mortality in A to that in B?

iii) Show that the crude birth rate in a stationary population corresponding to a life table is equal to $(1/e_0)$ where e_0 is the life expectation at birth. [4 + 2 + 4 = 10]

3. (a) Express $_nL_x$ in terms of l_x, $_na_x$ and $_nd_x$.

(b) Why q_x cannot be calculated directly? Describe how it gets estimated.

(c) If the crude birth rate in a country remains constant over a number of years, but the general fertility rate increases steadily, what does this tell you about the dynamism of the country's population?

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

4. Explain how Chandra Sekar and Deming have estimated the number of vital events, missed by both the sample registration system and the sample survey. State clearly the conditions that are to be satisfied for validity of the estimate.

[4 + 3 = 7]

5. What is the purpose of El-Badry's procedure? Explain the procedure. All symbols must be defined clearly.

[1 + 9 = 10]

6. The table below gives the parity progression ratios for a number of birth cohorts in a country.

(i) Assuming that no woman in any of these birth cohorts had a fifth child, calculate

(a) the proportion of women in each birth cohort who had exactly 0, 1, 2, 3 and 4 children,

(b) the total fertility rate for women in each birth cohort.

(ii) Comment on your results.

Calendar years of birth	Parity Progression Ratios			
	0 -1	1 – 2	2 – 3	3 – 4
1931-33	0.861	0.804	0.555	0.518
1934-36	0.885	0.828	0.555	0.489
1937-39	0.886	0.847	0.543	0.455
1940-42	0.890	0.857	0.516	0.416
1943-45	0.892	0.854	0.458	0.378
1946-48	0.885	0.849	0.418	0.333

[(6 + 3) + 1 = 10]

Group B: Economic and Official Statistics

[Note: Answers should be to the point and brief. Group B has two parts – I and II]

Part I: *Answer Question no. 1 and any two from the rest*

1. State the exact forms and elasticities of Engel Curve related to the following non-linear forms: Log inverse, Semi- log, Double- log. [3+3+3=9]

2. a) State the forms of Cobb Douglas and CES Production Functions along with all assumptions.

   b) Derive the elasticity of substitution for Cobb Douglas Production Function. [4+4=8]

3. a) Explain the properties of Price Index numbers.

   b) Briefly discuss about the types of errors observed while computing such index numbers.

   [5+3=8]

4. Explain the Homogeneity Hypothesis for effects of variation in household size in Engel curve analysis. Also state the implications of this hypothesis. [5+3=8]

Part II: *Answer Question no. 1 and any two from the rest*

1. Explain the Sampling Registration Scheme followed by the Office of the Registrar General and Census Commissioner, India. [9]

2. Explain the types of Poverty Data published by the World Bank. [8]

3. Mention, in brief, the functions of Niti Aayog. [8]

4. State the key indicators of All India Debt and Investment Survey conducted by NSSO (70th Round). [8]

5. Explain, in brief, the type of Industrial Statistics compiled by CSO, Ministry of Statistics and Programme Implementation, Government of India. [8]

INDIAN STATISTICAL INSTITUTE

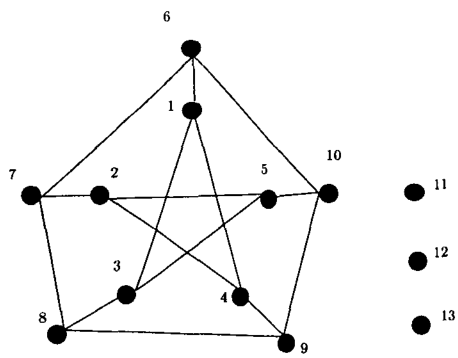
Semestral Examination

B.Stat - 2nd Year (Semester - II) *Discrete Mathematics*

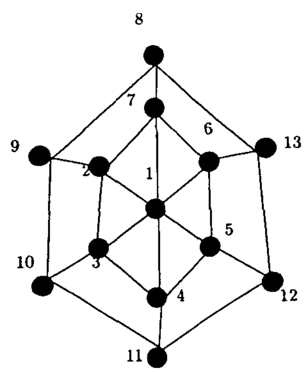
Date : 26 April, 2017 Maximum Marks : 100 Duration : 3 Hours

Note : You may answer any part of any question, but maximum you can score is 100. Notations used is as defined in class.

1. Find the intersection number of the following graph. [5]



2. Is it possible to draw the following planar graph in a plane such that the vertices 1, 2 and 3 are the only vertices in the outer face? [5]



3 A tree with 100 vertices have maximum degree 10. There exist a vertex whose degree is 5. Suppose the number of leaf nodes (that is vertices having degree 1) in the tree is n. Which of the following is/are true?

(a) $13 \leq n \leq 72$

(b) $15 \leq n \leq 72$

(c) $15 \leq n \leq 99$

(d) $13 \leq n \leq 88$

(e) $15 \leq n \leq 100$

[5]

4 For any graph $G(V, E)$, which of the following is/are true?

(a) $\chi(G) \leq \frac{|V|}{\text{Cardinality of maximum independent set in } G}$,

(b) $\chi(G) \leq |V| - 2$(Cardinality of maximum independent set in G),

(c) $\chi(G) \leq |V|-$(Cardinality of maximum independent set in G)-2,

(d) $\chi(G) \leq |V|-$(Cardinality of maximum independent set in G)$+1$,

(e) $\chi(G) \leq$ Maximum degree among vertices in G.

[5]

5 In a graph $G(V, E)$, every subgraph $H(V_H, E_H)$ of G has an independent set consisting of at least half of $|V_H|$. Which of the following is/are true?

(a) G is planar

(b) G is bipartite

(c) G is chordal

(d) Degrees of vertices of G are all equal

[5]

6 G is a triangle free planar graph. Let $\{d_1, d_2, d_3, \cdots, d_n\}$ be its degree sequence where, $d_1 \leq d_2 \leq d_3 \leq \cdots d_n$. The maximum value of d_1 is

(a) 1

(b) 2

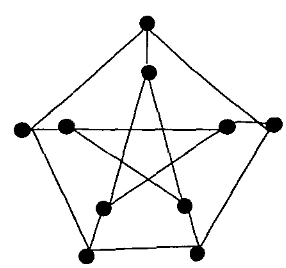
(c) 3

(d) 4

(e) 5

[5]

7 In a graph G, a *k cycle* is an ordered list of k distinct vertices $v_1, v_2, \cdots, v_k$ such that all $v_{i-1}v_i$, $2 \leq i \leq k$ and $v_k v_1$ are edges of G. What is the number of 6 cycles in the following graph? [5]



8 Draw all graphs satisfying all of the following properties.

(i) Plane graph with all vertices lying on a particular face

(ii) Triangle free

(iii) Not a tree

(iv) Non adjacent vertices have at least one common neighbour

[5]

9 Let G be a graph having 30 vertices and minimum size of vertex cover is 12. (*Vertex cover* of a graph G is a set S of vertices such that S contain at least one end points of every edge of G.) Which of the following is/are true?

(a) Maximum size of matching in G is 18

(b) Maximum size of matching in G is 12

(c) Maximum size of independent set in G is 18

(d) Chromatic number of graph G is greater than 12

[5]

10 Draw all graphs satisfying all of the following properties:

(i) Number of vertices is 10

(ii) Triangle free

(iii) Degree of all vertices is 3

(iv) Non adjacent vertices have exactly one common neighbour

[5]

11 Two people play a game on a graph G by alternately selecting distinct vertices $v_1, v_2, \cdots$ forming a path. The last player to select a vertex wins. If the graph satisfies property X, the second player wins; otherwise, the first player wins. Describe property X using a graph class (like bipartite, chordal etc.). [5]

12 Draw a triangle free graph whose chromatic number is 4. [5]

13 For any graph $G(V, E)$, which of the following is true?

(a) $\chi(G) + \chi(\overline{G}) \leq |V|$

(b) $\chi(G) + \chi(\overline{G}) \geq |V|$

[5]

14 Present the smallest graph G which does not have perfect elimination ordering but $\chi(G)$ is equal to the size of maximal complete subgraph. [5]

15 Suppose T is a tree. Suppose $T_1, T_2, \cdots, T_k$ are k different connected subgraphs of the tree T. Let $G(V, E)$ be a graph where $V = \{v_1, v_2, \cdots, v_k\}$ and $(v_i, v_j) \in E$ if some vertex in T is common in both T_i and T_j. Which of the following is/are true?

(a) G is bipartite,

(b) G is planar,

(c) G is chordal,

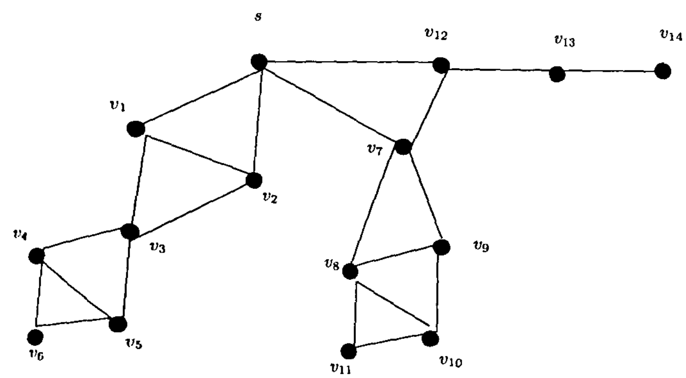
(d) G is a tree,

(e) G is Hamiltonian.

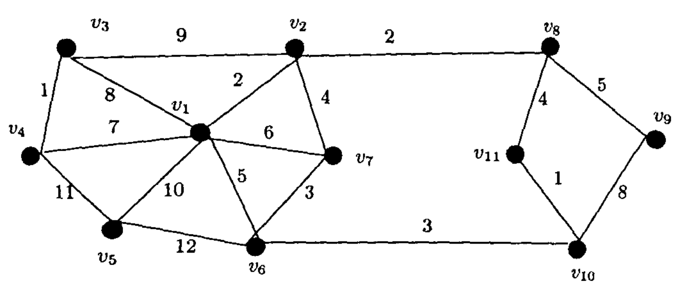
[5]

16 An integer x be such that $R_4(3,3,3,3) \leq x$. Find the possible value of x from the set $\{6, 18, 66, 72, 100\}$.

17 Draw breadth first tree and depth first tree for the graph given below starting from distinguished source vertex s. [5]

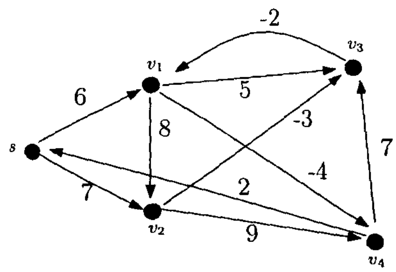


18 Draw a minimum spanning tree for the following graph. [5]



19 From the source vertex s, compute a single source shortest path to all vertices for the weighted directed graph given below. [5]

20 A triangulation of an n-gon is a division of the interior into triangles. Diagonal triangulation is a triangulation where it makes use only of



internal nonintersecting diagonals of the n-gon. Let T_n be the number of different diagonal triangulations of an n-gon. Find a recurrence formula for T_n. [5]

21 Let $(a_1, a_2, \ldots, a_n)$ be a permutation of the numbers $1, 2, \ldots, n$, such that no element is back in its original place, that is, $a_1 \neq 1, a_2 \neq 2, \ldots, a_n \neq n$. Such a permutation is called a derangement. Let D_n denote the number of derangement of the set $\{1, 2, \ldots, n\}$. Find the recurrence relation for D_n. [5]

22 Let $S_1, S_2, \ldots, S_k$ be k subsets of the set $S = \{a_1, a_2, \ldots, a_n\}$ such that $|S_i|$ is even and $|S_i \cap S_j|$ is odd for every i, j where, $i \neq j$, $1 \leq i, j \leq k$. Find the maximum value of k in terms of n where such k subsets can be generated. [5]

INDIAN STATISTICAL INSTITUTE

Second Semester Examination : 2016-17

Course Name : B. STAT. II Year (Second Semester)

Subject Name : Physics II

Date : 28|04|2017 **Maximum Marks :** 50 **Duration :** 3 hours

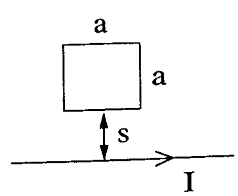
Use separate answer sheet for each group

Group A

Answer any **Five** Questions

1. Suppose two metal objects are embedded in weakly conducting material of conductivity σ.
(a) How is the resistance between them related to the capacitance of the arrangement?
(b) A battery is then connected between the two objects and there is a potential difference V_0. What will happen if you disconnect the battery?
(c) How the voltage will change with time? Write your answer in terms of permittivity and conductivity.
[2+1+2]

2. Suppose there is a very long straight wire carrying current I. Assume that a square loop of wire with side a is placed at a distance s above the wire, as shown in figure below.



(a) Find the flux passing through the loop.
(b) What electromotive force is generated if the square loop is directly pulled away from the straight wire at a speed v?
(c) State the direction of the current flow.
[2+2+1]

3. (a) Two co-axial long cylinders, one inside the other, of radii $a = 4$cm and $b = 8$cm are separated by a material of conductivity $\sigma = 1.68 \times$

P.T.O

10^{10}(ohm-cm)$^{-1}$. If they are maintained at a potential difference $V = 6$ Volts, what current will flow from one to the other, in a length $L =$ 0.75m?

(b) Find the resistance for this configuration in proper unit.

[4+1]

4. (a) Suppose you have two loops of wire at rest: loop 2 is above loop 1 some distant apart. If you pass a steady current I_1 around loop 1, show that the flux through loop 2 is proportional to I_1. Name the proportionality constant and state some of its properties.

   (b)Find the self-inductance per unit length of a long solenoid, of radius R, carrying n turns per unit length.

   (c) Let the amount of charge per unit time passing down a wire be I. What is the work done per unit time in terms of its self inductance? What will be the total work done, if you start with a zero current and build it up to a final value I?

[(1+1)+1+(1+1)]

5. Suppose in a region an electromagnetic field is characterized by the vector potential $\mathbf{A} = \hat{\mathbf{i}}A_0 \sin(kz - \omega t)$ and electric field $\mathbf{E} = \hat{\imath}E_0 \cos(kz - \omega t)$. Here, $\hat{\imath}$ is the unit vector in the x-direction, A_0 and ω are known constants. Using the Maxwell's equations find the values of $\mathbf{B}$, E_0 and k.

[5]

6. (a) Suppose there is some charge and current configuration, which at time t, produces fields **E** and **B**. Find the rate at which work is done on all the charges in a volume V and make a comment on work-energy theorem of electrodynamics.

   (b) Find the energy per unit time per unit area transported by the fields $\mathbf{E} = [E_x\hat{\mathbf{i}} + E_z\hat{\mathbf{k}}](\cos ky + \omega t)$ and $\mathbf{B} = -\frac{k}{\omega}[E_z\hat{\mathbf{i}} + kE_x\hat{\mathbf{k}}(\cos ky + \omega t)]$. Here, $\hat{\mathbf{i}}$ and $\hat{\mathbf{k}}$ are the unit vectors in x and z direction respectively and ω and k are known constants.

[3+2]

Group B

Answer any **Five** questions

1. Show that
   (a) $\left(\frac{\partial C_p}{\partial p}\right)_T = -T\left(\frac{\partial^2 V}{\partial T^2}\right)_p$
   (b) C_p of an ideal gas is a function of T only, the symbols having their usual meaning.

   [3+2]

2. Using Maxwell distribution of velocities, find the fraction of the molecules having kinetic energy in the range E and $E + dE$. Hence obtain the most probable value of the kinetic energy.

   [3+2]

3. Consider a system of simple harmonic oscillators, the energy of an individual oscillator being $E_i = \hbar\nu(i + \frac{1}{2}), i = 0, 1, 2, \ldots.$ Compute the partition function and hence determine the average energy per oscillator. The symbols having their usual meaning.

   [3+2]

4. A system consisting of two particles, each of which can be in any of three quantum states of energies 0, ϵ and 3ϵ, is in thermal equilibrium at temperature T. Determine the partition function of the system when the particles obey (a) Bose-Einstein statistics (b) Fermi-Dirac statistics.

   [3+2]

5. Let $g(E)$ and E_F denote respectively the density of states and Fermi energy in a metal. Obtain an expression for the total number of electrons in the system at $T = 0$.

   [5]

6. Consider a photon gas in equilibrium contained in a cubical box of volume a^3. Calculate the number of states in the frequency range ν to $\nu + d\nu$.

   [5]

Indian Statistical Institute
End-Semester Examination 2017
Course Name: BStat Second Year
Subject Name: Economics II (Macroeconomics)

Date 28·04·17 Maximum Marks: 60 Time: 2.5 Hours

Answer the following questions:

1. (i) Consider the Simple Keynesian Model for an open economy with government. Suppose the government raises its expenditure on imported armaments and finances it with additional tax revenue. The tax is a lump sum. Examine its impact on domestic GDP.

(ii) Suppose a Simple Keynesian Model is given by the following equations:

$C = 160 + 0.8Y$; $I = 80 + 0.3Y$; $G = 30$; $X = 10$ and $M = 50 + 0.2Y$

Compute the multiplier effect of an exogenous increase in planned investment expenditure made entirely on imported goods,

(All symbols have their usual meanings.) [16 + 6 = 22]

2. (i) Consider a Simple Keynesial Model for a closed economy, with Government. Suppose it is given that $dC/dYD = 0.8$, $dI/dY = 0.3$. Can you design a tax function that will ensure stability of equilibrium in this model? Is this tax function unique? (Assume $R = 0$.)

(ii) A Simple Keynesian Model for a closed economy with government is given by the following set of equations: $C = 160 + 0.8YD$, $YD = Y - T + R$; $I = 120 + 0.3Y$; $G = 30$, $R = 0$; and $T = 20 + 0.5Y$.

(a) Compute the equilibrium level of Y.

(b) Consider another situation where all the functions given above remain the same except for the fact that a ceiling of 400 is imposed on the total tax collection. Tax collection continues to be given by the above tax function. However, total tax collection remains 400 for all values of Y higher than the one at which tax collection reaches the ceiling. What implications does it have for the aggregate demand function and the equilibrium of the model?

(c) How does your answer to (b) change if the ceiling on tax collection were 620 instead of 400? [4+10+8=22]

3. Consider an IS – LM Model given by the following set of equations: $C = 40 + 0.75Y$; $I = 300 - 50r$; $L = 0.25Y - 50r$; $M_0/P = 100$ (Take $P = 1$).

(i) Derive the equations of the IS and LM schedules.

(ii) Now, suppose M_0 is raised by 50 units. Drawing upon the assumption regarding the relative speeds of adjustment in Y and r, derive the immediate impact on r and I before any change had taken place in Y. State what will happen to Y and r finally.

Suppose the government wants aggregate planned investment to remain at the level that it attained at the initial equilibrium level of Y following the given increase in M_0. Which policy parameter the government has to change to achieve this without disturbing the expansion in Y? Derive the value of the needed change in the policy parameter. [4 +18=22]

Or

4. Suppose the government takes an additional loan of Rs.100 from the central bank. Explain how this will lead to an increase in the stock of high-powered money by the same amount. Assuming CRR to be unity, show how people will come to hold an additional amount of money of the same amount at the end of the operation of the money multiplier process when high-powered money goes up by Rs.100. [22]

INDIAN STATISTICAL INSTITUTE
Second Semestral Examination: (2016 - 2017)
B. Stat II Year
Agricultural Science

Date 28.04.17 Maximum Marks 50 Duration 3:00 hours.

(Attempt any five questions)

(Number of copies of the question paper required 15)

1. Draw a suitable rice calendar with the following data.

Week No.	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Rainfall (mm) at 0.5 Prob.	0	15	21	35	19	45	64	87	18	92	113	145	90	83	0	22	11	0	0
PET (mm)	45	30	28	26	35	27	24	22	34	23	19	17	24	24	35	30	31	35	36

(2+8)

2. Briefly describe the cultural practices associated with direct seeded rice. Why the Boro rice yields more than the Kharif rice

(7+3)

3. What are the criteria for the essentiality of plant nutrients. Calculate the quantity of VC, DAP, SSP and KCL required for 1 hectare wheat crop to supply the nutrient requirement of 180 kg N, 120 kg P_2O_5 and 120 kg K_2O per hectare. Note that 50% of required N should be given through VC.

(4+6)

4. What are the growth related and yield attributing chatercters of rice. Estimate the yield per hectare of mustard crop from the following data.
   (i) Average no. of branches/plant –15, (ii) Average no. of silique/branch –45, (iii) Average no. of seeds/siliqua –20, (iv) Average siliqua length -9 cm, (v) Test weight -48 g.

(4+6)

5. What is intercropping? Rice-Mustard intercropping experiment was done in 2:1 and 2:2 row replacement series system and yield data are given in the following table. In your opinion, which combination is the best intercropping system ?

(3+7)

Cropping System	Rice yield in kg/ha	Mustard yield in kg/ha
Rice Sole	3255	-
Mustard Sole	-	5876
Rice+Mustard (2:2)	1988	4125
Rice+Mustard (2:1)	2057	3197

6. Define irrigation. What are the different sources of irrigation water. Write in brief about different methods of irrigation.

(3+3+4)

7. Write the differences between:
   a) Manures and Fertilizers
   b) Inter cropping and mixed cropping
   c) Soil texture and soil structure
   d) Macro and micro nutrients

(2.5 x 4)

Introduction to Markov Chains
B. Stat. IInd Year Semester 2
Indian Statistical Institute

Semestral Examination
Time: 3 Hours Full Marks: 50
Date: 2nd May, 2017

This is an OPEN NOTE examination. You may use any handwritten notes that you have brought with you, but you are not allowed to share them. Also no notes in printed, photocopied, cyclostyled or any other form (including electronic storage) is allowed.
Calculators and any other computing devices are strictly not allowed.

1. Consider a simple symmetric random walk $\{X_n\}$ with barriers on the state space $\{-m, -m+1, \ldots, -1, 0, 1, \ldots, m-1, m\}$, where all internal states (other than $\pm m$) have equal probabilities of moving one step to the right or the left. For the barrier states $\pm m$, it stays put or goes to the neighbouring state with equal probability.
   (a) Write down the transition matrix of the walk. [2]
   (b) Is the chain $\{|X_n|\}$ Markov? In case it is, provide the transition matrix. In case it is not, give example of a transition which violates Markov property. [4]

2. **Susceptible-Infected-Susceptible (SIS) model:** Assume a population of n individuals. At the beginning of each day, each individual is either infected or susceptible (capable of contracting the flu). Suppose that each pair (i, j), $i \neq j$, independently comes into contact with one another during the daytime with probability p. Whenever an infected individual comes into contact with a susceptible individual, the former infects the latter. In addition, assume that overnight, any individual who has been infected for at least 24 hours will recover with probability $0 < q < 1$ and return to being susceptible, independently of everything else (i.e., assume that a newly infected individual will spend at least one restless night battling the flu).
   (a) Suppose that there are m infected individuals at daybreak. What is the conditional distribution of the number of new infections by day end? [2]
   (b) Draw a Markov chain with as few states as possible to model the spread of the flu for $n = 2$. [6]
   (c) Identify all recurrent states of the Markov chain obtained in 2(b). [2]

3. Let $f_1, f_2, \ldots$ be an infinite sequence of positive numbers adding up to 1. Define $F_n = \sum_{i=1}^{n} f_i$. Consider a Markov chain on the set of nonnegative integers given by $p_{i0} = f_{i+1}/(1 - F_i)$, $p_{i,i+1} = (1 - F_{i+1})/(1 - F_i)$, for all $i \geq 0$. Show that the chain is irreducible and recurrent. Find a necessary and sufficient condition for the chain to be positive recurrent. [2+3+5=10]

4. Consider a Markov chain $\{X_n\}$ on a finite state space S with probability transition matrix $\boldsymbol{P}$. Fix subset A of S, and define $\tau_A = \inf\{n \geq 0 : X_n \in A\}$. Also denote $p_A(x) = \mathrm{P}_x[\tau_A < \infty]$.
   (a) Show that $\mu = p_A$ satisfies the equation
$$\mu(x) = \begin{cases} 1, & \text{if } x \in A \\ \sum_{y \in S} p_{xy}\mu(y), & \text{otherwise.} \end{cases} \qquad (\star)$$
   [3]
   (b) Let q be any other nonnegative solution of the above equation $(\star)$. Show that for $x \notin A$ $q(x) = \mathrm{P}_x[\tau_A = 1] + \sum_{y \notin A} p_{xy} q(y)$. Show that, for any n, $q(x) \geq \sum_{i=1}^{n} \mathrm{P}_x[\tau_A = i]$. Prove that p_A is the smallest nonnegative solution of the equation $(\star)$. [4+4+2=10]

[P.T.O.]

5. Find a coupling (X, Y) such that $X \leq Y$ and X and Y follow Poisson distribution with parameters μ and ν respectively, with $0 < \mu < \nu$. [3]

6. **Knapsack problem:** Consider a knapsack with volume constraint b which needs to be packed with some of n items having volumes $a_1, a_1, \ldots, a_n$. Thus a packing configuration is a vector $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ satisfying $\sum_1^n a_i x_i \leq b$, where x_i is the indicator of i-th item in the knapsack. If $\sum_1^n a_i > b$, write down an algorithm, with justification, to generate a uniform sample from the packing configurations. How would you modify your algorithm if $\sum_1^n a_i \leq b$? [6+2=8]

INDIAN STATISTICAL INSTITUTE

Final Examination, 2^{nd} Semester, 2016-17

Statistical Methods IV, B.Stat 2^{nd} Year

Date: May 5, 2017 Time: 3 hours

The paper carries 65 marks. Answer all questions.
Use separate answer scripts for each group.

Group A: 33 marks

1. Consider two populations: $N_p(\mu_1, \Sigma)$ and $N_p(\mu_2, \Sigma)$. If an observation x is classified to that population which yields a higher likelihood, show that the above rule is equivalent to classifying x to the first population iff $(x-\mu_2)'\Sigma^{-1}(\mu_1-\mu_2) \geq \Delta^2/2$ where Δ^2 is the Mahalanobis Distance between the two populations. [5]

2. The Hardy-Weinberg Equilibrium law implies that at any locus with alleles $A_1, A_2, \ldots, A_k$ the probability distribution of the genotypes is as follows: $P(A_iA_i) = p_i^2 \ \forall\ i = 1, 2, \ldots, k$ and $P(A_iA_j) = 2p_ip_j$ $\forall\ i < j = 1, 2, \ldots, k$, where $\sum_{i=1}^{k} p_i = 1$. Data are collected on 146 patients suffering from Alcoholic Cirrhosis. The frequency distribution of the genotypes at a marker acid phosphatase (ACP) having three alleles is 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]

3. In a study to test whether the variability of systolic blood pressure (s.b.p.) is twice that of diastolic blood pressure (d.s.p), data were collected on s.b.p. and d.s.p. for 10 randomly chosen individuals. Based on the above sample, it was found that the mean and the standard deviation of s.b.p. were 126.4 mmHg and 13.8 mmHg, respectively; the mean and the standard deviation of d.s.p were 85.7 mmHg and 7.1 mmHg, respectively while the correlation between s.b.p. and d.s.p was 0.6. Stating your assumptions clearly, infer whether the hypothesis is validated by the data. [10]

4. Suppose $(0.6, -0.7, -1, 1.3, 0.2)$ is a random sample from $U(\theta, 2)$. Find the bias corrected jackknife maximum likelihood estimate of θ. Obtain the jackknife estimate of the variance of the m.l.e. of θ. [8]

Group B: 32 marks

5. (a) Distinguish between seasonal and cyclical components of a time series data.

   (b) What do you mean by a correlogram? Evaluate and draw a correlogram for a first order stationary auoregressive model. [4+2+6]

6. (a) Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution with mean μ and variance $\sigma^2 > 0$. Find the asymptotic distribution of $e^{\bar{X}}$.

   (b) Suppose $X_1, X_2, \ldots, X_n$ are *i.i.d.* $N(\theta, 1)$ variables. Consider the statistic $\Phi\{\frac{\sqrt{n}(c-\bar{X})}{\sqrt{n-1}}\}$ for estimating $\Phi(c - \theta)$ where c is a known constant and Φ is the c.d.f. of a standard normal variable. Examine whether this estimator is consistent and obtain its asymptotic distribution.

   (c) Using variance stabilizing transformation, describe how you would test $H_0 : \rho_1 = \rho_2$ against $H_1 : \rho_1 \neq \rho_2$ where ρ_1 and ρ_2 are the sample correlation coefficients of two independent bivariate samples. Assume that the sample sizes are large. [5+8+7]

Introduction to Markov Chains
B. Stat. IInd Year Semester 2
Indian Statistical Institute

Backpaper Examination

Time: 3 Hours Full Marks: 100

Date: July 10, 2017

This is an OPEN NOTE examination. You may use any handwritten notes that you have brought with you, but you are not allowed to share them. Also no notes in printed, photocopied, cyclostyled or any other form (including electronic storage) is allowed.
Calculators and any other computing devices are strictly not allowed.
You need to justify all your answers and claims.

1. Let $\{X_n\}$ be a Markov chain on the state space S and let $h : S \to S$ be a one-to-one function. Is $\{h(X_n)\}$ a Markov chain? Justify your answer. [7]

2. For any integer $m \geq 2$, let $m = \sum_{i=0}^{k} a_i 10^i$ be its expansion in base 10. (Note that k will depend on m.) Fix $0 < p < 1$ and $p + q = 1$. Then a Markov chain on the set of integers greater than 1 makes transition as follows: From any state m, it goes to $\max(2, \sum_{i=0}^{k} a_i^2)$ with probability p and to 2 with probability q. What is the smallest closed set containing 2? Classify the states into recurrent and transient ones. What are the recurrent classes? [9+6+2=17]

3. Consider a Markov chain $\{X_n\}$ on a finite state space S with probability transition matrix $\boldsymbol{P}$. Fix subset A of S, and define $\tau_A = \inf\{n \geq 0 : X_n \in A\}$. Also denote $t_A(x) = \mathrm{E}_x[\tau_A]$.
   (a) Show that $\mu = t_A$ satisfies the equation
   $$\mu(x) = \begin{cases} 0, & \text{if } x \in A \\ 1 + \sum_{y \in S} p_{xy}\mu(y), & \text{otherwise.} \end{cases} \qquad (\star)$$
   [6]
   (b) Prove that any nonnegative solution q of the equation $(\star)$ satisfies, for any $n \geq 1$ and $x \notin A$, $q(x) \geq n + \mathrm{P}_x[\tau_A = n]$. [6]

4. Consider a Markov chain $\{X_n\}$ on the state space $\{1, 2, 3\}$ with transition probability matrix given by
   $$\begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0 & 0.1 & 0.9 \end{pmatrix}.$$
   (a) Why does this chain have a unique limiting distribution? Obtain the limiting distribution. [3+5=8]
   (b) What is the stationary distribution for the ladder chain (X_n, X_{n+1})? (You may assume and need not prove that the ladder chain is Markov.) [3]
   (c) Define $Y_n = X_n - X_{n-1}$. Is $\{Y_n\}$ a Markov chain? Justify. What is $\lim_{n\to\infty} \mathrm{P}[Y_n = 1]$? [5]
   (d) What is $\lim_{n\to\infty} \mathrm{P}[Y_n = 1]$, if it exists? [6]

5. A fair coin is tossed repeatedly and independently. Find the expected number of tosses required to get the sequence HTH for the first time. [14]
   [Hint: Consider a chain with state space same as all sequences starting with H.]

6. Let $p_1, \ldots, p_n$ be numbers in $(0, 1)$. Define $\lambda_i = -\log(1 - p_i)$ and $\mu = \sum_1^n \lambda_i$. Consider independent Poisson random variables W_i with parameter λ_i and $X_i = \min(W_i, 1)$. Finally define $S_n = \sum_1^n X_i$. Show that the total variation distance of the distribution of S_n from Poisson distribution with parameter μ is bounded by $\frac{1}{2} \sum_1^n \lambda_i^2$. [14]
7. Prove or disprove with justification: For two finite sets S and V, and a probability distribution π on S^V, the Metropolis algorithm and Glauber dynamics will always coincide. [14]

INDIAN STATISTICAL INSTITUTE

Backpaper Examination, 2^{nd} Semester, 2016-17

Statistical Methods IV, B.Stat 2^{nd} Year

Date: 12.07.17 Time: 3 hours

The paper carries 100 marks. Answer all questions.
Use separate answer scripts for each group.

Group 1: 65 marks

1. Consider the bivariate density given by: $f(x,y) = c(x+y-1)^{-p-2}$ where c is a constant and $p > 0$. Stating suitable conditions, obtain the mean of x and the correlation coefficient between x and y. [15]

2. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N_p(\mu, \Sigma)$.
   (a) Obtain an unbiased estimator of $\mu'\Sigma^{-1}\mu$.
   (b) If $\mu = \alpha\mu_0$ and $\Sigma = \sigma^2 I$, where α and σ^2 are unknown constants and μ_0 is known, obtain the maximum likelihood estimators of α and σ^2. Find the M.S.E. of the m.l.e. of α. [10 +10]

3. It is claimed that the weight of the salt (in grams) produced in a chemical reaction increases on an average by 20 grams on increasing the temperature from 80^oC to 90^oC and doubles when the temperature is increased from 90^oC to 100^oC. An investigator who suspects that the claim is false carries out the chemical reaction five times independently at all the three temperatures and obtains the following results:

Experiment	Weight at 80^oC	Weight at 90^oC	Weight at 100^oC
1	238	257	503
2	264	291	590
3	253	278	554
4	227	243	485
5	259	285	560

Based on the above data, do you think that the investigator's suspicion is validated? [20]

4. In an association study on nature of diet and cholesterol levels based on 500 randomly chosen individuals from a population, the following data were observed:

Cholesterol Level	High	Low
Nature of diet		
Vegetarian	137	72
Non-vegetarian	204	87

Do the data provide any evidence of association between nature of diet and cholesterol levels based on Odds Ratio? [10]

Group B: 35 marks

5. (a) Derive a variance stabilising transformation of sample correlation coefficient r for testing $H_0 : \rho = \rho_0$ against $H_1 : \rho \neq \rho_0$ for a bivariate normal distribution and hence, obtain a 95% confidence interval for ρ assuming the sample size to be large.

   (b) Suppose $X_1, \cdots, X_n$ be *i.i.d.* $N(\mu, \sigma^2)$ variables. Find the asymptotic distribution of:

      i. the fourth central sample moment

      ii. the coefficient of variation of the sample

[8 + 7 + 7]

6. (a) While estimating the seasonal component of a time series data, suppose one of the values for a particular season of a year is unusually high. Describe a procedure to calculate seasonal indices in such a scenario.

   (b) Write down the first order autoregressive model. Show that the model is not stationary when the absolute value of the regression parameter is 1. [6+7]

INDIAN STATISTICAL INSTITTUE, KOLKATA

Second Semestral Examination 2016-17 (Back Paper)

B. Stat. (Hons.) II Year

Subject: Economic and Official Statistics and Demography

Date: 13·07·17 Maximum Marks 100 Duration 4 Hours

*(**Instructions**: This question paper has two Groups, A and B. Answer Group A and Group B on separate answer booklets. **Each Group carries a maximum of 50 marks**. Marks allotted to each question are given within parentheses. **The total duration is 4 hours**. Standard notations are followed.)*

Group A: Demography & Economic Statistics

(Answer all questions.)

1 (a) Define/ Explain the following.

i) Late Neo-natal Mortality Rate;

ii) Demographic Dividend;

iii) Parity Progression Ratio;

iv) Maternal Mortality Ratio.

(b) Explain step by step how a life table is constructed from the m-type mortality rates to obtain life expectancy at some age. [1+2+2+1+4 = 10]

2 (a) Express ${}_nL_x$ in terms of l_x , ${}_na_x$ and ${}_nd_x$.

(b) Why q_x cannot be calculated directly? Describe how it gets estimated.

(c) For age groups with width of n years, derive an equation that relates q-type mortality rate with m-type mortality rate. [2 + 1 + 4 + 3 = 10]

3. What is the purpose of El-Badry's procedure? Explain how the purpose is served. All symbols must be defined clearly. [1 + 9 = 10]

5. Define Gross Reproduction Rate (GRR). In what way is it different from Total Fertility Rate (TFR)? How is GRR related to TFR? Define and explain Net Reproduction Rate (NRR). Prepare a self-explanatory tabular format which allows calculation of TFR, GRR and NRR.

[1.5+1.5+1+2+4 = 10]

6. Write down the important properties of the Lorenz Curve with all used symbols fully explained. Show how Lorenz Ratio is related to the Gini Mean Difference.

[5+5 = 10]

Group B: Economic and Official Statistics

[Note: Answers should be to the point and brief. Group B has two parts - I and II.]

Part I: ***Answer Question no. 1 and any two from the rest***

1. Derive the projected demand for the future assuming that the Engel Curve remains invariant over time (use Engel Curve of Double-log form). [9]
2. Derive the profit maximisation conditions of a producing firm (assuming that producers maximise net revenue subject to constraint imposed on production function, existence of perfect competition in factor market and absence of perfect competition in output market). [8]
3. State and explain the types of error faced in constructing an index number. [8]
4. State and explain the difficulties faced in Engel Curve Analysis while combining two types of samples – cross section and time series. [8]

Part II

Write short notes on Question no. 1 and any two from the rest.

1. Consumer Confidence Survey, 2017, conducted by the Reserve Bank of India. [9]
2. Poverty Data published by the World Bank. [8]
3. Social Indicators upheld by the United Nations [8]
4. Methodology of collection of Price Statistics in India. [8]

Indian Statistical Institute
Back Paper Examination 2017
Course Name: BStat Second Year
Subject Name: Economics II (Macroeconomics)

Date 14·07·17 Maximum Marks: 100 Time: 3 Hours

Answer the following questions:

1. The following data regarding a firm in a given year are specified below:

(All figures are in lakhs of rupees)

Revenue earned from the sale of (5/6)th of output (The remaining part was unsold)	600
Raw materials purchased from other firms	100
Unused part of raw materials	20
Interest paid to households	5
Payment made to a labour contractor for supplying labour	5
Travelling and hotel expenses of the officials of the company	1
Land purchased by the company for construction of an additional shed	100
Wages and salaries	2
Depreciation	1
Dividend paid	100
Tax on profit	1
Subsidy received by the firm	2
Donations made to the chief minister's relief fund	1

From the data given above compute the firm's contribution to
(i) NDP
(ii) National Income
(iii) Personal Income
(iv) Private income
(v) Aggregate final expenditure [5x5 = 25]

P.T.O

2.(i) Consider the Simple Keynesian Model for a closed economy with government. State the equilibrium condition of the model and derive the condition for the stability of its equilibrium.

(ii) Suppose in the model given above, at Y = 1000, producers have to sell 20 units from their stock to meet the consumers' demand fully. It is also given that the equilibrium level of Y is 1040. Find out the impact of an increase in autonomous expenditure on the equilibrium level of Y, assuming the expenditure function to be linear. [9+16 = 25]

3.(a) Consider the IS-LM model. Explain how an autonomous shift in the money demand function will affect the equilibrium values of Y and r. Trace the adjustment paths of these two variables.

(b) Consider the IS-LM model where aggregate planned investment is a function of the interest rate alone and $dI/dr = -50$. Now, a shift in the money demand function such that at every (Y,r) demand for real balance increases by 1 unit causes the LM curve to shift horizontally and vertically by -4 units and 0.02 units respectively. Again, all other factors remaining the same, a shift in the consumption function is found to increase the equilibrium level of Y by 500 units. Using the information given above, derive (i) the slope of the LM curve, (ii) the change in the level of aggregate planned investment from the initial to the final equilibrium resulting from the aforesaid shift in the consumption function. [8+17 = 25]

4. Explain the process of money creation initiated by the central bank extending a loan to the government.

[25]