# INDIAN STATISTICAL INSTITUTE

## Mid- Semestral Examination: 2014-15

**Course Name: B. STAT. II YEAR**
**Subject Name: Molecular Biology**

**Date: 02.09.2014** **Maximum Marks: 40** **Duration: 2hrs 30 mins**

**Note: Answer any five out of first seven questions, and any five from Question no. 8 to 14. For Multiple choice questions only first attempted five will be considered. For descriptive questions best five will be considered.**

**Multiple choice questions: Answer any five out of seven questions (Each question carries 2 marks)**

1. If the percent of T in a DNA molecule is 30, what is the percent of G?
   a. 20
   b. 30
   c. 40
   d. 50

2. The RNA contamination in Genomic DNA can be determined by
   a. looking into the absorbance at 260nm
   b. looking into the absorbance at 280nm
   c. looking into the ratio of absorbance at 260nm and 280nm
   d. none of the above

3. DNA replication in prokaryotes and eukaryotes are
   a. exactly similar
   b. only different in the leading strand
   c. only different in the lagging strand
   d. different in several aspects

4. The amino acid, proline, disrupts α-helical structure in proteins because it is
   a. an acidic amino acid
   b. a basic amino acid
   c. an aromatic amino acid
   d. an imino acid

5. In a sequencing reaction, by mistake, ddATP was added in the reaction mixture instead of dATP. What would be the consequence?
   a. No DNA synthesis would occur
   b. Normal DNA synthesis would occur
   c. Synthesis would always stop at the position at which first A was incorporated
   d. Synthesis would terminate randomly regardless of the nucleotide incorporated

6. Which of the following enzymes is involved in nucleic acid synthesis without requiring any template DNA?

a. DNA polymerase 1
b. DNA polymerase γ
c. Poly (A) polymerase
d. RNA polymerase

7. Type II topoisomerases change the linking number by

a. 1
b. 2
c. ±1
d. ±2

**Descriptive questions: Answer any five out of seven questions**
**(Each question carries 6 marks)**

**8. a)** Calculate the length (in nm) of DNA having 64 base pairs that are in (i) B-DNA (ii) Z-DNA and (iii) A-DNA conformation.

**b)** Explain why DNA is negatively charged.

**c)** How does gel electrophoresis separate different DNA molecules?

**9.** a) What are the differences between alpha-helix and beta-sheet protein conformations?

b) Define tertiary structure of a protein?

c) Explain Ramachandran plot.

**10. a)** Establish the Henderson-Hasselbalch equation.

**b)** Define and explain buffer.

**c)** What is the pH when 10 mL of 1 M acetic acid is added to pure water so that the final volume is 100 mL? The $K_a$ for acetic acid is $1.74 \times 10^{-5}$?

**11. a)** Describe major differences between prokaryotic and eukaryotic cells.

**b)** Explain cell cycle checkpoint?

**c)** Briefly explain the regulations that are involved during cell cycle progression.

**12. a)** Explain why DNA replication is Semi-conservative in nature.

**b)**What are the major differences in prokaryotic and eukaryotic replication?

**c)** Briefly explain how telomeres are replicated?

**13. a)** What are the basic steps involved in the polymerase chain reaction (PCR)?

**b)** What reagents do you need to perform a PCR?

**c)** Briefly explain multiplex PCR.

**14. a)** What are the two main methods of DNA sequencing?

**b)** What is Central Dogma?

**c)** What are the major differences in eukaryotic and prokaryotic gene structures?

Mid-Semester Examination 2014

Course – B Stat II Year (2014)

Subject – Economics I (Microeconomics) 02.09.14

Maximum Marks 40 Duration 2.5 Hrs

Answer all the following questions

1.a. Derive and explain the Slutsky equation.

b. (i) Suppose the utility function of a consumer is given by $U = \min\{2x_1 + x_2, x_1 + 2x_2\}$

Derive the optimum solution when $p_1 = 2, p_2 = 3$ and $W = 30$.

(ii) Now, suppose $p_1$ rises to 3. Derive the optimum values of $x_1$ and $x_2$ in the new situation. Hence, describe the change in the optimum value of $x_1$ in terms of the Slutsky equation.

(iii) If $p_1$ had changed to 6 instead of 3, what would have been the maximum and minimum optimum values of $x_1$. [7+(5+7+6)=25]

2. a) Consider the utility maximisation problem of a consumer whose utility function is given by $U = \min\left(\frac{x_1}{10}, x_2\right)$. Prices of the commodities and the budget of the consumer are given by $p_1 = 5, p_2 = 10$ and $W = 100$. Write down the equation of the budget set and show it in a diagram. Draw the indifference curve corresponding to $U = \bar{U}$ in a separate diagram. Derive the equation of the income expansion path and show it in the same diagram as the one showing the indifference curve. Compute the optimum values of $x_1$ and $x_2$.

b. Now, suppose the consumer is given a coupon which is non-transferrable and redeemable for 3 units of commodity 2. Show the new budget set in a diagram and write down its equation. Compute the optimum quantities of $x_1$ and $x_2$.

[8+11=19]

# INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

First semester

B. Stat - Second year 2014-2015

Analysis III

Date: September 3, 2014 | Maximum Marks: 40

Duration: 2 hours

Answer all questions.

For full credit, you have to state the theorems clearly if you use them.

(1) Give brief answers (at most 2 lines) to the following questions: $4 \times 4 = 16$

(a) Show that if the open sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are diffeomorphic then $m = n$.

(b) Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a differentiable function and $\widetilde{f} : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^n$ be the graph of $f$. i.e. $\widetilde{f}(x) = (x, f(x))$. Show that $D\widetilde{f}_x = \widetilde{Df_x}$.

(c) Find a diffeomorphism between the open ball $B(0, R)$ and the (solid) ellipsoid

$$\{x \in \mathbb{R}^n \mid \sum_{i=1}^{n} \frac{x_i^2}{d_i^2} < 1\}$$

where $R, d_i$ are positive real numbers.

(d) Let $f : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$ is given by $f(A) = AA^T$. Find $Df_B$ for some $B \in \mathrm{M}_n(\mathbb{R})$ using the definition of directional derivative and its relation with the total derivative.

(2) Let $r$ be a positive integer. Determine if the sets

$$A = \{X \in \mathrm{M}_n(\mathbb{R}) \mid \mathrm{rank} X \geq r\}, \quad B = \{X \in \mathrm{M}_n(\mathbb{R}) \mid \mathrm{rank} X \leq r\}$$

are open / closed / neither open nor closed. 8

(3) (Notation: $f_x = \frac{\partial f}{\partial x}$, $f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$ etc.) Consider the function

$$g(x, y) = \frac{x^2 - y^2}{x^2 + y^2} \text{ for } (x, y) \neq (0, 0); g(x, y) = 0 \text{ for } (x, y) = (0, 0).$$

Let $f(x, y) = xyg(x, y)$. Find $f_x, f_y$ from first principle, show that they exist everywhere and $f_x = yg + xyg_x$; $f_y = xg + xyg_y$ for $(x, y) \neq (0, 0)$. Using them find $f_{xy}$ and $f_{yx}$ at $(0, 0)$ and show that $f_{xy}(0, 0) \neq f_{yx}(0, 0)$. What is your conclusion from this? 8

(4) Show that the minimum Euclidean distance of a matrix $A \in \mathrm{SL}(2, \mathbb{R}) \subset M_n(\mathbb{R}) \sim \mathbb{R}^{n^2}$ from origin is $\sqrt{2}$. 7

(5) Suppose that for a matrix $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$, there exists an $\epsilon > 0$ such that $|a_{ij}| \leq \epsilon$ for all $1 \leq i, j \leq n$. Show that there is a matrix $X \in M_n(\mathbb{R})$ such that $A = X^2 + X^T$. 6

# INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination : 2014

**B. Stat II year**

**Elements of Algebraic Structures**

Date :04.09.2014 Maximum Marks :60 Duration :2 hrs

**Note**: Answer as many as you can. The maximum you can score is 60. Notation is as used in the class.

1. (a) Let $H$ be a finite subset of a multiplicative group $G$. Show that $H$ is a subgroup of $G$ iff $H$ is closed under multiplication.

   (b) Let $H$ be a subgroup of a group $< G, . >$. Show that for $a, b \in G; Ha = Hb$ iff $ab^{-1} \in H$. [4+5]

2. • Let $n$ be a positive integer. Let

$$Z_n^* = \{a \in Z_n : gcd(a, n) = 1\}.$$

   Show that $Z_n^*$ is a group under multiplication modulo $n$
   Hence deduce Euler's Theorem. [5+2]

3. Prove that if $G$ has no non-trivial subgroups, then $G$ must be finite of prime order [6]

4. • Suppose $a$ is an element of a group $G$ of order $n$. Find the order of $a^i, 1 \leq i \leq n$. How many generators are there in a cyclic group of order $n$? [6]

5. Let $\mathbb{R}$ be the additive group of real numbers and let $\mathbb{Z}$ be the subgroup of integers. Show that $\mathbb{R}/\mathbb{Z}$ is isomorphic to the group of complex numbers on the unit circle under multiplication. [6]

6. Let $G$ be a cyclic group of order $n$ and suppose $G =< a >$. Let $\phi$ be an automorphism of $G$. Show that $\phi(a) = a^t$ for some integer $0 < t < n$ with $gcd(t, n) = 1$. Hence conclude that

$$\mathcal{A}(G) \sim Z_n^*.$$

[7]

7. • Suppose a permutation $\pi \in S_n$ has cycle decomposition $\{n_1, n_2, \ldots, n_k\}$. Show that its sign

$$sg(\pi) = (-1)^{n+k}.$$

• Show that the product of 2 tranpositions can be written as a product of 3-cycles. Hence prove that for $n \geq 3$, the subgroup generated by 3-cycles is $A_n$.

[6+6]

8. • State Sylow's Theorem

• Let $o(G) = pq$, $p$ and $q$ distinct primes and $p < q$. Show that if $p$ does not divide $(q-1)$ then $G$ is cyclic.

• Show that a group of order 12 is not simple. [4+6+6]

INDIAN STATISTICAL INSTITUTE

**Mid-Sem Examination, $1^{st}$ Semester, 2014-15**

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

Date: September 5, 2014 Time: 2 hours

**This paper carries 30 marks. Each question carries 10 marks. Answer all questions.**

1. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution with mean $\mu$ and variance $\sigma^2$. Consider linear functions of the m.l.e. of $\mu^2/\sigma^2$. Obtain an unbiased estimator in the above class of estimators. Examine whether the above estimator is consistent for $\mu^2/\sigma^2$.

2. In order to estimate the prevalence ($p$) of a rare disorder, $k$ families are randommly selected, each having at least one affected individual. Suppose the $i^{th}$ family has $n_i$ individuals, of whom $r_i$ are affected, $i = 1, 2, \ldots, k$. Assuming that individuals within a family manifest the disorder independently of each other, explain how you would use the Fisher's Scoring method to obtain the m.l.e. of $p$. Show all your computational steps clearly.

3. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from an exponential distribution with mean $\lambda$. Consider a test for $H_0 : \lambda \leq \lambda_0$ vs $H_1 : \lambda > \lambda_0$ that rejects $H_0$ iff $\overline{X}$ is greater than a real number $c$. If the size of the test is $\alpha$, determine the value of $c$ in terms of the quantiles of a chi-squares distribution. Give a rough sketch of the power function of the above test. If $n$ is large, show that the power of the test at $\lambda = \lambda_1$ ($> \lambda_0$) can be approximated by $\Phi\{\frac{\sqrt{n}(\lambda_1 - \lambda_0) + \lambda_0 z_\alpha}{\lambda_1}\}$ where $\Phi$ is the c.d.f. and $z_\alpha$ is the $\alpha^{th}$ quantile of a standard normal variable.

# INDIAN STATISTICAL INSTITUTE

## Mid-Semestral Examination: 2014-15 (First Semester)
## Bachelor of Statistics (B. Stat.) II Year
## Probability Theory III

### Teacher: Parthanil Roy

**Date: 06/09/2014** **Maximum Marks: 40** **Duration: 2:30 pm - 5:30 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. Suppose $X \sim Gamma(n, \lambda)$. Compute the characteristic function of $X$. Using this or otherwise, calculate the characteristic function of $Y \sim \chi^2_m$. [4 + 2]

2. Suppose a sequence of random variables $X_n$ with distribution functions $F_n$, $n \geq 1$, converges in distribution to a random variable $X$ with a continuous distribution function $F$. Show that $F_n$ converges to $F$ uniformly on $\mathbb{R}$. In this case, for a sequence of real numbers $x_n \to x$ as $n \to \infty$, compute, with justification, $\lim_{n\to\infty} F_n(x_n)$. [10 + 4]

3. 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) A sequence of random variables $\{X_n\}_{n\geq 1}$ converges in probability to a random variable $X$ if and only if
   $$E\left(\frac{|X_n - X|}{1 + |X_n - X|}\right) \to 0$$
   as $n \to \infty$. [5]

   (b) If $X$ is any random variable and $A$ is an event such that $P(A) = 0$, then $E(XI_A) = 0$. [5]

   (c) If $X$ and $Y$ are two integrable random variables defined on the same probability space such that $E(XI_A) = E(YI_A)$ for all events $A$, then $X = Y$ almost surely. [5]

   (d) If $\{X_n\}_{n\geq 1}$ is a sequence of independent and identically distributed Cauchy random variables, then the sequence of random variables $\{n^{-1/2}(X_1 + X_2 + \cdots + X_n)\}_{n\geq 1}$ is tight. [5]

# INDIAN STATISTICAL INSTITUTE

## Final- Semestral Examination: 2014-15

**Course Name: B. STAT. II YEAR**

**Subject Name: Molecular Biology**

**Date: 03.11.2014** **Maximum Marks: 50** **Duration: 2½ hrs**

**Note: Answer any five from Group A, and any five from Group B. In Group A, only first attempted five answers will be considered. In Group B, best five answers will be considered.**

## Group A

**Multiple choice questions: Answer any five out of seven questions (Each question carries 2 marks)**

1. The enzyme that catalyzes the synthesis of DNA is called
   (a) DNA polymerase
   (b) DNA gyrase
   (c) DNA ligase
   (d) helicase

2. The replication of chromosomes by eukaryotes occurs in a relatively short period of time because
   (a) the eukaryotes have more amount of DNA for replication
   (b) the eukaryotic replication machinery is 1000 times faster than the prokaryotes
   (c) each chromosome contains multiple origin of replications
   (d) eukaryotic DNA is always single stranded

3. Which of the following enzymes unwind short stretches of DNA helix immediately ahead of a replication fork?
   (a) DNA polymerase
   (b) Single-stranded binding proteins
   (c) helicase
   (d) Topoisomerase

4. An allele is
   (a) another word for a gene
   (b) a homozygous genotype
   (c) a heterozygous genotype
   (d) one of the several possible forms of a gene that arise by mutation and are found at the same place on a chromosome.

5. The synthesis of DNA by DNA polymerase occurs in the
   (a) 3' to 5' direction
   (b) 5' to 5' direction
   (c) 3' to 3' direction
   (d) 5' to 3' direction

6. The idea that for any particular trait, the pair of alleles of each parent separate and only one allele from each parent passes to an offspring is Mendel's principle of
   (a) independent assortment
   (b) unit inheritance
   (c) hybridization
   (d) segregation

7. RNA contains which bases?
   (a) adenine, thymine, guanine, cytosine, uracil
   (b) adenine, thymine, guanine, cytosine
   (c) thymine, guanine, cytosine, uracil
   (d) adenine, guanine, cytosine, uracil

## Group B

**Answer any five out of the following seven questions**

**(Each question carries 8 marks)**

**8. (a)** How do prokaryotic ribosomes recognize the 5′ end of the messenger RNA?
**(b)** How do eukaryotic ribosomes recognize the 5′ end of the messenger RNAs?
**(c)** Briefly explain tRNA charging? **[2+2+4]**

**9. (a)** What type of cross would produce the following genetic ratios? (i) 3:1 (ii) 1:1 (iii) 1:2:1 (iv) 9:3:3:1 (v) 2:1
**(b)** Black & white coloration in cows is dominant to brown & white. Black & white cows were crossed over several years and the following progeny were produced: 8 black & white and 3 brown & white. (i) What is the most probable genotype of each parent? (ii) what is the expected genotypic and phenotypic ratio for the progeny? **[5+3]**

**10.** In garden peas, long stems are dominant to short stems, and yellow seeds are dominant to green seeds. 100 long/yellow pea plants, all of which had one short/green parent, are interbred (bred to each other). 1600 progeny result. Answer the following questions about these progeny. **(a)** Assuming that these two genes are unlinked, about how many long/green pea plants would you expect to find among the offspring? **(b)** What ratio of yellow to green seed color would you expect among the offspring? **(c)** What would you expect the overall phenotypic ratio among the 1600 offspring to be (taking into consideration both traits)? **[2+3+3]**

**11.(a)** Why is Mendel's Second Law of Genetics called the Law of Independent Assortment? **(b)** A short-tailed mutant of mouse was discovered. Multiple crosses of this mouse to long-talied mice produced 27 long-talied mice and 25 short-tailed mice. A series of crosses among short-tailed mice were made and 21 short-tailed mice and 11 long-tailed mice were produced? Study these results and determine which phenotype is dominant and explain the ratios observed with regards to the genotypes of the parents in each cross. **[2+6]**

**12.** **(a)** Where do (i) transcription and (ii) translation occur in a eukaryotic cell? **(b)** What are the start signal/sequence for (i) transcription and (ii) translation? **(c)** What is the sequence of the mRNA produced from the DNA sequence given below?
5'-GAGCCATGCATTATCTAGATAGTAGGCTCTGAGAATTTATCTC-3'
**(d)** Distinguish between intron and exon in a eukaryotic DNA. **[2+2+2+2]**

**13.** Explain the process of translation in four stages (i) initiation, (ii) elongation, (iii) translocation and (iv) termination. **[2+2+2+2]**

14. **(a)** In a dihybrid cross, AaBb x AaBb, what fraction of the offspring will be homozygous for both recessive traits? **(b)** In Mendel's experiments, the spherical seed character (SS) is completely dominant over the dented seed character (ss). If the characters for height were incompletely dominant, such that TT is tall, Tt is intermediate and tt is short, what would be the phenotypes resulting from crossing a spherical-seeded, short (SStt) plant to a dented-seeded, tall (ssTT) plant?
**(c)** Suppose you have two rose plants, both with pink flowers. You cross the two plants and are surprised to find that, while most of the offspring are pink, some are red and some are white. (i) You decide that you like the red flowers and would like to make more. What cross would you perform to produce the most red flowered plants? (ii) Your sister decides to have pink flowered roses. Which cross would give you the most pink flowered plants? **[1+3+4]**

**Indian Statistical Institute**
**Semester Examination**
**Course Name: B Stat II Year 2014**
**Subject Name: Economics I (Microeconomics)**

Date: 3.11. 2014

Maximum marks: 60 Duration 2.5 hours

**Answer only three questions.**

**1.** A firm has the option of selling in two different markets A and B segregated by law. In market A, the firm is the only supplier and the market demand function is given by $p^A = 120 - \frac{q^A}{10}$. Market B is perfectly competitive and the firm in question has to sell in market B at a given market price of Rs.80. The government imposes a tax of Rs.15 per unit of production of the firm. The cost function of the firm (excluding the tax) is given by $C(Q) = 50Q + \frac{Q^2}{20}$, where $Q$ denotes the total output of the firm. Will the firm sell in both the markets? Explain your answer. [20]

**2.** Suppose there are two firms, firm 1 and firm 2, in an oligopoly market. They sell a differentiated product. The demand functions of the two firms are given by $p_1 = 100 - 2q_1 - q_2$ and $p_2 = 95 - q_1 - 3q_2$, while the cost functions are $C_1 = \frac{5}{2}q_1^2$ and $C_2 = 25q_2$. The price and quantity combinations $p_1 = 70, p_2 = 55, q_1 = q_2 = 10$ were found to prevail for quite some time.

(i) Derive the kinked oligopoly demand curve for firm 2 and indicate the kink point as explained in the Sweezy model.

(ii) Using this demand function, explain why firm 2 will have no incentive to change its initial price-quantity combination despite changes in the cost of production within certain limits. [10+10=20]

**3. a.** Consider the utility maximisation problem of a consumer whose utility function is given by $U = \min\left\{\frac{x_1}{10}, x_2\right\}$. Prices of the commodities and the budget of the consumer are given by $p_1 = 5, p_2 = 10$ and $W = 100$ respectively. Write down the equation of the budget set and show it in a diagram. Draw the indifference curve corresponding to $U = \overline{U}$ in a separate diagram. Derive the equation of the income expansion path and show it in the same diagram

P. T. O

as the one showing the indifference curve. Compute the optimum values of $x_1$ and $x_2$ and explain the consumer's choice.

b. Now, suppose the consumer is additionally given a coupon which is non-transferable and redeemable for 3 units of commodity 2. Show the new budget set in a diagram and write down its equation. Compute the new optimum quantities of $x_1$ and $x_2$. [10+10=20]

4. a. Consider a Cournot model with two firms selling a homogeneous product. The market demand and cost functions are given by $Q = 120 - p$ and $C_i = 10q_i$ i = 1,2 respectively. Find out the Cournot solution. Also compute the profit earned by each firm.

b. Find out what the Stackelberg solution will be with the same demand and cost functions specified above, when firm 1 acts as the leader and firm 2 is the follower.

c. Do you think that the profit earned by firm 1 in the Stackleberg solution in general can never be less than that in the Cournot solution? Explain your answer. [7+7+6=20]

# INDIAN STATISTICAL INSTITUTE

Semestral Examination

First semester

B. Stat - Second year 2014-2015

Analysis III

Date: November 7, 2014

Maximum Marks: 60

Duration: 3 hours

Answer all questions.

For full credit, you have to state any theorems/results you use.

(1) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a compactly supported continuous function and $g : \mathbb{R} \to \mathbb{R}$ be continuous. Define $h(y) = \int_{\mathbb{R}^n} f(x)g(\langle x, y\rangle)dx$.

(a) Suppose that $f$ is a homogeneous function of degree $k > 0$ i.e. $f(rx) = r^k f(x)$ for any $r > 0$, $x \in \mathbb{R}^n$. Show that $h$ is also homogeneous and find its degree of homogeneity.

(b) Suppose $f$ satisfies $f(Ax) = f(x)$ for all $A \in O(n)$ and $g$ is an odd function. Show that $h(y) = 0$ for all $y \in \mathbb{R}^n$.

(c) Show that the only function $f$ on $\mathbb{R}^n$ to $\mathbb{R}$ which is both radial and homogeneous is $C|x|^\alpha$ for some $\alpha \in \mathbb{R}$ and a real constant $C$.

$5 \times 3$

(2) Let $\phi(x, y)$ be a $C^2$-function on $\mathbb{R}^2 \setminus \{(0,0)\}$ which satisfies the relation $\phi_{xx} + \phi_{yy} + \phi_x = 0$ on $\mathbb{R}^2 \setminus \{(0,0)\}$. Let $\Gamma_R$ be the circle $x^2 + y^2 = R^2$. Show that the integral

$$\int_{\Gamma_R} e^x(\phi_x dy - \phi_y dx)$$

does not depend on the radius $R$. 8

**Hint:** Take two different radii $R$ and $r$ and change it to integration over annulus.

(3) Let $f : \mathbb{R} \to \mathbb{R}^n$ be a $C^1$-function which satisfies $f'(t) = Af(t)$, where $A$ is an $n \times n$ matrix and $A + A^T$ is positive definite. Show that the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(t) = \|f(t)\|^2$ is an increasing function for $t > 0$. 7

**Hint:** Write $g$ using inner product. Find $g'(t)$ and use $f'(t) = Af(t)$.

P. T. O.

(4) Let $E$ be an open convex set of $\mathbb{R}^n$ and $w$ be a closed 1-form on $E$. Show that $w$ is exact. Find $f$ such that $df = w$. Using this algorithm find a real valued function $\Phi$ on $\mathbb{R}^2$ such that,

$$d\Phi = (1+y)\cos x dx + (\sin x + y)dy.$$

8+7

(5) Consider a 2-surface $S$ in $\mathbb{R}^3$ parametrized as:

$$(\theta, \phi) \mapsto ((2+\cos 2\pi\theta)\cos 2\pi\phi, \sin 2\pi\theta, (2+\cos 2\pi\theta)\sin 2\pi\phi),$$

for $(\theta, \phi) \in [0,1] \times [0,1]$. Show that for any exact 2-form $w$, $\int_S w = 0$. 7

(6) Let $C$ be a right circular cone of height $L$ in $\mathbb{R}^3$ with its axis along $z$-axis, and its base is a circle of radius $r$ in the $xy$-plane.

(a) Express the surface $S$ of $C$ as a 2-surface, i.e. a map from $[0,1]^2$ to $\mathbb{R}^3$.

(b) Find the boundary $\partial S$ of $S$.

(c) Let $w$ be a 1-form of class $C^2$ on $S$. Show that $\int_S dw = \int_{\mathbb{D}} dw$ where

$$D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}.$$

3+5+4

**INDIAN STATISTICAL INSTITUTE**

**Final Examination, $1^{st}$ Semester, 2014-15**

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

Date: November 10, 2014 Time: 3 hours

**This is an open notes examination. The paper carries 70 marks.**

1. (a) 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 $x_2 = x_3 = \ldots = x_{10} \neq x_1$. Under what conditions will the least squares regression line of $Y$ on $X$ be identical to the least absolute deviation line?

   (b) Suppose $X_1, X_2, X_3, X_4$ is a random sample from an exponential distribution with mean $\lambda$. Show that the variance of any unbiased estimator of $1/\lambda^2$ cannot be less than $1/\lambda^4$.

   (c) Suppose $Z_1, Z_2, \ldots, Z_n$ is a random sample from $N(0, \sigma^2)$. What is the distribution of the statistic $\overline{Z}^2 / \frac{1}{n} Z_i{}^2$? [6 + 4 + 5]

2. (a) Suppose $X_1, X_2, \ldots, X_{10}$ is a random sample from the density $f(x) = \theta/x^2, x \geq \theta$. Consider estimators of the form $\alpha\widehat{\theta}$, where $\alpha$ is a real number and $\widehat{\theta}$ is the m.l.e. of $\theta$. Which estimator in the above class has the minimum mean squared error?

   (b) Consider a set of bivariate observations $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, *)\}$ where * denotes a missing value and $x_n$ is the mean of $x_1, x_2, \ldots, x_{n-1}$. Suppose we pretend that the above set of observations represents a random sample from a bivariate normal distribution and our aim is to obtain the maximum likelihood estimates of the different parameters. Show that an EM algorithm that uses the maximum likelihood estimates of the parameters based on the first $(n-1)$ observations as initial values converges in two cycles of iteration. [7+8]

3. (a) Assume that the conditional distribution of the number of pea plants ($Y$) given the area of a field ($X$) is Poisson with mean $\beta X$. In a survey on $n$ independent fields, it was found that the area of the $i^{th}$ field is $x_i$ and the number of pea plants in it is

$y_i$, $i = 1, 2, \ldots, n$. Show that the likelihood ratio test for $H_0 : \beta = \beta_0$ vs $H_1 : \beta \neq \beta_0$ based on the above sample rejects $H_0$ iff $\overline{y}\{log_e\, \overline{y} - log_e\, \beta_0\overline{x}\} - (\overline{y} - \beta_0\overline{x})$ is sufficiently large.

(b) The Council of Scientific and Industrial Research (CSIR) organizes the National Eligibility Test (NET) every six months for entitlement to fellowships for carrying out graduate research in Indian universities. If two students required 4 and 6 attempts, respectively, to be succesful in NET, develop a suitable test with level 0.05 to conclude whether the success rates of the two students are same or not. State your assumptions and show all computations clearly. [7 + 8]

4. (a) Suppose we want to estimate the difference in the proportions of male and female airline passengers who prefer window seats to aisle seats using an asymptotic 99% equal tail confidence interval. What is the minimum <u>combined</u> sample size required such that the error in estimation based on the confidence interval is at most 0.05?

(b) Consider data on two binary variables $X$ and $Y$. Show that testing whether the slope coefficient of a logistic regression model of $Y$ on $X$ is 0 or not is equivalent to testing whether the Odds Ratio based on the above model is 1 or not. [6 + 4]

*Class Tests and Quizes carry 15 marks*

# INDIAN STATISTICAL INSTITUTE

## Semestral Examination: 2014-15 (First Semester)
## Bachelor of Statistics (B. Stat.) II Year
## Probability Theory III

### Teacher: Parthanil Roy

**Date: 12/11/2014** **Maximum Marks: 60** **Duration: 10:30 pm - 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.**
- **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. Let $\{a_n\}_{n\geq 1}$ be a sequence of numbers in $(0,1)$ such that $a_n \to 0$, but $\sum_{n\geq 1} a_n = \infty$. Suppose $X_1, X_2, \ldots$ are independent random variables with $P(X_n = 1) = P(X_n = -1) = a_n/2$ and $P(X_n = 0) = 1 - a_n$ for all $n \geq 1$. Define $Y_1 = X_1$ and for all $n \geq 2$,

$$Y_n = \begin{cases} X_n & \text{if } Y_{n-1} = 0, \\ nY_{n-1}|X_n| & \text{if } Y_{n-1} \neq 0. \end{cases}$$

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) $Y_n$ converges in probability. [6]

   (b) $Y_n$ does not converge almost surely. [6]

2. Suppose that $X$ is an integrable random variable and $\varphi(t)$ is its characteristic function. Show that $\varphi(t)$ is a differentiable function and for all $t \in \mathbb{R}$, $\varphi'(t) = iE(Xe^{itX})$. [12]

3. Suppose $U$ and $V$ are two independent and identically distributed random variables following a uniform distribution on $(0,1)$. Compute the conditional distribution of $U$ given $Z := \max(U, V)$. [12]

**[P. T. O]**

4. Suppose $\{X_n\}_{n\geq 1}$ is a sequence of random variables converging to $X$ almost surely and $E|X_n| \leq 23$ 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 $E|X| \leq 23$. [6]

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

5. For any nonzero column vector $\mathbf{x} \in \mathbb{R}^k$, let $\tilde{\mathbf{x}} = \left(\sqrt{\mathbf{x}^T\mathbf{x}}\right)^{-1} \mathbf{x}$. If $\mathbf{x} = \mathbf{0} \in \mathbb{R}^k$, then define $\tilde{\mathbf{x}}$ to be $\mathbf{1}$, the column vector of all ones in $\mathbb{R}^k$. For any random vector $\mathbf{X} \sim N(\mathbf{0}, I_k)$ and any symmetric idempotent $k \times k$ matrix $A$ with $1 \leq \text{trace}(A) < k$, find the distribution of $\widetilde{\mathbf{X}}^T A \widetilde{\mathbf{X}}$. You have to justify all the steps of your calculation. [12]

# INDIAN STATISTICAL INSTITUTE

Semester Examination : 2014( First Semester)

**B. Stat. II year**

**Elements of Algebraic Structures**

Date :–14.11.14 Maximum Marks :90 Duration :3 hrs

Note: Answer as many as you can. The maximun you can score is 90. Notation is as used in the class.

1. (a) Let $p$ be an odd prime. An integer $a \in \mathbf{Z}_p^*$ is called a *quadratic residue* modulo $p$ if there is an integer $x$ such that $x^2 \equiv a \bmod p$. Let QR(p) denote the set of quadratic residues modulo $p$. Show that QR(p) is a group of order $(p-1)/2$ under multiplication modulo $p$. Show that if $a \notin QR(p)$ and $b \in QR(p)$ then $ab \notin QR(p)$. Also show that, for $a \in \mathbf{Z}_p^*$, $a \in QR(p)$ iff

$$a^{(p-1)/2} \equiv 1 \bmod p.$$

   (b) Write down the Euclidean Algorithm for finding the GCD of two integers $a, b$. Prove its correctness. [10+10]

2. State the Chinese Remainder Theorem(for 2 congruence relations). Show that if $n, m$ are positive integers such that $GCD(n, m) = 1$, then

$$\phi(nm) = \phi(n)\phi(m),$$

   where $\phi$ is the Euler's function.
   (Hint: Use CRT.) [7]

3. Show that every ideal $I$ in the ring $F[x]$, where $F$ is a field, is generated by a single polynomial i.e. $I = < p(x) >$ for some $p(x) \in F[x]$.
   Show that if $p(x)$ is irreducible, then $I$ is maximal. [7+8]

4. Prove that $x^2 + 1$ is irreducible over $\mathbf{Z}_7$.
   Hence, or otherwise, construct a field with 49 elements, showing how the elements of the field are represented and how addition and multiplication are defined. [8]

5. • Show that the Gausssian integers $\mathbf{Z}[i]$ form a Euclidean ring.

• Find the units of $\mathbf{Z}[i]$.

• Find the GCD of $11 + 7i$ and $18 - i$ in $\mathbf{Z}[i]$.

[8+3+7]

6. • Let $F$ be a field and $K$ an extension of $F$. When is an element $a \in K$ said to be algebraic over $F$?
Show that the set of algebraic elements forms a subfield of $K$.

• Find an element $u \in I\!R$ such tha $\mathbf{Q}(\sqrt{2}, \sqrt[3]{5}) = \mathbf{Q}(u)$.

• Show that if $\alpha$ is an algebraic integer and $m \in \mathbf{Z}$ then $\alpha + m$ is an algebraic integer.

• Let $f(x) \in F[x]$ be an irreducible polynomial. Show that if the characteristic of $F$ is $p \neq 0$, then $f(x)$ has a multiple root only if it is of the form $f(x) = g(x^p)$.

[7+7+6+7]

# INDIAN STATISTICAL INSTITUTE

## Backpaper Examination: 2014-15 (First Semester)
## Bachelor of Statistics (B. Stat.) II Year
## Probability Theory III

**Teacher: Parthanil Roy**

Date: 26/12/2014 **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.**
- **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. 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 $\mu$ 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 $\phi$ denote the probability density function of standard normal distribution and $\Phi$ denote its cumulative distribution function. Prove that $\lim_{x\to\infty} \dfrac{1-\Phi(x)}{\phi(x)/x} = 1$. [5]

   (b) If $\{X_n\}$ is a sequence of independent and identically distributed standard normal random variables, show that $P\left[\limsup_{n\to\infty} \dfrac{|X_n|}{\sqrt{\log n}} = \sqrt{2}\right] = 1$. [20]

# INDIAN STATISTICAL INSTITUTE

Back Paper Examination : 2014( First Semester)

**B. Stat. II**

**Elements of Algebraic Structures**

Date 30/12/14 Maximum Marks :100 Duration :3 hrs

Note: Answer as many as you can. The maximun you can score is 100. Notation is as used in the class.

1. (a) Define a cyclic group. Show that a subgroup of a cyclic group is cyclic.

   (b) Condider the additive group of integers $\mathbf{Z}$. Let $H$ be a non-trivial subgroup of $\mathbf{Z}$. Show that there exits a positive integer $d$ such tha $H = d\mathbf{Z} = \{dn : n \in \mathbf{Z}\}$.

   (c) Let $G_1, G_2$ be two multiplicative groups. Consider the cartesian product $G_1 \times G_2$. For $(x_1, y_1), (x_2, y_2) \in G_1 \times G_2$ define their product to be $(x_1x_2, y_1y_2)$. Show that $G_1 \times G_2$ is a group. This is called the **direct product** of $G_1$ and $G_2$.

   Show that the direct product of two cyclic groups need not be cyclic.

[6+6+(5+6)]

2. When is a subgroup $H$ of $G$ said to be normal? Let $H$ be a normal subgroup of $G$. Define the quotient group $G/H$.

   Show that if $G$ is finite, then $|G/H| = |G|/|H|$.

   Let $f : G \to G'$ be a homomorphism of the group $G$ *onto* $G'$. Show that the kernel $ker(f)$ is normal. Prove that $G/ker(f)$ is isomorphic to $G'$.

[6+5+10]

3. Show that every permutation of a finite set can be expressed as a product of transpositions.

Express the following permutation as a product of disjoint cycles.

$$\pi(1) = 3, \pi(2) = 5, \pi(3) = 1, \pi(4) = 2, \pi(5) = 4.$$

Find the order of the permutation.

If $\pi$ is an $r$-cycle, then what is its sign? Justify. [8+6+6]

4. (a) For finite groups, write down the *class equation* or *class formula*. Explain all notation.

   (b) Use the class equation to show that if a group $G$ has order $p^2$, where $p$ is prime, the $G$ is commutative.

   (c) Define an ideal in a ring $R$. When is it said to be maximal? Show that if $M$ is a maximal ideal in a commutative ring $R$ then the quotient ring $R/M$ is a field.

   Hence, or otherwise, show that $\mathbf{Z}_p$, $p$ prime, is a field.

   (d) Show that $p\mathbf{Z}$ is a prime ideal in the ring of integers $\mathbf{Z}$, where $p$ is prime.

[5+7+(8+5)+6]

5. Show that if $p(x)$ is an irreducible polynomial in $F[x]$ of degree $n \geq 1$, then there is an extension $K$ of $F$, such that $[K : F] = n$, in which $p(x)$ has a root. [12]

# INDIAN STATISTICAL INSTITUTE

**Backpaper Examination, $1^{st}$ Semester, 2014-15**

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

Date: 01/01/2015 Time: 3 hours

**This paper carries 100 marks. Answer all questions.**

1. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $U(\theta_1, \theta_2)$. Suppose $\theta = \theta_2 - \theta_1$.

   (a) Obtain a method of moments estimator of $\theta$ based on the above sample.

   (b) Consider estimators of the form $k\widehat{\theta}$, where $k$ is a real number and $\widehat{\theta}$ is the m.l.e. of $\theta$.. Which estimator in the above class has the minimum mean squared error?

   (c) Examine if $\widehat{\theta}$ is a consistent estimator of $\theta$. [6 + 12 + 7]

2. 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 $\rho$. Show all computational steps clearly with one cycle of iteration. [20]

3. Consider the regression model:

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

   where, $x_i$s are fixed and assume values 1, 0 or -1; while $e_i$s are random errors with mean 0 and the same variance. Consider two possible distributions of the errors: normal and double exponential. Explain whether it is possible that the maximum likelihood estimates of $\beta$ are identical under both error distributions. [15]

4. (a) Suppose $X_1, X_2, \ldots, X_5$ is a random sample from a Poisson distribution with mean $\theta$. Show that the variance of the sample variance is greater than $0.128\ \theta$.

(b) Suppose $X_1, X_2, \ldots, X_n$ is a random sample from an exponential distribution with mean $\lambda$. Consider a test of size $\alpha$ for $H_0 : \lambda \leq \lambda_0$ vs $H_1 : \lambda > \lambda_0$ that rejects $H_0$ iff $max\{X_1, X_2, \ldots, X_n\}$ is sufficiently large. Give a rough sketch of the power function of the above test. What is the minimum sample size required such the power of the test at $\lambda = \lambda_1$ ($> \lambda_0$) is at least $1 - \beta$? [8 + 12]

5. (a) Suppose $X_1, X_2, \ldots, X_7$ is a random sample from $N(\mu_1 . {\sigma_1}^2)$ and $Y_1, Y_2, \ldots, Y_8$ is a random sample from $N(\mu_2, {\sigma_2}^2)$, independent of the first sample. Based on the above samples, the maximum likelihood estimates of $\mu_1, \mu_2, \sigma_1$ and $\sigma_2$ were obtained as 4.83, 3.31, 1.39 and 1.28, respectively. Stating your assumptions clearly, obtain a 95% equal tail confidence interval for $(\mu_1 - \mu_2)$. Based on your confidence interval, would you conclude that the means of the two populations are equal?

(b) The government claims that at least 75% of the population support a particular legislation. In a survey on 10000 randomly chosen individuals from the population, 6852 individuals responded that they support the legislation. Based on the above data, would you conclude that the government's claim is justified? State your assumptions clearly. [12 + 8]

# INDIAN STATISTICAL INSTITUTE

Backpaper Examination

First semester

B. Stat - Second year 2014-2015

Analysis III

Date: 02.01.2015 ~~2014~~

Maximum Marks: 100

Duration: 3 hours

Answer all questions.

You must state clearly any result/theorem you use.

(1) Let $X \in \mathrm{GL}(n, \mathbb{R})$. Show that the ball $B_R(X) \subset \mathrm{GL}(n, R)$ when the radius $R = \frac{1}{2\|X^{-1}\|}$. The norm $\|\cdot\|$ used here is operator norm. 7

(2) Verify compactness/non-compactness of these sets:

$$S^1 \times S^1 \subset \mathbb{R}^2,\ SO(2,\mathbb{C}) = \{A \in M_2(\mathbb{C}) \sim \mathbb{R}^4 \mid AA^T = I\},\ SL(2,\mathbb{R}).$$

9

(3) On $\mathbb{R}^n$ define a new distance $d(x,y) = \frac{\|x-y\|}{1+\|x-y\|}$. Show that it is complete. 7

(4) (a) Let $f : \mathrm{GL}(n,\mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$ be given by $f(X) = X^T$. Find $Df_A$ for $A \in \mathrm{GL}(n,\mathbb{R})$.

(b) Let $g : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ be given by $g(X) = X^4$. Find $Dg_A$ for $A \in \mathrm{GL}(n,\mathbb{R})$.

$4 \times 2 = 8$

(5) Let $F : \mathbb{R} \to M_2(\mathbb{R})$ given by $F(t) = (f_{ij}(t))$ be a differentiable function. Find the derivative of the map $t \mapsto \mathrm{Trace}[F(t)^3]$ in terms of $F(t)$ and $F'(t)$. 8

(6) $F : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ is given by $F(X) = X + X^2$. Show that the range of $F$ contains an open neighbourhood of $O$. 9

P. T. O.

(7) Let $\mathfrak{so}(n)$ be the set of $n \times n$ real skew symmetric matrices.

(a) Show that if $A \in \mathfrak{so}(n)$, then $e^A \in SO(n)$. You may assume that $A \mapsto e^A$ is continuous or use the formula $\det e^A = e^{Trace\, A}$.

(b) Let $T_eSO(n)$ be the tangent space of $SO(n)$ at identity $e$. Show that $T_eSO(n)$ is $\mathfrak{so}(n)$.

(c) Show that for any point $p \in SO(n)$, the tangent space at $p$, $T_pSO(n)$ is a vector space over $\mathbb{R}$ and is isomorphic to $T_eSO(n)$.

$5 \times 3$

(8) Find the critical points of the function $f(x, y, z) = x^2 + y^2 + z^2$ on the surface $z = xy + 1$. Is it a local maxium/minimum? 8

(9) Find $\int_S e^{x+y} dxdy$ where $S = \{(x, y) \in \mathbb{R}^2 \mid |x| + |y| \leq 1, xy \geq 0\}$. 9

(10) Let $C$ be the right circular cone in $\mathbb{R}^3$ specified by the following data. Height of $C$ is $L$, vertex of $C$ is at origin, axis of $C$ is the $z$-axis and its base is a disc of radius $r$ in the plane $z = L$. Verify that $C$ is a 3-surface in $\mathbb{R}^3$ and find its volume. 10

(11) Let $A$ be the region in $\mathbb{R}^2$ defined by

$$A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \geq r, x^2 + y^2 \leq R\}$$

where $R > r > 0$. Prove that

$$\int_A dxdy = \int_{\gamma_r} xdy - \int_{\gamma_R} xdy = \int_{\gamma_R} ydx - \int_{\gamma_r} ydx$$

through the following steps.

(a) Take $\gamma_r : t \mapsto (r\cos 2\pi t, r\sin 2\pi t)$ and $\gamma_R : t \mapsto (R\cos 2\pi t, R\sin 2\pi t)$, $0 \leq t \leq 1$.

(b) Find $\Phi(u, t) : [0, 1] \times [0, 1] \to A \subset \mathbb{R}^2$ such that $\Phi(0, t) = \gamma_r(t)$, $\Phi(1, t) = \gamma_R(t)$.

(c) Show that the Jacobian determinant $J_\phi > 0$, i.e. $A$ is positively oriented.

(d) Find $\partial A$.

(e) State and use Stokes' theorem.

10

# INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : 2014-15( Second Semester)

**B. Stat. II year**

**Discrete Maths.**

Date: 23.02.15 Maximum Marks : 80 Duration : 2 1/2 hrs

Note: Answer as many as you can. The maximum you can score is 80. Notation is as used in the class.

1. Suppose $n+1$ are chosen from the set $\{1, 2, \ldots, 2n\}$. Show that thre exists a pair $a, b$ such that

$$(i) gcd(a, b) = 1$$

$$(ii) a|b \text{ or } b|a$$

[5+5]

2. Show that in any sequence of 111 real numbers, there is a monotonically increasing subsequence of length 11 or a monotonically decreasing subsequence of length 12.
[7]

3. State the Inclusion-Exclusion Principle.
Hence, find $\phi(n)$, where $n = p^2q^3r^4$ , $p, q, r$ are three distinct primes and $\phi$ is the Euler's function. [4+6]

4. Show thaty the expected number of fixed points of a random permutation drawn from $S_n$ is 1. [7]

5. (a) Let $G(x)$ denote the generating function for the Fibonacci sequence $\{F_n\}$. Show that $G(x) = \frac{1}{1-x-x^2}$.
Suppose $G(x) = \frac{1}{(1-\lambda x)(1-\mu x)}$. What are the values of $\lambda, \mu$? Hence, derive a formula for $F_n$.

   (b) Let $S_{n,t}$ denote the number of ways to partition an $n$-element set into $t$ non-empty subsets. Show that

$$S_{n,t} = tS_{n-1,t} + S_{n-1,t-1}, 1 \leq t < n.$$

   Compute $S_{5,3}$

(c) i. Show that for $|x| < 1$,

$$(1-x)(1+x)(1+x^2)\dots(1+x^{2^k})\dots = 1.$$

Hence, deduce that for $|x| < 1$,

$$1 + x + x^2 + x^3 + \dots = (1+x)(1+x^2)\dots(1+x^{2^k})\dots$$

ii. Conclude that any integer can be written uniquely in the binary form.

[8+6+8]

6. Suppose $\{a_n\}$ satisfy the following recurrence, for $n \geq 2$

$$na_n = 2(a_{n-1} + a_{n-2}),$$

with $a_0 = e$ and $a_1 = 2e$. Let $A(x)$ denote the ordinary generating function for $\{a_n\}$. Show that

$$A'(x) = 2(1+x)A(x).$$

Find $A(x)$ [8]

7. (a) Consider the recurrence

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_p a_{n-p}, \quad n \geq p,$$

where $a_i \in I\!\!R$ and $a_p \neq 0$. Show that the set of solutions of the recurrence relation forms a vector space over $I\!\!R$ What is its dimension?

(b) A codeword over the alphabet is said to be legitimate if no two 0's appear consecutively. Let $b_n$ denote the number of legitimate codewords of length $n$. Find a recurrence for the number $b_n$.

[6+4]

8. Define the Stirling's numbers $s(n,k)$.
Derive the generating function for the (unsigned) Stirling numbers $|s(n,k)|$.

[4+7]

9. Let $p(n, k)$ denote the number of partitions of $n$ into (exactly) $k$ parts. Derive the following recurrence relation

$$p(n, k) = p(n-1, k-1) + p(n-k, k).$$

[5]

**INDIAN STATISTICAL INSTITUTE**
MID-SEM Examination : (2014 – 2015)
B. Stat. II Year
AGRICULTURAL SCIENCE

Date : 24/02/15 Maximum Marks : 30 Duration :Two hours

**(Attempt any three questions)**

**( Number of copies of the question paper required : 6 )**

1. Define onset and cessation of southwest monsoon? Draw a suitable rice calendar with the following data. 2+8

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***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.*** | ***5*** | ***17*** | ***25*** | ***30*** | ***49*** | ***37*** | ***28*** | ***35*** | ***18*** | ***20*** | ***98*** | ***142*** | ***95*** | ***80*** | ***15*** | ***32*** | ***22*** | ***0*** | ***0*** |
| ***PET (mm)*** | ***35*** | ***32*** | ***30*** | ***27*** | ***33*** | ***25*** | ***22*** | ***20*** | ***22*** | ***20*** | ***19*** | ***21*** | ***24*** | ***25*** | ***35*** | ***28*** | ***31*** | ***33*** | ***34*** |

2. Write in brief about different types of rice. Briefly describe the cultural practices associated with System of Rice Intensification. 4+6

3. Name different meteorological variables that are related to crop production. Write the names of apparatus used to measure those variables. Why the ambient humidity is considered as "RELATIVE HUMIDITY" 4+4+2

4. Write short notes on any two
   a) PET b) Dapog method d) CGR

   2 X 5

**Indian Statistical Institute**

**Mid-Semester Examination 2015**

**Course Name: BStat Second Year**

**Subject: Economics II (Macroeconomics)**

**Date** 24·02·15 **Maximum Marks: 40** **Duration: 1.5 Hours**

**Answer the following questions**

**1.** The following data regarding a firm in a given year are specified below. All figures are in crores of rupees. The firm in the given year sold three-fourths of its output to earn its revenue.

| | |
|---|---|
| Revenue earned | 300 |
| Raw materials purchased from other firms | 20 |
| Unused part of the raw materials purchased | 5 |
| Interest paid on bank loans | 5 |
| Payment made to a labour contractor for supplying labour | 10 |
| Payments made to a car rental company for hiring cars for company officials | 1 |
| Wages and salaries | 20 |
| Indirect taxes | 2 |
| Capital goods purchased by the firm | 20 |
| Depreciation | 1 |
| Dividend paid | 5 |
| Tax paid on profit | 1 |
| Subsidy received by the firm | 10 |
| Donation to a political party | 1 |
| Purchase of a minister's dress in an auction | 2 |

From the data given above compute the firm's contribution to the economy's

**(i) GDP**

**(ii) National Income**

**(iii) Personal income** , Private income

**(iv) Aggregate final expenditure** [8 x 4 = 32]

**P.T.O.**

**2.** Consider the following information regarding an economy in a given year: (All figures are in crores of rupees)

| | |
|---|---|
| Government administration and defence's purchases from other firms | 60 |
| Wage bill in government administration and defence | 20 |
| Government Budget Deficit | 10 |
| NDP | 1000 |
| Net Factor Income from Abroad | 10 |
| Net Foreign Transfers | 2 |

**Derive private disposable income.** **[8]**

# INDIAN STATISTICAL INSTITUTE

Mid-semester Examination: 2014-2015

B. Stat. (Hons.) 2nd Year. 2nd Semester

Statistical Methods IV

Date: February 25, 2014 | Maximum Marks: 30 | Duration: 1 and 1/2 hours

---

- Answer all the questions.

- You should provide as much details as possible while answering a question. You must state clearly any result stated (and proved) in class you may need in order to answer a particular question.

---

1. Suppose $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\mathbf{A}$ be a $p \times p$ symmetric matrix and let $Q \stackrel{def}{=} \boldsymbol{X}^{\mathrm{T}}\mathbf{A}\boldsymbol{X}$. Also, let $\boldsymbol{Y} \stackrel{def}{=} \boldsymbol{X} - \boldsymbol{\mu}$. Show the following.

   (a) $\mathrm{Cov}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{Y}, \boldsymbol{Y}^{\mathrm{T}}\mathbf{A}\boldsymbol{Y}) = 0$, for any fixed $\underset{\sim}{a} \in \mathbb{R}^p$. [3]

   (b) $\mathrm{Var}(Q) = 2\,\mathrm{trace}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}^{\mathrm{T}}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu}$. [11]

2. Consider $k$ samples consisting of $n_i$, $i = 1, \ldots, k$, $p$-variate observations from $k$ populations. Let $n := n_1 + \cdots + n_k$.

   (a) State and prove a result which enables us to decompose the overall variability of all $n$ observations into between- and within-group variabilities. [5]

   (b) Suppose now that the $n_i$ observations of the $i$-th sample constitute a random sample from $\mathrm{N}_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ distribution and that all the $n$ observations are independent. Assume that $\boldsymbol{\Sigma}$ is an unknown positive definite matrix.

   Use (a) to develop a test for the hypothesis that all the $\boldsymbol{\mu}_i$'s are equal. You must state explicitly the test statistic, indicate the critical region and mention the null distribution in terms of distributions we have discussed in class. [11]

* * * * ** *Best of Luck!* * * * * **

Indian Statistical Institute

MID-SEM EXAM 2014

26.02.2015

Course Name BSTAT 2nd Year

Subject : Economic and Official Statistics and Demography

Maximum Marks 20 Time 1 hour

Answer Question 1 and any two from Question2 : 10+(5+5)=20

1. What do you understand by Runs of Index numbers? Derive the Divisia Price Index number and state its advantages : 2+7+1=10

2. Write short notes (on any two) with the following restrictions : 5*2=10

a) Survey Rounds of NSSO **(not for Group B students)**

b) Database of Reserve Bank of India **(not for Group C students)**

c) Index numbers constructed and maintained by Labour Bureau, Shimla **(not for Group A students)**

D) Structure of the Sample Registration System of Census of India **(not for Group E students)**

E) State Statistical Plan (SSP) under ISSP **(not for Group D students)**

G) World Bank Development Indicators for Environment **(not for group F students)**

# INDIAN STATISTICAL INSTITUTE

Mid-semester Examination: 2014-2015

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

***Subject Name : Economic and Official Statistics and Demography***

*(Answer Group A and Group B on two separate answer booklets)*

26/02/2015

## Group B

Maximum Marks: 40 Time : 1.5 Hours

1. What do you understand by neo-Malthusian thinking. Explain briefly the statement "population pressure contributed to (induced) technological development and thus led to a rise in per capita income and production". Describe the population growth pattern in the post-independence period and explain how the neo-Malthusian thinking persisted in the period. [1 + 4 + 5 = 10]

2. Explain how demographic dividend arises. How high is high enough to be called a dividend? When had India crossed the 60 per cent marks in the share of population in the age group of 15 to 64 years. When is it expected to reach the peak share of 69 per cent? For how long the share would remain above 65 per cent? [3 + 4 + 1 + 1 +1 = 10]

3. Explain the logic behind the construction of the Whipple index (W). State the assumption behind the use of Sex Ratio Score (SRS) for measuring accuracy of data on population age distribution. Write the mathematical expressions for SRS along with explanations for the symbols. [4 + 2 + 4 = 10]

4. Write short notes on the following.

(a) Age heaping and age shifting.

(b) U.N. Joint Score

(c) Significance of population age distribution

[3 + 4 +3 =10]

== END ==

B.Stat. II / Introduction to Markov Chain

Midsem. Exam. / Semester II 2014-15

Date - February 27, 2014 / Time - 2 hours

Maximum Score - 30

**NOTE : SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.**

1. Let $P$ be the transition probability matrix of a Markov chain on the states space $S = \{1, 2, 3, 4, 5, 6, 7\}$ and it is given by

$$P = \begin{bmatrix} 0.3 & 0.2 & 0.3 & 0 & 0.1 & 0.1 & 0 \\ 0.1 & 0.3 & 0 & 0.2 & 0.1 & 0 & 0.3 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.3 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 1.0 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0 & 0.6 & 0 \\ 0 & 0 & 0.6 & 0 & 0 & 0 & 0.4 \end{bmatrix}$$

(a) (4 marks) Describe, with a diagram, if this chain is irreducible and finding its communicating class(es). Classify the states according to transience, null recurrence and positive recurrence, with very brief justification.

(b) (3 marks) For each communicating class(es) find the stationary distribution (separately), whenever they exist.

(c) (3 marks) Find, with justification, $P(X_n = 3 \text{ or } 7, \text{ for some } n \geq 1 \mid X_0 = 1)$.

(d) (5 marks) Find, with justification, the matrix $\Pi$ such that, $P\Pi = \Pi P = \Pi = \Pi\,\Pi$ where each entries of $\Pi$ is nonnegative and $0 \leq \sum_j \pi_{ij} \leq 1$.

2. Let $\{X_n\}$ be a Markov chain on the state space $I = \{0, \pm1, \pm2, \ldots\}$ with stationary transition probability $p_{i,i+1} = \alpha_i > 0$, $p_{i,i-1} = \beta_i > 0$ and $p_{i,i} = \gamma_i = 1 - \alpha_i - \beta_i \geq 0$ and $p_{i,j} = 0$ otherwise. Note that the chain is irreducible.

(a) (3 marks) Let $c < d$ be two integers. Show that, for $c+1 \leq x \leq d-1$,

$$P(\{X_n\} \text{ reaches } c \text{ before } d \mid X_0 = x) = \frac{\sum_{j=x}^{d-1} \frac{\beta_j\beta_{j-1}\cdots\beta_{c+1}}{\alpha_j\alpha_{j-1}\cdots\alpha_{c+1}}}{1+\sum_{j=c+1}^{d-1} \frac{\beta_j\beta_{j-1}\cdots\beta_{c+1}}{\alpha_j\alpha_{j-1}\cdots\alpha_{c+1}}}.$$

(b) (3 marks) Using (2a) or otherwise, show that, for all $x > c$,

$$P(X_n = c \text{ for some } n \geq 1 \mid X_0 = x) = 1 \quad \text{iff} \quad \sum_{j=x}^{\infty} \frac{\beta_j\beta_{j-1}\cdots\beta_{c+1}}{\alpha_j\alpha_{j-1}\cdots\alpha_{c+1}} = \infty$$

$$\text{iff} \quad \sum_{j=1}^{\infty} \frac{\beta_j\beta_{j-1}\cdots\beta_1}{\alpha_j\alpha_{j-1}\cdots\alpha_1} = \infty.$$

(c) (4 marks) Using (2b) or otherwise, give a similar condition,

$$P(X_n = d \text{ for some } n \geq 1 \mid X_0 = x) = 1 \quad \text{for all } x < d,$$

Hence, combining these or otherwise, give the necessary and sufficient condtion for the state 0 to be recurrent.

3. Suppose the balls labeled $1, 2, \ldots, N$ are initially distributed between two boxes labeled I and II. At each time step a ball is randomly (uniformly) selected from the numbers $1, 2, \ldots, N$. Independently, of the ball selected, box I or II is selected with respective probability $p_1 > 0$ and $p_2 = 1 - p_1 > 0$. The ball selected is placed in the ball selected. Let $\{X_n\}$ be the number of balls in box I at time $n$. Note that it is an irreducible Markov chain.

(a) (5 marks) Write down the state space and find the stationary transition probabilities $p_{ij}$ (for all $i, j$) for this chain.

(b) (5 marks) Find its invariant probability distribution (vector) $\pi$ such that $\pi_j = \sum_i \pi p_{ij}$ with $\sum_j \pi_j = 1$.

All the best.

Indian Statistical Institute
Second Semester Examination 2015
Course Name: B Stat II Year
Subject : Economics II (Macroeconomics)

Date 11/04/2015 Maximum marks: 60 Duration 2.5 Hours

Answer all questions

**1.** Consider a Simple Keynesian Model for a closed economy with government, where investment is given by $I = \bar{I}$. Examine how an increase in the tax rate will affect GDP, private disposable income, private saving, government saving and total tax collection in this model. [22]

**2.** a) Explain why an increased desire to save may lead to a decline in aggregate planned saving in the Simple Keynesian Model.
b) Consider a Simple Keynesian Model for a closed economy with government, where aggregate planned saving is a proportional function of GDP ($Y$) and marginal propensity to invest with respect to $Y$ is 0.3. Suppose, following a parallel downward shift of the saving function by 10 units, equilibrium value of aggregate planned saving increases by 30 units. Compute the new consumption function. [22]

**3.** Consider an IS-LM model where aggregate planned investment is a function of rate of interest ($r$) alone. Following an increase in money supply in this model, the equilibrium levels of $Y$ and $r$ are known to have changed by 500 units and -4 units respectively. It is also known that, following an increase in $\bar{G}$ by 50 units, IS curve shifts vertically by 1 unit. Given the information noted above, derive the change in the equilibrium value of $Y$ if the government raises $\bar{G}$ by 50 units and at the same time changes money supply by such an amount that $I$ remains unaffected. (Assume all functions to be linear) (22)

Indian Statistical Institute

Second Semestral Examination 2014-15

Course Name : B.STAT II Year

Subject : Economic and Official Statistics and Demography

Date : 22.04.15 Maximum Marks : 50 Time 2 hrs

**GROUP A ( To be attempted first ; use separate answer scripts for Group A and B)**

**Answer as many Questions as you can in Section I and two questions from Section II (as per instructions given):**

**Section I:**

1. a) State the specific form of Cobb-Douglas Production function and show that the elasticity of substitution is unity at every point .
   b) Explain how does it (elasticity of Substitution) differ in case of CES production function .

5+5=10

2. Explain the notion of constant utility price index number and state in this connection why true price index number can not be estimated. 10

3. a) What are the difficulties in combining cross section and time series data for Engel curve construction?

b) State a possible method to pool cross section and time series data for Engel Curve (use linear equation).

6+4=10

4. a)State the axiomatic properties of an ideal Index number represented by function $P_{0T}=P(q_0,p_o,q_t,p_t)$ where $q_0,p_o,q_t,p_t$ are n vector base quantity, base price, current quantity, current price respectively.

b) State the Fisher's tests for consistency of an index number formula.

5+5=10

**Section II: Write Short notes on the following (any two) as per the guideline given .** 10+10

1. Indian Economic Census -Students belonging to Groups A ,B,C not permitted
2. Urban Frame Survey of NSSO (including sampling frame and estimation procedure)- Students belonging to Groups B,C,D not permitted
3. Exchange Rate Management of Reserve Bank of India – Students belonging to Groups C,D,E not permitted
4. Key Issues of Election Statistics- Students belonging to Groups D,E,F not permitted
5. Findings of Census 2011 (Highlights)- Students belonging to Groups E ,F,A not permitted
6. World Bank Data on Poverty- Students belonging to Groups F,A,B not permitted

# INDIAN STATISTICAL INSTITTUE, KOLKATA

## Second Semestral Examination 2014-15

**B.Stat (Hons.) II Year**

**Subject: Economic & Official Statistics and Demography**

Group B: Demography

Date: 22.04.2015 Maximum Marks: 50 Duration: 2 hours

*(Answer as much as you can.)*

1(a) Describe the Brass method of evaluation of birth and death registration using age distribution and child survivorship data.

(b) Explain step by step how a life table is constructed from the m-type mortality rates to obtain life expectancy at some age. (6+4 = 10)

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 of an age group of width n years.

ii) The standardized death rate for the 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. The data below relate to fertility in a country in 1976 and 1993.

(a) Calculate age specific fertility rates for the two years.

(b) Using the 1976 population as the standard, calculate the standardized fertility rate for 1993.

(c) Comment on your results.

| Age Group (in years) | 1976 | | 1993 | |
|---|---|---|---|---|
| | No. of births ('000) | Mid-year female population ('000) | No. of births ('000) | Mid-year female population ('000) |
| 15 - 19 | 57.9 | 1809 | 45.1 | 1455 |
| 20 - 24 | 182.2 | 1672 | 152.0 | 1831 |
| 25 - 29 | 220.7 | 1855 | 236.0 | 2070 |
| 30 - 34 | 90.8 | 1593 | 171.1 | 1967 |
| 35 - 39 | 26.1 | 1374 | 58.8 | 1729 |
| 40 - 44 | 6.5 | 1300 | 10.5 | 1750 |

(3 + 6 + 1 = 10)

4. (a) Write how Newton's halving formula is used for tackling errors due to inaccurate age reports or faulty enumeration.

(b) Explain in detail a method of rectifying an erroneously noted census count for a particular age group.

(4 + 6 = 10)

5. Mention three important features that make NFHS-4 different from the earlier rounds of the NFHS. Explain the formula which is being used in NFHS-4 for calculation of sample size. Describe the sample design of the survey. Mention three steps that are being used for controlling non-sampling error.

(2 + 2 + 3 + 3 = 10)

6. Exhibit the table format for simultaneous computation of Total Fertility Rate (TFR), Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR) inclusive of the computational steps. Interpret the three rates and order them in decreasing values of their estimates. Age survival rate, represented by $({}_5L_x/l_0)$, indicates survival of children up to which age?

(6 + 3 + 1 = 10)

# END #

# INDIAN STATISTICAL INSTITUTE

Second Semestral Examination: 2014–2015
B.Stat. (Hons.) 2nd Year. 2nd Semester
Statistical Methods IV

Date: April 27, 2015 Maximum Marks: 75 Duration: 3 and 1/2 hours

---

- This question paper carries 80 points. Answer as much as you can. However, the maximum you can score is 75.
- You should provide as much details as possible while answering a question. You must state clearly any result stated (and proved) in class you may need in order to answer a particular question.

---

**1.** Suppose $\boldsymbol{X} = (X_1, \ldots, X_p)^{\mathrm{T}} \sim \mathrm{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathbb{R}^p$ and $\boldsymbol{\Sigma}$ are both unknown. We wish to test the hypothesis

$$\mathrm{H}_0 : X_1 \textit{ and } (X_2, \ldots, X_p) \textit{ are independent} \quad \text{against} \quad \mathrm{H}_1 : \mathrm{H}_0 \textit{ is false},$$

based on $n$ i.i.d. realizations of $\boldsymbol{X}$, denoted by $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$. Denote by $R$, the sample multiple correlation coefficient. Let $\lambda$ denote the LRT for testing $\mathrm{H}_0$ against $\mathrm{H}_1$.

(a) Show that $\lambda = (1 - R^2)^{n/2}$.

(b) Show that under $\mathrm{H}_0$, $[R^2/(1 - R^2)] \cdot [(n-p)/(p-1)] \sim F_{p-1, n-p}$. [7+8 = 15]

**2.** Suppose i.i.d. observations $X_i$'s ($i \geq 1$) are being drawn from a gamma population with known shape parameter $\alpha$ ($> 0$) and unknown scale parameter $\theta$, where $\theta > 0$.

(a) Find the MLE $\hat{\theta}_n$ of $\theta$ based on $X_1, \ldots, X_n$.

(b) Find the asymptotic distribution of $\hat{\theta}_n$.

(c) Obtain the variance stabilising transformation for the distribution in (b) and describe how you can use it to obtain an approximate $100(1-\alpha)\%$ confidence interval for $\theta$.

[3+4+(3+5) = 15]

**3.** Suppose we know the sample median $M_n$ of $n$ i.i.d. observations drawn from a distribution having pdf $f(x) \stackrel{def}{=} \alpha x^{\alpha-1} \exp(-x^\alpha)$, $x \geq 0$, $\alpha > 0$. Describe how you can obtain an estimator of $\alpha$ which is based on $M_n$, and is asymptotically normal with mean $\alpha$ and variance, to be obtained by you. Is your estimator consistent? Give reasons.

[4+7+3 = 14]

[P.T.O.]

4. Suppose a multiple linear regression model is being modelled with **independent** observations denoted by $Y_1, \dots, Y_n$ and two non-random co-variates taking values $(x_1, \dots, x_n)$ and $(z_1, \dots, z_n)$. Assume that the $n \times 3$ matrix

$$\begin{pmatrix} 1 & x_1 & z_1 \\ 1 & x_2 & z_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & z_n \end{pmatrix}$$

is of full column rank. Suppose that $\text{Var}(Y_i) = \sigma^2$ for $i = 1, \dots, n$. We are interested in testing if the contribution of the covariate $z$ is significant. Argue using an appropriate analysis of variance and Gram-Schmidt orthogonalization process that

$$T := \left|\text{Cov}(Y, z) - \frac{\text{Cov}(x, z)}{\text{Var}(x)}\text{Cov}(Y, x)\right|$$

is expected to be large in that scenario. Suggest a suitable test statistic based on $T$. Also, assuming normality of the $Y_i$'s, obtain the null distribution of your test statistic.
[**Note.** You may assume distribution theoretic facts about quadratic forms of a multivariate normal vector.]

[10+4+4 = 18]

5. Consider $n$ i.i.d observations, denoted $X_1, \dots, X_n$, from an exponential distribution with location parameter $\theta$ and scale parameter $\sigma$, $\theta \in \mathbb{R}, \sigma > 0$. Denote the order statistics by $X_{(1)} \leq \dots \leq X_{(n)}$. Suppose now that we have a Type II censored sample where only the first $r$ order statistics $X_{(1)} \leq \dots \leq X_{(r)}$ are observed.

(a) Find the MLE of $(\theta, \sigma)$, denoted by $(\hat{\theta}, \hat{\sigma})$.

(b) Show that $W_1 := 2n(\hat{\theta} - \theta)/\sigma \sim \chi^2_2$, $W_2 := 2r\hat{\sigma}/\sigma \sim \chi^2_{2r}$, and that $W_1$ and $W_2$ are independent.

(c) Discuss how you can obtain a $100(1 - \alpha)\%$ confidence interval for $\theta$.
[**Note.** You may assume distribution theoretic facts about exponential spacings.]

[6+8+4 = 18]

****** *Best of Luck!* ******

B.Stat. II / Introduction to Markov Chain
Final Exam. / Semester II 2014-15
Time - 3 hours
Maximum Score - 50

**NOTE : SHOW ALL YOUR WORK TO GET THE FULL CREDIT. RESULTS USED MUST BE CLEARLY STATED.**

1. Let $P$ be the transition probability matrix of a Markov chain on the states space $S = \{1, 2, 3, 4, 5, 6, 7\}$ and it is given by

$$P = \begin{bmatrix} 0.2 & 0.2 & 0.3 & 0 & 0.2 & 0.1 & 0 \\ 0.1 & 0.2 & 0 & 0.2 & 0.2 & 0 & 0.3 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.2 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 1.0 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0 & 0.7 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 \end{bmatrix}$$

   (a) (4 marks) Describe, with a diagram, if this chain is irreducible and find its communicating class(es). Classify the states according to transience, null recurrence and positive recurrence, with very brief justification.

   (b) *(3 marks) For each communicating class(es) find the stationary distribution (separately), whenever they exist.*

   (c) (3 marks) Find, with justification, $P(X_n = 4 \text{ or } 6, \text{ for some } n \geq 1 \mid X_0 = 2)$.

   (d) (5 marks) Find, with justification, the matrix $\Pi$ (non-trivial one) such that, $P\Pi = \Pi P = \Pi = \Pi\,\Pi$ where each entries of $\Pi$ is nonnegative and $0 \leq \sum_j \pi_{ij} \leq 1$.

2. During each unit of time $k(\geq 0)$ number of customer arrives for service and joins a single line. The probability of such event $(1-\lambda)\lambda^k$. Also during each unit of time independently of new arrivals, a single service is completed with probability $p$ or continues on to the next period with probability $1 - p$. Let $X_n$ be the total number of customer (waiting in line or being serviced) at the $n$th unit of time.

   (a) (3 marks) Argue briefly that $X_n$ is a Markov chain. Describe its state space and determine its transition probability.

   (b) (3 marks) Determine the condition for which it is transient, null recurrent or positive recurrent.

   (c) (3+3 marks) Calculate the invariant distribution when it is positive recurrent. Find expected length of queue in equilibrium (i.e., under stationary distribution).

3. An individual with a highly contagious disease enters a population. During each subsequent period, either the carrier will infect a new person or be discovered and removed by public health officials. A carrier is discovered and removed with probability $q = 1 - p$ $(0 < q < 1)$ at each unit of time. An

unremoved infected individual is sure to infect a new individual in each time unit (independently of other remaining infected individual in the population). Let $X_n$ be number of infected individual at time $n$, with $X_0 = 1$. Note it can be assumed to be a Markov Chain.

(a) (3 marks) Determine its transition probability mentioning its states space.

(b) (3 marks) Determine the probability that eventually there would be no infectious people in the population.

(c) (4 marks) Would the expected time for containing the infection be finite or infinite? Justfy your answer.

Calculate the expected total number of people infected when there are no infectious people in the population.

4. Two friends (A and B) are asked to choose a number (between 1 to 10) each. They have chosen their numbers uniformly (from 1 to 10) and independently. Call them $A_0$ and $B_0$, respectively. Then a magician friend is drawing cards at random from a deck showing A and B both, one after another. If $A_0 = 3$, and $B_0 = 5$, then they would look at the number on the 3rd card and the 5th card respectively. If those numbers are 4 and 6 then $A_1 = 4$ and $B_1 = 6$, and they would wait to see the next 4th and the 6th card respectively and it goes on like this. Assume all picture cards, i.e., Jack, Queen and King, are removed from the deck. Let $A_n$ and $B_n$ be the number of the $n$th card of $A$ and $B$ respectively.

(a) (3+4 marks) Let $X_0 = A_0 - B_0$, and for $k \geq 1$, let $X_k = X_{k-1} - N_k$ if $X_{k-1} > 0$, $X_k = X_{k-1} + N_k$ if $X_{k-1} < 0$, and $X_k = 0$ if $X_{k-1} = 0$, where $N_k$ is the number on the $k$th card that is seen (by $A$ or $B$). Note $\{X_k\}$ is a Markov chain with one absorbing state.

(i) Determine its transition probabilities describing its state space. (ii) Show that, starting from any state, the expected time to reach to the absorbing state is 10.

(b) (3 marks) Calculate the expected number of card that the magician need to draw for $A$ and $B$ to see the same card for the first time.

[Throughout, assume the deck is infinitely large (or drawing is done with replacement).]

5. Write TRUE or FALSE and justify your answer (the justification should be Mathematical as much as possible).

(a) (4 marks) Markov chain with a stationary transition probability on countably infinite state space may have all inessential states but on finite state space it must have at least one essential state.

(b) (4 marks) Let $P = ((p_{ij}))$ be the stationary transition probability of a Markov chain on a countably infinite state space. Then $x_j = \sum_i x_i p_{ij}$ for all $i$ in the state space has a non-trivial solution $\{x_i\}$ implies the chain is recurrent.

(c) (5 marks) Let $\{X_n\}$ be an irreducible MC on a state space $\{0, 1, \ldots\}$ with stationary transition probability $P = ((p_{ij}))$. Then the state 0 is recurrent if, $y_i \geq \sum_{j \neq 0} p_{ij} y_j$ for $i \neq 0$, has a solution $\{y_j\}$ such that $y_j \to \infty$ as $j \to \infty$.

All the best.

# INDIAN STATISTICAL INSTITUTE

Semester Examination : 2014-15( Second Semester)

**B. Stat. II year**

**Discrete Maths.**

Date: 08.05.15 Maximum Marks : 100 Duration : 3 hrs

Note: Answer as many as you can. The maximum you can score is 100. Notation is as used in the class.

1. Suppose an open necklace consists of a string of 3 beads, each being either blue or red. Two such necklaces $x$ and $y$ are considered same if $x = y$ or if $y$ can be obtained from $x$ by reversing. Find the number of distinct necklaces by enumeration. Verify by using Burnside's Lemma. [(4+3)=7]

2. (a) Suppose the edges of $K_4$ are coloured blue or red. Show that there is a monochromatic triangle.

   (b) If $p \geq 2$ and $q \geq 2$ are integers then show that the Ramsry number $R(p, q)$ satisfies
$$R(p,q) \leq \binom{p+q-2}{p-1}.$$
[(4+5)=9]

3. Let $M$ be the incidence matrix of a $k$-regular simple graph. Show that $k$ is an eigen value of $M$.

[5]

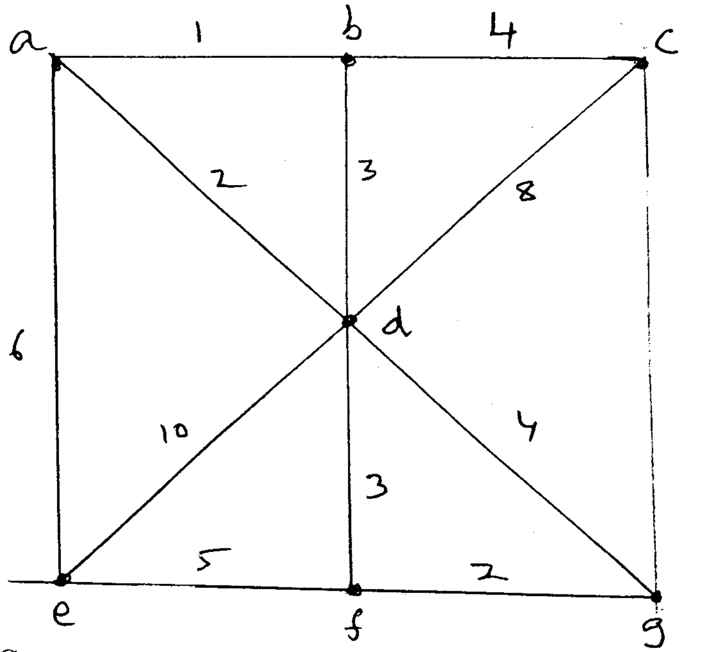
4. (a) Let $\tau(G)$ denote the number of spanning trees of a graph $G$. Show that if $e$ is an edge of a loopless graph $G$, then
$$\tau(G) = \tau(G - e) + \tau(G/e).$$
   Hence or otherwise, evaluate the number of spanning trees of $K_{3,3}$

(b) Find the number of spanning trees of a wheel with 6 vertices

[(5+5)+5=15]

5. (a) If the weights of the edges of a graph $G$ are distinct, then show that the minimum spanning tree is unique.

(b) Find a minimum cost spanning tree of the following using Prim's algorithm starting at $a$.




[6+8=14]

6. (a) Construct a graph $G$ which is 4-chromatic and contains no triangle. Justify.

(b) Find the chromatic polynomial of a cycle of length $n$.

(c) Show that no graph has chromatic polynomial $\lambda^4 - 3\lambda^3 + 3\lambda^2$.

[7+6+4=17]

7. (a) With usual notation, prove Euler's formula for planar graph viz.

$$n - e + f = 2.$$

Hence, show that for a planar graph with at least 3 vertices, $e \leq 3n - 6$. Show that the equality holds if G is a plane triangulation i.e. an embedding where each face is a triangle.

(b) Show that $K_{3,3}$ is non-planar.

(c) Why is a planar graph 6-colourable?

[(5+3+2)+4+3]

8. (a) Let $G[X : Y]$ be a bipartite graph. State the necessary and sufficient condition for $G$ to have a matching that saturates $X$. Hence, show that a system of distinct representation $a_1, a_2, \ldots, a_n$ with $a_i \in A_i$ can be chosen from the collection $A_1, A_2, \ldots, A_n$ of sets iff

$$|\bigcup_{i \in I} A_i| \geq |I|, \quad \text{for all } I \subseteq \{1, 2, \ldots, n\}.$$

(b) Suppose $G$ has $2n$ vertices and for all $a \neq b$ such that $\{a, b\}$ is not an edge of $G$

$$deg(a) + deg(b) \geq 2n.$$

Show that $G$ has a perfect matching.

[(3+4)+10=17]

9. (a) Show that in a directed network, the value of any $(s, t)$ flow is less than equal to the capacity of any $(s, t)$ cut.

(b) State the Max -Flow Min- Cut Theorem.

[6+3]=9

**INDIAN STATISTICAL INSTITUTE**
Second Semestral Examination: (2014 – 2015)
B. Stat II Year
AGRICULTURE

Date 11/05/15 Maximum Marks 50 Duration 3:00 hours .
(Attempt any five questions)
(Number of copies of the question paper required 10 )

1. Write a short note on about the System of Rice Intsification (SRI) 10

2. What are the differences between Manure and Fertilizer? Calculate the quantity of VC, Urea, SSP and KCL required for the 1 hectare Rice crop to supply the nutrient requirement of 160 kg N, 100 kg $P_2O_5$ and 100 kg $K_2O$ per hectare. Note that 50% of required N should be given through VC.
3+7

3. What are the different growth promoting and yield attributing characters of rice plant. Estimate the expected yield of rice grain in t/ha from the following data:
(i) Spacing – 20x20cm, (ii) Average no. of tillers/hill –35, (iii) Average no. of effective tillers/hill –30, (iv) Average no. of grain/panicle –65 (v) Average no. of unfilled grain/panicle –7 (vi) Test weight - 26g.
4+6

4. Write in brief (any two): 5 x 2
   a) Why the Boro rice yields more than the Kharif rice
   b) Organic farming
   c) Soil formation

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

| Cropping System | Rice yield in kg/ha | Soybean yield in kg/ha |
|---|---|---|
| Rice Sole | 4535 | - |
| Soybean Sole | - | 6752 |
| Rice+Soybean (2:2) | 2856 | 3050 |
| Rice+Soybean (2:1) | 3122 | 2572 |

10

6. Write short notes on <u>any five</u> of the following : 2 x 5
   a) Monsoon onset
   b) Potential Evapo-transpiration
   c) Cup counter anemometer
   d) Reproductive stages in rice
   e) Permanent wilting point
   f) Field capacity
   g) Capillary water