Indian Statistical Institute, Kolkata

Final examination: Second semester 2013 -'14

B.Stat I year

Subject	:	**Analysis II**
Time	:	3 hours 30 minutes
Maximum score	:	65

Instructions:

- *Justify every step in order to get full credit of your answers, stating clearly the result(s) that you use. Points will be deducted for missing arguments. Partial credit will be given for your approach to the problem.*
- *Switch off and deposit your mobile phones to the invigilator during the entire examination.*

(1) Let f be a Riemann integrable function over the interval $[a,b]$ not necessarily continuous. Prove that there exists a sequence $\{g_n\}_{n\in\mathbb{N}}$ of continuous functions on $[a,b]$ such that $\lim_{n\to\infty}\int_a^b |g_n(x) - f(x)|\,dx = 0$.

[15 marks]

(2) Let $f : [0,1] \to \mathbb{C}$ be continous such that $\int_0^1 x^n f(x)dx = 0$ for all $n \geq 2$. What is the best that you can say about f? Justify every step.

[11 marks]

(3) Let $a_n = \frac{1}{n^2+3n+2}$ for all integers n. Show that there exists a continuous 2π-periodic function f on $\mathbb{R}$ whose n-th Fourier coefficient is given by a_n for all n. Does the Fourier series converge to the continuous function f and if so, what is the kind of convergence?

[11 marks]

(4) Find the most general form of real valued solutions of the differential equation

$$y'' + y' + y = e^{-2x} + x^3 \quad \text{where } x \in \mathbb{R}.$$

[11 marks]

(5) Consider the differential equation $y'' - y = x(y'' - y')$.

(a) Find a particular solution φ by guessing.

(b) Using φ, find a linearly independent solution ψ.

(c) Write down the most general form of solution.

[11 marks]

(6) Use power series to find the most general form of solution to the differential equation $y'' - xy' - y = 0$ where $x \in \mathbb{R}$.

[11 marks]

INDIAN STATISTICAL INSTITUTE

B-Stat I, Second Semester, 2013-14

STATISTICAL METHODS II

Date: 13·05·14 Semester Examination Time: 2 1/2 hours

Total Point 100

1. Consider the following $n = 20$ observations from the joint distribution of (X_1, X_2, X_3). It is believed that X_1 and X_2 have a very strong natural correlation. However a new claim states that the high observed correlation between X_1 and X_2 arises from the large correlation which each of these two variables enjoys with X_3, and when adjusted for the latter, the net correlation between X_1 and X_2 is low. By studying the partial and/or multiple correlations of these variables (together with the total correlations), state whether you agree or disagree with this claim.

X_1	X_2	X_3
0.79736193	−0.34287378	0.77650455
−0.97092802	−0.63681782	−1.27794940
−0.16362481	−0.24208279	0.24385236
−0.05191644	−0.74491306	0.04270265
−1.23896641	−1.19593420	−1.12998282
−2.52847641	−1.90194863	−2.26449413
−1.34575814	−1.80489361	−1.23854370
−0.58718616	−1.02312883	−0.52612468
−0.58974340	−1.25802984	−0.90512423
−0.98235479	−1.74539648	−1.42751654
−0.84281780	−0.13454252	−0.28501330
−0.21302151	−0.04902253	0.06547209
0.25851092	0.76697194	0.06313120
1.00546297	0.31334488	0.88140972
0.16046021	0.25485487	0.33813630
−1.14600872	−1.34525017	−0.87229566
−0.37071964	−0.60438862	−0.78135154
0.52905052	1.03908493	0.68256995
−0.43542312	−0.03492247	−0.33095538
−0.14260157	0.18844682	0.36285226

You may use the following summary statistics:

$$\sum_{i=1}^{20} X_{1i} = -8.8587; \quad \sum_{i=1}^{20} X_{2i} = -10.50144; \quad \sum_{i=1}^{20} X_{3i} = -7.58272;$$

$$\sum_{i=1}^{20} X_{1i}^2 = 16.8047; \quad \sum_{i=1}^{20} X_{2i}^2 = 19.18105; \quad \sum_{i=1}^{20} X_{3i}^2 = 16.42928;$$

$$\sum_{i=1}^{20} X_{1i}X_{2i} = 15.18203; \quad \sum_{i=1}^{20} X_{1i}X_{3i} = 15.78126; \quad \sum_{i=1}^{20} X_{2i}X_{3i} = 15.39986.$$

[15]

2. Consider the A-B-O blood group gene frequency data given by the vector of cell frequencies

$$\boldsymbol{y} = (n_O, n_A, n_B, n_{AB}) = (176, 182, 60, 17),$$

where the individual components of $\boldsymbol{y}$ are the frequencies of the corresponding phenotypes. The parameter vector is $\theta = (p, q, r)$, where the individual components of θ represent the proportion of the gene counts (that is, count of the alleles A, B and O), with $p + q + r = 1$. We want to find the maximum likelihood estimators of the parameters, and want to use the EM algorithm for this purpose.

(a) Explain what would be the complete data in this example.

(b) Given estimates $p^{(k)}, q^{(k)}$ and $r^{(k)}$ at the k-th step of the iteration, explain how the E and M steps will be performed to get the parameter estimates at the $(k+1)$-th step.

(c) Given current estimates 0.264, 0.093 and 0.643 of p, q and r, find the parameter estimates after the next iteration. [4+8+8=20]

3. The probability density functions of the Exponential(θ), Gamma(α, θ) and the Beta (α, β) distributions are given by

$$f(x) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0$$

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta}, \quad x > 0,$$

and

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1,$$

where $B(\alpha, \beta)$ represents the beta function.

(a) Suppose that $X \sim$ Gamma(α, θ) and $Y \sim$ Gamma(β, θ) and X and Y are independent. Find the distribution of $\frac{X}{X+Y}$.

(b) Suppose you can generate any number of independent Uniform(0, 1) random variables. Show how you can simulate an observation from an Exponential(θ) distribution, a Gamma(α, θ) distribution, and a Beta(α, β) distribution. [6+14=20]

4. Let X_1 and X_2 be independently distributed as Poisson(θ) random variables.

   (a) Find a sufficient statistic for this model.

   (b) Find the minimum varianced unbiased estimator of θ.

   (c) Check whether the estimator obtained in the previous part attains the Cramer-Rao lower bound.

   (d) Show, from first principles, that $X_1 + 2X_2$ is not a sufficient statistic for this model. [2+4+4+5=15]

5. Assignments. [30]

Indian Statistical Institute

Semester Examination 2013-14

B. Stat. I st Year — Vectors and Matrices II

Date : 9.5.14. Maximum Marks : 50 Time: 3 hrs.

1. Prove or disprove: H is a bilinear form on a finite dimensional vector space V with dimension $(V) > 1$. For any $x \in V$, there exists $y \in V$ such that $y \neq 0$, but $H(x, y) = 0$. [5]

2. Describe all linear operators T on $\mathbb{R}^2$ such that T is diagonalisable and $T^3 - 2T^2 + T$ is the zero transformation. [5]

3. Let T be a linear operator on V whose characteristic polynomial splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distict eigenvalues of T. Then show that T is diagonalisable if and only if $rank(T - \lambda_i I) = rank((T - \lambda_i I)^2)$ for $1 \leq i \leq k$. *Hence* show that if T is a diagonalisable linear operator on a finite dimensional vector space V and W is a T invariant subspace of V, then T_W is diagonalisable. [10]

4. Prove or disprove : A linear operator is diagonalisable if its minimal polynomial splits. [4]

5. Let $K : \mathbb{R}^2 \to \mathbb{R}$ be a quadratic form defined by $K(x, y) = -2x^2 + 4xy + y^2$. Find a symmetric bilinear form H such that $K(u) = H(u, u)$ for all $u \in \mathbb{R}^2$. Find an orthonormal basis Γ of $\mathbb{R}^2$ such that $\Psi_\Gamma(H)$ is a diagonal matrix. [7]

6. Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$. Find a singular value decomposition of A. Use the decomposition to find the Moore-Penrose G-inverse of A. [12]

7. Let T be a linear operator on $P_2(\mathbb{R})$ defined by $T(f(x)) = 2f(x) - f'(x)$. Find a basis for each generalised eigenspace of T consisting of a union of disjoint cycles of generalised eigenvectors. Hence find a Jordan canonical form J of T. [12]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2013-14 (Second Semester)
Bachelor of Statistics (B. Stat.) I Year
Probability Theory II

Teacher: Parthanil Roy

Date: 05/05/2014 **Maximum Marks: 70** **Duration: 3 Hours**

Note:

- **Please write your roll number on top of your answer paper.**
- **You may use any theorem proved or stated in the class but do not forget to quote the appropriate result.**
- **You are NOT allowed to use class notes, books, homework solutions, list of theorems, formulas etc. If you are caught using any, you will get a zero in this examination.**
- **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 (U, V) is a bivariate normal random vector with $E(U) = E(V) = 0$, $Var(U) = Var(V) = 1$ and $Cov(U, V) = 0.5$. Compute the correlation coefficient between U^2 and V^2. [8]

2. Let $Y_1 < Y_2 < Y_3$ be an order statistic of size 3 from the distribution with probability density function $f(x) = 2x$ for $0 < x < 1$. For $i = 1, 2$, define $U_i = Y_i/Y_{i+1}$. Compute the joint probability density function of U_1 and U_2, and show that $U_1 + U_2$ is independent of Y_3. [8 + 2]

3. Let X, Y and Z be three independent standard uniform random variables and $S = X + Y + Z$. Find (a) the conditional density of S given $Y = y$ and $Z = z$ and (b) the conditional density of (Y, Z) given $S = s$. [6 + 6]

4. 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]

5. Other evaluations: presentations and quizzes. [10 + 20 = 30]

Wish you all the best

INDIAN STATISTICAL INSTITUTE
Semester Examination
B. Stat I year, 2nd Sem: 2013–2014
Numerical Analysis

Date: 29/4/2014, Time: 2 Hours

Please try to write all the part answers of a question at the same place.
Answer as many questions as you can. The maximum you can score is 70.

1. Assignments. [20]

2. (a) Show that in the interval $[-1,1]$, the Chebyshev polynomials can be wrriten as
$$T_n(x) = \cos(n \cos^{-1} x),$$
where n denotes the degree of the polynomial.

   (b) Define $Q_n(x) = \frac{1}{2^{n-1}} T_n(x)$. Choose $x_i = \cos(\frac{i\pi}{n})$, $i = 0, 1, \ldots, n$. Show that for any $i = 0, 1, \ldots, n$, we have
$$Q_n(x_i) = \frac{(-1)^i}{2^{n-1}}.$$

   (c) Show that for any monic polyomial P of degree n,
$$\max_{-1 \leq x \leq 1} |P(x)| \geq \frac{1}{2^{n-1}}.$$
(**Hint**: Use (b)).

   (d) Use the above results to show that $\Pi_n(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$ is minimized on $[-1, 1]$, if x_i's are chosen as the zeros of the Chebyshev polynomial T_{n+1}.

[4 + 3 + 5 + 3 = 15]

3. (a) Show that the parameters $M_i = s''(x_i)$, $i = 0, 1, \ldots, n$ of the complete cubic spline interpolant $s(x)$ can be obtained by solving a linear system $HM = Y$, where H is a tridiagonal matrix.

   (b) Show that this system can be solved efficiently using about $8n$ arithmetic operations.

[10 + 4 = 14]

4. (a) Why is $f[x_0, x_1]$ a better approximation to $f'(x)$ at the midpoint $a = \frac{x_0+x_1}{2}$ than at x_0 or x_1? In general, when is $n! f[x_0, \ldots, x_n]$ a good approximation to the n-th derivative of $f(x)$?

   (b) Compute $\int_0^{0.8} \left(1 + \frac{\sin x}{x}\right) dx$ using Simpson's 1/3rd rule with $h = 0.2$.

   (c) Why is Gaussian Quadrature more efficient than Cote's integration formula?

[(2 + 2) + 4 + 3 = 11]

5. (a) Derive the expressions for both *local* and *global* errors in Euler's method of solving the first order ODE (initial value problem).

   (b) For increasing the accuracy of numerical solution of an ODE, why do we prefer Runge-Kutta methods over generalizing Euler's method into a Taylor algorithm of order m?

   (c) Using Adams-Bashforth Predictor and Adams-Moulton Corrector methods of order 3, obtain the solution of $\frac{dy}{dx} = x^2 y + x^2$ at $x = 1.4$, given the following values.

x	1	1.1	1.2	1.3
y	1	1.233	1.548488	1.978921

[(3 + 5) + 4 + 8) = 20]

INDIAN STATISTICAL INSTITUTE

B-Stat I, Second Semester, 2013-14

STATISTICAL METHODS II

Date: 28.02.2014 Midsemester Examination Time: 2 1/2 hours

Total Points ~~80~~. 100

1. (a) Suppose that $p_1, p_2, \ldots, p_k$ are known, real, nonnegative constants, and they satisfy the condition $\sum_{i=1}^{k} p_i = 1$; let $f_1, f_2, \ldots, f_k$ represent k probability density functions. Consider the function

$$f_{\mathrm{M}}(x) = p_1 f_1(x) + p_2 f_2(x) + \ldots + p_k f_k(x). \quad (1)$$

   i. Show that $f_{\mathrm{M}}(\cdot)$ is a valid probability density function (it is called the mixture of the densities $f_1, f_2, \ldots, f_k$ with mixing weights $p_1, p_2, \ldots, p_k$).

   ii. Assuming that we know how to generate random nubmbers of each of the component densities $f_1, f_2, \ldots, f_k$, describe how to draw a random number from the mixture density in (1).

   (b) We wish to draw a random number X from the Laplace distribution with density

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

   Suppose that we know how to generate random numbers from the Uniform (0, 1) distribution. Starting from random variables distributed uniformly on (0,1), describe how you can draw a random number from the Laplace distribution presented above. [(2+8)+10=20]

2. Consider a random sample $X_1, X_2, \ldots, X_n$ from the parametric class of densities $\{f_\theta | \theta \in \Theta\}$. We are interested in estimating the function $\tau(\theta)$.

   Suppose that we have an unbiased estimator $T = T(X_1, X_2, \ldots, X_n)$ of $\tau(\theta)$ which is efficiency (meaning that its variance attains the Cramer-Rao lower bound). Show that in this case T must be the maximum likelihood estimator of $\tau(\theta)$. [15]

3. Let $r_{1(2.34\ldots p)}$ be the correlation between x_1 and $x_{2.34\ldots p}$ (the residual after regressing x_2 on $x_3, x_4, \ldots, x_p$). Show that

   (a) $r^2_{1(2.34\ldots p)} \le r^2_{12.34\ldots p}$.

   (b) $r^2_{1.23\ldots p} = r^2_{1p} + r^2_{1(p-1.p)} + \ldots + r^2_{1(2.34\ldots p)}$. [5+15=20]

4. Consider the multiple linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y}$ and ϵ are each $n \times 1$ vectors, $\mathbf{X}$ is an $n \times p$ matrix and β is a $p \times 1$ vector; let $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 I_{n \times n})$. Let $\mathrm{SSE} = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^T(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ be the error some of squares, where $\hat{\boldsymbol{\beta}}$ is the least squares estimator of $\boldsymbol{\beta}$.

   (a) Show that SSE/σ^2 has a $\chi^2(n-p)$ distribution.

   (b) Let $S^2 = \mathrm{SSE}/(n-p)$. Find the value of c so that the cS^2 has the minimum mean square error as an estimator of σ^2. [10+10=20]

5. Computer Assignment. To be assigned, collected and graded separately. [25]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination
B. Stat I year, 2nd Sem: 2013–2014
Numerical Analysis

Date: 27. 02. 2014, Maximum Marks: 90, Time: 3 Hours (10:30 AM to 1:30 PM)
Please try to write all the part answers of a question at the same place.

1. (a) Show that the error sequence $\{e_n\}$ in *Secant* method satisfies the condition $K|e_n| \leq \delta^{a_n}$, where K is a suitably defined constant in the interval of interest containing the root, $\delta < 1$ (assuming convergence) and $a_n \in \mathbb{N}$.

   (b) Prove that the sequence $\{a_n\}$ of exponents form a *Fibonacci* sequence.

   [8 + 6 = 14]

2. (a) What is the geometric interpretation of *Newton-Raphson* method?

   (b) Derive a recursive algorithm for finding the square root of a non-negative real number using Newton-Raphson method and write the pseudocode for your derived algorithm.

   (c) Extend the above algorithm for finding the n-th root of a non-negative real number.

   (d) Use your extended algorithm for finding the positive real cube-root of 10 correct up to eight decimal places.

   [4 + (6 + 4) + 4 + 8 = 26]

3. (a) What is the order of convergence of the *bisection* method? If we modify the bisection method so that $c = \frac{2a+b}{3}$ (the variables having usual meaning), what would be the effect on the order of convergence?

   (b) Find a real root of $f(x) = 2^x - 3x = 0$ correct up to 3 decimal places by the bisection method.

   [4 + 8 = 12]

4. (a) If $V(x_1, \ldots, x_n)$ is an $n \times n$ *Vandermonde* matrix, show that its determinant is equal to $\Pi_{1 \leq j < i \leq n}(x_i - x_j)$.

   (b) Usually, we interpolate n given points by a polynomial of degree at most $n-1$. Discuss two "unusual" situations when we interpolate n given points by a polynomial of degree $2n - 1$ and when we interpolate a single point by a polynomial of degree $n - 1$. What are the expressions for errors in both these "unusual" cases.

   [8 + (4 + 4) = 16]

5. (a) Compare the time complexities of ordinary *Lagrange's method* and *Newton's divided difference method* of polynomial interpolation for calculating the coefficients and for evaluation at a new point.

(b) How can the Lagrange's method be improved to make it as efficient as Newton's method?

[(4 + 4) + 6 = 14]

6. Consider interpolation with equally-spaced points with h as the equal space between consecutive points.

(a) Establish the following relations between the *forward* and the *backward difference operators*: $\nabla = \frac{\Delta}{1+\Delta}$.

(b) Define two operators

$$\delta f(x) = f(x + \frac{h}{2}) - f(x - \frac{h}{2})$$

and

$$\mu f(x) = f(x + \frac{h}{2}) + f(x - \frac{h}{2}).$$

Show that $\delta = 2\sqrt{\mu^2 - 1}$.

[6 + 8 = 14]

7. (a) Given the following table, approximate $f(x)$ as a polynomial in x using *Newton's divided difference formula.*

x	-1	0	3	6	7
$f(x)$	3	-6	39	822	1611

(b) Estimate $f(10)$ for the above function.

[10 + 4 = 14]

Indian Statistical Institute

Mid-Semester Examination 2013-14

B. Stat. I st Year — Vectors and Matrices II

Date : February 26, 2014 — Maximum Marks : 40 — Time: 2 hrs.

1. Let $x_1, x_2, \ldots, x_k$ be an orthonormal basis of an inner product space and $y_1 = x_1, y_2 = x_1 + x_2, \ldots, y_k = x_1 + \ldots + x_k$. Apply the Gram-Schmidt orthogonalization process to $y_1, y_2, \ldots, y_k$. [8]

2. Let P_4 be the vector space of all polynomials of degree less than 4 with real coefficients. The inner product in P_4 is defined by $< y_1(x), y_2(x) >= \int_{-1}^{1} y_1(x)y_2(x)dx$ for $y_1(x), y_2(x) \in P_4$. Let S be the vector space spanned by the polynomials $y_1(x) = 1, y_2(x) = x$, and $y_3(x) = x^2$. Determine the best approximation of the polynomial $2x + 3x^2 - 4x^3$ by a polynomial from S. [10]

3. Let T be a linear operator on a finite dimensional inner product space V. Suppose that the characteristic polynomial of T splits. Then show that there exists an orthonormal basis Γ for V such that $[T]_\Gamma$ is upper triangular. [7]

4. Let $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$. Find an orthogonal matrix P such that $P^t AP = D$, where D is a diagonal matrix. [10]

5. Let T be a linear operator on R^2 such that $T(x) = x \ \forall \ x \in L$ and $T(x) = -x \ \forall \ x \in L^{\perp}$, where L is a line through origin and $L^{\perp}$ is the orthogonal complement of L. Show that T is an orthogonal operator.[8]

6. Let T be a normal operator on a finite dimensional complex inner product space V. Show that there exists a normal operator U such that $U^2 = T$. [7]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2013-14 (Second Semester)
Bachelor of Statistics (B. Stat.) I Year
Probability Theory II

Teacher: Parthanil Roy

Date: 25/02/2013 **Maximum Marks: 30** **Duration: 10:30 am - 12:30 am**

Note:

- **Please write your roll number on top of your answer paper.**
- **There are three problems each carrying 10 marks with a total of 30 marks. Solve as many as you can. Show all your works and write explanations when needed. Maximum you can score is 30 marks.**
- **You may use any fact proved in the class but do not forget to quote the appropriate result.**
- **You are NOT allowed to use class notes, books, homework solutions, list of theorems, formulas etc. If you are caught using any, you will get a zero grade in this exam.**

1. Fix two positive real numbers t and λ and a nonnegative integer n. Suppose $T \sim Gamma(n+1, \lambda)$ and $N \sim Poi(\lambda t)$. Show that $P(T > t) = P(N \leq n)$. [10]

2. Range of a sample is defined as the absolute difference between the minimum and the maximum sample point. For example, if the sample consists of the values 6, 8, 1, 4 and 7, then the range of the sample is 7. Let R be the range of a random sample of size n from uniform distribution on $(0, 1)$. Show that R is an absolutely continuous random variable and find its probability density function. [10]

3. Let $Z_1, Z_2, \ldots Z_{20}$ be independent and identically distributed random variables with $Z_1 \sim N(0,1)$. Find the joint probability density function of $\sum_{i=1}^{20} Z_i^2$ and $\frac{\sum_{i=1}^{8} Z_i^2}{\sum_{i=1}^{20} Z_i^2}$. Are they independent? [9 + 1]

Wish you all the best

INDIAN STATISTICAL INSTITUTE, KOLKATA

MID-SEMESTER EXAMINATION: SECOND SEMESTER 2013 -'14

B.STAT I YEAR

Subject	:	**Analysis II**
Time	:	3 hours
Maximum score	:	30
Date	:	24.02.2014

Instructions:

- *Justify every step in order to get full credit of your answers, stating clearly the result(s) that you use. Points will be deducted for missing arguments. Partial credit will be given for your approach to the problem.*
- *Switch off and deposit your mobile phones to the invigilator during the entire examination.*
- *Total marks carried by the questions turns out to be 50 which is more than 30. The total marks obtained will be multiplied with* $\frac{3}{5}$.

1) Show that the function $f : [0,1] \to \mathbb{R}$ defined by $f(x) = \begin{cases} 1, & \text{if } x = \frac{1}{n}, \\ 0, & \text{otherwise,} \end{cases}$

is Riemann integrable over $[0,1]$ from the definition. Justify your answers with proper arguments.

[10 marks]

2) Let $a < b$. For $n \in \mathbb{N}$, consider the partition $P_n = \{x_i = a + \frac{i}{n}(b-a) : 0 \le i \le n\}$. Will a bounded fucntion $f : [0,1] \to \mathbb{R}$ be always Riemann integrable if $\lim\limits_{n\to\infty} \frac{1}{n}\sum\limits_{i=1}^{n} f(\frac{x_i + x_{i-1}}{2})$ exists? If yes, prove it, and if not, find a counter example. Justify every step.

[10 marks]

3) Find a sequence of **continuous functions** on a **compact interval** converging to a **continuous function pointwise** but **not uniformly**. Stating a correct example without justification will fetch partial credit. So, justify every statement.

[10 marks]

4) Prove that $\int\limits_1^\infty \frac{\sin(t)}{t} dt$ **is an improper integral** which converges but $\int\limits_1^\infty \left|\frac{\sin(t)}{t}\right| dt$ does not. (*Hints: For the second integral, consider intervals of length* π.)

[10 marks]

5) For $n \in \mathbb{N}$, define $f_n : [0,\infty) \to \mathbb{R}$ by $f_n(t) = te^{-nt}$.

(i) Prove that $\sum\limits_{n=1}^{\infty} f_n$ is convergent pointwise.

(ii) Show that $\left[\sup\limits_{x\ge 0} \sum\limits_{k=n+1}^{2n} f_k(x)\right] \ge e^{-2}$ for all $n \in \mathbb{N}$. (*Hints: Consider the value at* $x = \frac{1}{n}$.)

(iii) Prove that $\sum\limits_{n=1}^{\infty} f_n$ is not uniformly convergent. (*Hint: Use part (ii) even if you are unable to prove it.*)

[10 marks]

INDIAN STATISTICAL INSTITUTE

Back-Paper Examination, 1st Semester, 2013-14

Statistical Methods I, B.Stat I

Total Points 100

Date: 29.01.14 Time: 3 hours

1. Consider the following sample of five numbers 1.0, 2.3, 3.0, 4.2 and 1.5.

   (a) Find the sample mean and sample variance based on the above five numbers.

   (b) Find five numbers which have the same variance as the above five observations, but have a mean three units higher.

   (c) Find five numbers which have the same mean as that of the five original numbers but has a variance four times larger than that of the original sample.

   (d) Find five numbers which have a mean three units larger than that of the five original numbers and has a variance four times larger than that of the original sample.

   [2+4+7+7=20 points]

2. The probability density function of a gamma distribution with parameters α and θ is given by

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta}, \; 0 \leq x \leq \infty.$$

   Suppose that X_1 and X_2 have independent gamma distributions with parameters α, θ, and β, θ respectively. Find the probability density function of $Y = X_1/(X_1 + X_2)$.

   [18 points]

3. Suppose that 10% of the people in a certain population have the eye disease glaucoma. For persons who have glaucoma, measurements of eye pressure X will be normally distributed with a mean of 25 and a variance of 1. For persons who do not have glaucoma, the pressure X will be distributed with a mean of 20 and a variance 1. Suppose that a person is selected at random.

(a) Determine the conditional probability that the person has has glaucoma given $X = x$.

(b) For what values of x is the conditional probability in part 3a greater than 1/2?

[6+6=12]

4. Suppose that $X_1, X_2, \ldots, X_n$ represent an independently and identically distributed sample from the $N(\mu, 1)$ distribution.

   (a) Suppose we want to test the null hypothesis $H_0 : \mu = 0$ versus $H_1 : \mu > 0$. We will reject the null hypothesis when $\bar{X} > b$, for some constant b. For $n = 25$, find the value of b so that the test has a size equal to 0.05.

   (b) What sample size will be necessary for the test in part (4a) has power 0.8 or higher?

[6+9=15]

5. Suppose that the joint distribution of (X, Y) is uniform over a circle in the xy plane. Determine the correlation of X and Y. [10]

6. Suppose we observe the random variable X which has a binomial distribution with parameters n and p. Based on the value of X, find an unbiased estimator of the variance of X. [12]

7. Consider independent random samples from two normal distributions, $X_i \sim N(\mu_1, \sigma_1^2)$ and $Y_j \sim N(\mu_2, \sigma_2^2)$, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$. Assuming that μ_1 and μ_2 are known, derive a $100(1-\alpha)\%$ confidence interval for σ_2^2/σ_1^2. [13]

Indian Statistical Institute, Kolkata

Back paper examination: First semester 2013 -'14

B.Stat I year

Subject	:	**Analysis I**
Time	:	3 hours
Maximum score	:	100
Date	:	27.01.2014

Instructions:

- *Justify every step in order to get full credit of your answers, stating clearly the result(s) that you use. Points will be deducted for missing arguments. Partial credit will be given for your approach to the problem.*
- *Switch off and deposit your mobile phones to the invigilator during the entire examination.*
- *Total marks carried by the questions turns out to be 72 which is less than 100. The total marks obtained will be multiplied by* $\frac{100}{72}$.

(1) Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be two sequences in $\mathbb{R}$ which are bounded above. Set $s = \limsup_{n\to\infty} a_n$, $t = \limsup_{n\to\infty} b_n$ and $u = \limsup_{n\to\infty} (a_n + b_n)$.
Prove that $s + t = u$ if and only if there exists subsequences $\{a_{n_k}\}_{k\in\mathbb{N}}$ and $\{b_{n_k}\}_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} a_{n_k} = s$ and $\lim_{k\to\infty} b_{n_k} = t$. (*Hint: For the 'only if' part, first prove that: given* $\varepsilon > 0$, *there exists* $n \in \mathbb{N}$ *such that (i)* $|(a_n + b_n) - u| < \varepsilon/2$, *(ii)* $a_n < s + \frac{\varepsilon}{2}$ *and (iii)* $b_n < s + \frac{\varepsilon}{2}$, *and then, solve the inequalities for* $x = a_n - s$ *and* $y = b_n - t$.)

[16 marks]

(2) Give an example of a function $f : (0,1) \to \mathbb{R}$ such that

(i) f is continuous on a dense subset D of $(0,1)$,

(ii) $E = (0,1) \setminus D$ is dense in $(0,1)$, and

(iii) f is discontinuous at every point in E.

Justify your answer.

[12 marks]

(3) If $\sum_n |a_{n+1} - a_n|$ converges, then show that $\{a_n\}_n$ is convergent. Is the converse true? Justify your answer.

[4 marks]

(4) Let $f : [0,1] \to \mathbb{R}$ be a twice differentiable function such that $f(0) = 0 = f(1)$ and $f'' + 2f' + f \geq 0$. Prove that $f \leq 0$. (*Hint: Consider* $g(x) = e^x f(x)$.)

[4 marks]

(5) If $f : (a,b) \to \mathbb{R}$ is an injective continuous map, then show that f is strictly monotone.

[8 marks]

(6) Let $f : \mathbb{R} \to \mathbb{R}$ be continuous in a neighbourhood of $c \in (a,b)$. Define $g : \mathbb{Q}\setminus\{0\} \to \mathbb{R}$ by $g(r) = \frac{f(c+r)-f(c)}{r}$ for all $r \in \mathbb{Q}$. Prove that $\lim_{r\to 0} g(r)$ exists if and only if f is differentiable at c.

[8 marks]

(7) Let the function $f : (a,b) \to \mathbb{R}$ have continuous second derivative. **Use Taylor's expansion to prove** $f''(x) = \lim_{h\to 0} \frac{f(x+h)-2f(x)+f(x-h)}{h^2}$ for all $x \in (a,b)$.

[8 marks]

(8) Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function satisfying $\lim_{x\to\infty} f'(x) = c \in \mathbb{R}$. **Use Mean Value Theorem to prove** $\lim_{x\to\infty} \frac{f(x)}{x} = c$.

[12 marks]

Indian Statistical Institute

First Semestral Back Paper Examination (2013-2014)

B. STAT I: Vectors and Matrices –I

Date: 29 ·01 ·2014 **Maximum Marks: 100** **Time: 3 Hrs.**

1. Let $\mathcal{V} = \{(\alpha,\beta): \alpha > 0 \text{ and } \beta > 0\}$. Define vector addition and scalar multiplication in $\mathcal{V}$ as follows.
(1) $(\alpha_1,\beta_1) + (\alpha_2,\beta_2) = (\alpha_1\alpha_2,\beta_1\beta_2)$ for every (α_1,β_1) and (α_2,β_2) in $\mathcal{V}$.
(2) $\delta(\alpha,\beta) = (\alpha^\delta,\beta^\delta)$ for every δ in $\mathcal{R}$ and (α,β) in $\mathcal{V}$.

   Show that $\mathcal{V}$ is a vector space over the field $\mathcal{R}$ of real numbers. What is its dimension? 12

2. Let $V = \{(x_1,x_2,x_3,x_4,x_5) \in R^5: x_1 - 2x_2 + 3x_3 - x_4 + 2x_5 = 0\}$. Show that $S = \{(0,1,1,1,0)\}$ is a linearly independent subset of V. Extend S to a basis for V. 10

3. Let $S = \{u_1, u_2, \dots, u_n\}$ be a finite set of vectors. Prove that S is linearly dependent if and only if $u_1 = \theta$, the null vector, or $u_{k+1} \in span(\{u_1, u_2, \dots, u_k\})$ for some $k(1 \le k < n)$. 10

4. Show that the cosets $\alpha + W$ and $\beta + W$ are equal if and only if $\alpha - \beta \in W$, where W is subspace of a vector space V. 6

5. Let $\mathcal{T}$ be a linear map from V to W. Show that this $\mathcal{T}$ is surjective if and only if $\mathcal{T}^t$ is injective. 6

6. Let $\mathcal{V}$ be the set of all complex numbers regarded as a vector space over the field of real numbers. Let $\mathcal{T}$ be the map from $\mathcal{V}$ to $\mathcal{V}$ defined by $\mathcal{T}(x + iy) = x - iy$. Show that $\mathcal{T}$ is an isomorphism of $\mathcal{V}$. Determine $\mathcal{T}^{-1}$. 6+6

7. Obtain the rank factorization of the matrix
$$\begin{bmatrix} -1 & 2 & 4 \\ 2 & -1 & 2 \\ 0 & 3 & 10 \end{bmatrix}.$$
10

8. Find a non-singular matrix P such that P^tAP is a diagonal matrix where
$$A = \begin{bmatrix} 2 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 10 \end{bmatrix}$$
12

9. Let V and W be finite dimensional vector spaces over a field F with $\dim V = n$ and $\dim W = m$. Let $\mathcal{T}$ be a linear map from V to W. Relative to two different pairs of ordered bases of V and W, let A and C be two matrix representations of $\mathcal{T}$. Show that there exists non-singular matrices P and Q such that $C = P^{-1}AQ$. 12

P.T.O.

10. State and prove the Frobenius inequality for ranks of matrices. 10

INDIAN STATISTICAL INSTITUTE, KOLKATA

FINAL EXAMINATION: FIRST SEMESTER 2013 -'14

B.STAT I YEAR

Subject	:	**Analysis I**
Time	:	3 hours
Maximum score	:	65

Instructions:

- *Justify every step in order to get full credit of your answers, stating clearly the result(s) that you use. Points will be deducted for missing arguments. Partial credit will be given for your approach to the problem.*
- *Switch off and deposit your mobile phones to the invigilator during the entire examination.*
- *Total marks carried by the questions turns out to be 130 which is more than 65. The total marks obtained will be halved.*

(1) Let $\{q_n\}_{n\in\mathbb{N}}$ be an enumeration of the rationals. Define $f:\mathbb{R}\to\mathbb{R}$ by

$$f(x)=\sum_{\{n\in\mathbb{N}\,:\,q_n<x\}}\frac{1}{2^n}\quad\text{for } x\in\mathbb{R}.$$

Prove that f is continuous at every irrational and discontinuous at every rational.

[20 marks]

(2) Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. What is the relation between

$$\limsup_{n\to\infty} a_n \quad\text{and}\quad \lim_{n\to\infty}\left[\sup\{a_k : k\geq n\}\right]?$$

Justify your answer.

[15 marks]

(3) Show that $\sum_{n=1}^{\infty}\left(\frac{a}{n}+\frac{1}{n+1}\right)$ (where $a\in\mathbb{R}$) is convergent **only when** $a=-1$.

[15 marks]

(4) (a) Let ξ be irrational and $S=\{m+n\xi : m,n\in\mathbb{Z}\}$.

(i) Show that the map f taking $n\in\mathbb{Z}$ to $f(n)=n\xi-\lfloor n\xi\rfloor$ is a one-to-one map from $\mathbb{Z}$ to $S\cap(0,1)$ where $\lfloor\cdot\rfloor$ denotes the greatest integer function.

(ii) Show that 0 is a limit point of S. (*Hint: Use part (i).*)

(iii) Show that S is dense in $\mathbb{R}$.

(b) Show that the set of all subsequential limits of the sequence $\{\sin n\}_{n\in\mathbb{N}}$, is dense in $[-1,1]$. (*Hint: Use part (a) for an appropriate* ξ.)

[30 marks]

(5) Prove that a surjective continuous map $f:\mathbb{R}\to\mathbb{R}$ which takes any value at most twice, must be bijective.

[15 marks]

(6) Let $f:[0,1]\to\mathbb{R}$ be a differentiable function such that there does **not** exist any $x\in[0,1]$ satisfying $f(x)=0=f'(x)$. Prove that $f^{-1}\{0\}$ is a finite set.

[10 marks]

(7) Show that there does **not** exist any differentiable function $f : \mathbb{R} \to (0, \infty)$ such that $f' = f \circ f$. (*Hint: Suppose not. First show that* $f'(x) > f(0)$ *for all* x *and then use the fact to arrive at a contradiction.*)

[15 marks]

(8) Let f be a function with continuous third derivative on $[0, 1]$ such that $f(0) = f'(0) = f''(0) = f'(1) = f''(1) = 0$ and $f(1) = 1$. Prove that there exists $c \in [0, 1]$ such that $f'''(c) \geq 24$. (*Hint: Consider Taylor expansions about both* 0 *and* 1*, and then evaluate them at* $\frac{1}{2}$.)

[10 marks]

INDIAN STATISTICAL INSTITUTE
First Semestral Examination: 2013-14

Course Name: B. STAT. I YEAR

Subject Name: Introduction to Programming and Data Structure

Date: 22. 11. 2013

Maximum Marks: 100 Duration: 3 hours

Answer as much as you can.

1. a) Consider the following 'C' function that expects a non-negative integer argument:

```
unsigned int magic ( unsigned int n )
{
     unsigned int m;
     m = 0;
     while (n > 0) {
     m = (m*10) + (n%10);
     n = n/10;
     }
 return m;
}
```

What does magic(34243) return? Describe the return value of magic as a function of the input argument n. (2+4 = 6)

b) Write a C program that scans a positive integer $n > 2$ and prints the largest proper divisor m of n. Example: For $n = 60$, your program should print $m = 30$. For $n = 61$, your program should print $m = 1$. (10)

c) What integer value is printed by the following 'C' program? (4)

```
int main ()
{
     int A[5] = {4,8,12,16,20}, *p;
     p = A + A[0];
     printf("%d", *p);
}
```

2. a) What will be the output of the following program and why? (4)

```
struct node {
                int data;
                struct node *next;
                }
main( )
{
      struct node N1, N2, N3;
      N1.data = 1; N2.data = 20; N3.data = 200;
      N1.next = &N2; N2.next = &N3; N3.next = &N1;
```

```
        printf("%d,%d"), N2.next -> data, N2.next -> next -> data);

}
```

b) Consider the type definitions given below. Suppose that a linked list is made up of nodes of type struct node. The last node points to NULL.

```
struct node {
                int key;
                struct node * next;
                };
typedef struct node *link;
```

i) Write a function with prototype void *second_delete*(link head) to delete the second node of the list, given "head" which points to the first element of the list. Assume that there are at least two nodes on the list.

ii) Write a function with prototype int count (link head) that takes the head of a linked list as input, and returns the number of nodes in the linked list. (8+8 = 16)

3. a) Give a suitable *typedef* for representing a stack of integers using a linked list. (2)

b) For the stack in part (a), write C functions for implementing the following stack operations:

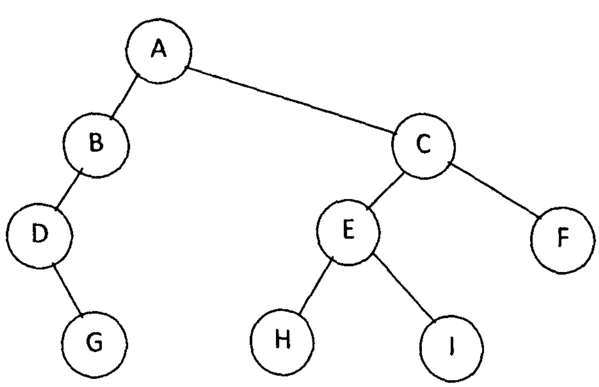
`init_stack`	- which constructs and returns an empty stack.
`isempty`	- which returns a value indicating whether a given stack is empty or not.
`push`	- which pushes a given integer on to a given stack.
`top`	- which returns the top element of a given stack and deletes the top element.
`delete`	- which deletes the top element of a given stack and returns the head pointer.

(2+2+3+2+3)

c) Write an iterative C function which takes an unsigned integer and prints its representation to the base 7 using a stack. The function should use the data type defined by you for representing a stack and only the functions defined in part (b). (6)

4. a) Prove that a full binary tree with n leaves contains $2n$-1 nodes. (6)

b) Give the sequence of the nodes of the following tree when the nodes are traversed in preorder and postorder manner. (2+2 = 4)



c) Construct a binary search tree from the following input sequence:

14, 15, 4, 9, 7, 18, 3, 5, 16, 4, 20, 17, 9, 14, 5.

What kind of traversal of the binary search tree retrieves the integers in ascending order? Assuming the following *typedef* for a tree representation, write down a recursive C function for such tree traversal.

```
struct nodetype {
                        int info;
                        struct nodetype *left;
                        struct nodetype *right;
                }
typedef struct nodetype *NODEPTR;
```

(5+1+4 = 10)

5. a) Determine the time complexity of the following iterative function and express in big Oh notation: (5)

```
int f ( int A[SIZE][SIZE], int n )
{
        int i, j, sum = 0;
        for (i=0; i<n; ++i)
        {
                if (i % 2 == 0)
                for (j=0; j<=i; j = j+1) sum = sum + A[i][j];
                else
                for (j=n-1; j>=i; j=j-1) sum = sum - A[i][j];
        }
}
```

b) Explain the use of the following C library functions with one example use for each: (3×4 = 12)

i) fscanf
ii) feof
iii) strcat
iv) strcmp

c) Can you use the *fprintf* function to print a string on your monitor screen or display unit? Explain briefly. (3)

6. a) The mode of an array of numbers is the number *m* in that array that is repeated most frequently. If more than one number is repeated with equal maximum frequencies, there is no mode. Write a C function that accepts an array of numbers and returns the mode or an indication that the mode does not exist. (10)

b) Write a C function that accepts a real number *x* and computes the cosine of *x* by summing up the following series up to 4 decimal places:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots$$

Do not use any math library functions.

(10)

7. a) What will be printed upon execution of the following line and why? (4)

```
printf("%d", printf("\n Hello World"));
```

b) Suppose we define a two-dimensional array in the following fashion:

int (*B)[100];

How can you assign memory dynamically to this array so that it may contain 30 rows? (3)

c) Consider the following declaration: int (*A)[20]; If A points to the memory location x, which memory location does A+1 point to? Assume that sizeof (int) = 4. (3)

d) With example, discuss the differences between the *break* and *continue* statements in context to loop termination in C. (5)

e) What will be the output of the following program? If you feel that there is any compile time or run time error, point out the same. Justify your answer. If your answer is yes, then can you correct this error by doing a minimal change in the program? (3+2 = 5)

```
#include<stdio.h>
main( )
        {
                int a[5]={1,2,3,4,5};
                for(i=0;i<5;i++) printf("%d\n", *a++);
        }
```

INDIAN STATISTICAL INSTITUTE

Semestral Examination, 1st Semester, 2013-14

Statistical Methods I, B.Stat I

Total Points 108. Maximum you can score is 100.

Date: 20.11.2013 Time: 3 hours

1. Suppose that the random variables X and Y have the following pdf:

$$f(x) = \begin{cases} 8xy & \text{for } 0 \le x \le y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Also let $U = X/Y$, and $V = Y$.

   (a) Are X and Y independent?

   (b) Determine the joint pdf of U and V.

   (c) Are U and V independent?

   (d) Find the probability density function of U.

[6+3+4+5=18]

2. Suppose that $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$, and $Z_i \sim N(0, 1)$, $i = 1, 2, \ldots, k$, and all variables are independent. State the distribution of each of the following variables with the appropriate parameters.

   (a) $X_1 - X_2$.

   (b) $X_2 + 2X_3$.

   (c) $\dfrac{X_1 - X_2}{\sigma S_Z \sqrt{2}}$, where S_Z represents the samples standard deviation based on the Z observations (with divisor $k-1$).

   (d) $\dfrac{\sqrt{n}(\bar{X} - \mu)}{\sigma S_Z}$

   (e) $Z_1^2 + Z_2^2$.

   (f) $\dfrac{Z_1}{\sqrt{Z_2^2}}$.

(g) $\dfrac{Z_1^2}{Z_2^2}$.

(h) $\dfrac{\sqrt{nk}(\bar{X} - \mu)}{\sigma\sqrt{\sum_{i=1}^{k} Z_i^2}}$

(i) $\dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} + \sum_{i=1}^{k}(Z_i - \bar{Z})^2$

(j) $\dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} + \sum_{i=1}^{k} Z_i^2$

(k) $\dfrac{\bar{X}}{\sigma^2} + \dfrac{\sum_{i=1}^{k} Z_i}{k}$

(l) $k\bar{Z}^2$.

(m) $\dfrac{(k-1)\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)\sigma^2\sum_{i=1}^{k}(Z_i - \bar{Z})^2}$

[1 + 1 + 2 + 2 + 1 + 1 + 1 + 2 + 2 + 2 + 2 + 1 + 2=20]

3. It is known that 30 percent of small steel rods produced by a standard process will break when subjected to a load of 3000 pounds. In a random sample of 50 similar rods produced by a new process, it was found that 21 of them broke when subjected to a load of 2000 pounds. Test the hypothesis that the breakage rate for the new process is the same as the rate for the old process. [12]

4. Suppose that a box contains a large number of chips of three different colours, red, green and blue, and that it is desired to test the null hypotheses H_0 that chips of the three colours are present in equal proportions against the alternative hypothesis H_1 that they are not present in equal proportions. Suppose that three chips are drawn at random from the box with replacement, and that H_0 is to be rejected if and only if at least two chips are of the same colour.

   (a) Determine the size of the test.

   (b) Determine the power of the test if 1/7 of the chips are red, 2/7 are green and 4/7 are blue.

[7+8=15]

5. (a) Suppose that the random variables X_1 and X_2 are independent, and that each has a normal distribution with mean μ and variance σ^2. Prove that the random variables $X_1 + X_2$ and $X_1 - X_2$ are independent.

(b) Suppose that the random variables X_1 and X_2 are independent, and that each has a normal distribution with mean 0 and variance σ^2. Determine the value of

$$P\left(\frac{(X_1 + X_2)^2}{(X_1 - X_2)^2} < 4\right).$$

[6+6=14]

6. Suppose that the random variables X_1, X_2 and X_3 are i.i.d. and that each of them has a standard normal distribution. Also suppose that

$$\begin{aligned} Y_1 &= 0.8X_1 + 0.6X_2, \\ Y_2 &= \sqrt{2}\,(0.3X_1 - 0.4X_2 - 0.5X_3)\,, \\ Y_3 &= \sqrt{2}\,(0.3X_1 - 0.4X_2 + 0.5X_3)\,. \end{aligned}$$

Find the joint distribution of Y_1, Y_2 and Y_3. [14]

7. Let the variables $Y_1, Y_2, \ldots, Y_n$ be defined as $Y_i = \beta x_i + \epsilon_i$, $i = 1, \ldots, n$, where $x_1, x_2, \ldots, x_n$ are known constants and ϵ_i are independently distributed as normal with mean 0 and variance σ^2.

(a) Derive the least square estimator $\hat{\beta}$ of β based on (x_i, Y_i), $i = 1, 2, \ldots, n$.

(b) Show that $\hat{\sigma}^2 = \sum(Y_i - \hat{\beta}x_i)^2/(n-1)$ is an unbiased estimator of σ^2.

[7+8=15]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2013-14 (First Semester)
Bachelor of Statistics (B. Stat.) I Year
Probability Theory I

Date: 18/11/2013 **Maximum Marks: 70** **Duration: 10:30 am - 02:00 pm**

- **Please write your roll number on top of your answer paper.**
- **Justify ALL your steps. You can use any result proved in the class but you should clearly mention the result that you are using.**
- **You are NOT allowed to use class notes, books, homework solutions, list of theorems, formulas etc. If you are caught using any, you will get a zero grade in the course.**

1. Let N_0 be the number of empty poles when r flags of different colours are displayed randomly on n poles arranged in a row (here $r, n \in \mathbb{N}$). Assuming that there is no limitation on the number of flags on each pole, compute the expectation and variance of N_0. In the same setup, let N_1 denote the number of poles with exactly one flag on it. Calculate the covariance between N_0 and N_1. [(3 + 5) + 5 = 13]

2. Assume that $1 \leq k < m < n$ are positive integers and $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables following geometric distribution with parameter $p \in (0, 1)$. For all $j \geq k$, define $I_j = \{(i_1, i_2, \ldots, i_k) \in \mathbb{N}^k : 1 \leq i_1 < i_2 < \cdots < i_k \leq j\}$ and

$$S_j = \sum_{(i_1, i_2, \ldots, i_k) \in I_j} \ln\left(2 + X_{i_1}^3 + X_{i_2}^3 + \cdots + X_{i_k}^3\right).$$

   Show that S_m/S_n has finite mean and compute its mean. [2 + 5 = 7]

3. Suppose that the initial number M of bacteria in a bacteria colony follows binomial distribution with parameters $n = 100$ and $p = 0.5$. Assume that the bacteria behave independently of each other, they do not replicate and each of them die within an hour with probability 0.8 independently of M. Let Z denote the number of surviving bacteria in the colony after one hour.

   (a) Calculate $E(s^Z|M)$, where $s \in \mathbb{R}$. [5]

   (b) Using (a) or otherwise, find the probability mass function of Z. [5]

4. Your teachers Dr. A. Basu and Dr. S. Ghosh decide to play the following game. They toss a fair coin. If head appears, then Dr. Basu gives one rupee to Dr. Ghosh and if tail appears, Dr. Ghosh gives one rupee to Dr. Basu. They continue this game till the 1000^{th} toss and they both end up with the equal amount of money in their pockets. Assuming that the outcomes of all the tosses are independent, and Dr. Ghosh had Rs. 1000 and Dr. Basu had Rs. 1600 in their pockets before the game started, calculate the probability that Dr. Ghosh never had more money in his pocket than Dr. Basu during the entire game. [10]

5. Other evaluations: homework assignments, presentations and quizzes. [5 + 10 + 15 = 30]

First Semester Examination (2013-14)
B. Stat – 1 yr
Remedial English
100 Marks
1 ½ hours

Date : 12.11.2013

1. Write an essay *on any one* of the following topics. Five paragraphs are expected.
    a. Pollution
    b. An Unforgettable Character
    c. School Days

(60 marks)

2. Fill in the blanks with appropriate prepositions :

He was a professor ______ English and an expert ______ Slavic languages. I met him ______ London ______ a conference. He was ______ a white suit and ______ lunch preferred vegetarian food ______ the non-vegetarian fare. ______ enquiry I found he had diverse interests ranging ______ astronomy ______ biochemistry. The conference lasted ______ two days ______ ______ the last day was set aside ______ interaction.

I stayed back. He left ______ Bath, his home town, but before he left I thanked him ______ ______ ______ the team ______ being the life and soul ______ the meet.

(20 marks)

3. Fill in the blanks with appropriate words :

A letter is a substitute for ______ contact. One ______ ______ ______ instead of being ______ ______ or ______ ______ ______ on the telephone. It ______ ______ a way of making ______ that ______ ______ ______ record of ______ which ______ proof. This why a business telephone call is ______ followed by ______ ______ of information.

(20 marks)

Indian Statistical Institute, Kolkata

Mid-semester examination: First semester 2013 -'14

B.Stat I year

Subject	:	**Analysis I**
Time	:	2 hours
Maximum score	:	30

06·09·13

Instructions:

- *Justify every step in order to get full credit of your answers, stating clearly the result(s) that you use. Points will be deducted for missing arguments. Partial credit will be given for your approach to the problem.*
- *Switch off and deposit your mobile phones to the invigilator during the entire examination.*
- *Total marks carried by the questions turns out to be 70 which is more than 30. The total marks obtained will be multiplied with $\frac{3}{7}$.*

(1) Let $E \subset \mathbb{R}$ and $f : E \to \mathbb{R}$ be a map. Set

$$\Delta_{\varepsilon,a} = \{\delta > 0 : N_\delta(f(a)) \supset f(E \cap N_\delta(a))\} \text{ for } \varepsilon > 0,\ a \in E, \text{ and}$$

$$A_{n,\delta} = \left\{a \in E : E \cap N_\delta(a) \subset f^{-1}(N_{\frac{1}{n}}(f(a)))\right\} \text{ for } n \in \mathbb{N},\ \delta > 0$$

where $N_r(x)$ denotes the interval $(x - r, x + r)$.

(i) If $\bigcup\limits_{\varepsilon>0} \bigcap\limits_{a\in E} \Delta_{\varepsilon,a} = (0, \infty)$, then what is the best that you can say about f without any extra assumption on the hypothesis?

(ii) If $\bigcap\limits_{n\in\mathbb{N}} \bigcup\limits_{\delta>0} A_{n,\delta} = E$, then what is the best that you can say about f without any extra assumption on the hypothesis?

Justify your answers with proper arguments.

[(7+7) marks]

(2) (i) If $f(x) = 7x^2+4x+5$ for $x \in \mathbb{R}$, then prove that $\lim\limits_{x\to 2} f(x)$ exists from the definition.

(ii) If $g(x) = \frac{2x^2-3x-2}{|x^2-x-2|}$ for $x \in \mathbb{R} \setminus \{2\}$, then show that $\lim\limits_{x\to 2} g(x)$ fails to satisfy the definition of its existence. (*Hint: Suppose* $\lim\limits_{x\to 2} g(x) = l$. *Then, establish contradictions by considering two cases:* $l \geq 0$ *and* $l < 0$.)

[(7+8) marks]

(3) Let $f : X \to Y$ be a map where $X, Y \subset \mathbb{R}$. Then, prove that

$$f \text{ is continuous} \iff f^{-1}(A^o) \subset (f^{-1}(A))^o\ \forall A \subset Y$$

where E^o denotes the interior of E.

[8 marks]

(4) If E is a compact subset of $\mathbb{R}$, then prove that E has a countable dense subset.

[6 marks]

(5) (i) Prove that $A_x = \{(x, y) : y \in \mathbb{R}\}$ and $B_y = \{(x, y) : x \in \mathbb{R}\}$ are connected subsets of $\mathbb{R}^2$ for all $x, y \in \mathbb{R}$.

(ii) Let $\{X_i\}_{i\in I}$ be a collection of connected subsets of $\mathbb{R}$ such that $X_i \cap X_j \neq \emptyset$ for $i, j \in I$. Prove that $\bigcup\limits_{i\in I} X_i$ is connected.

(iii) Let $(x_0, y_0) \in \mathbb{R}^2$. Show that $\mathbb{R}^2 \setminus \{(x_0, y_0)\}$ is connected. (*Hint: Use parts (i) and (ii)*)

(iv) Prove that there does not exist any continuous injective map from $\mathbb{R}^2$ to $\mathbb{R}$. (*Hint: Use part (iii)*)

[(4+10+4+9) marks]

Indian Statistical Institute

First Semestral Examination (2013-2014)

B. STAT I: Vectors and Matrices –I

Date: 11.11.2013 **Maximum Marks: 60** **Time: 3 Hrs.**

This paper contains questions for 72 marks. The maximum you can score is 60.

1. Let $\mathcal{R}$ be the set of all real numbers. Consider $\mathcal{R}$ as a vector space over the field $\mathcal{Q}$ of rational numbers. Show that $\sqrt{2}$ and $\sqrt{3}$ are linearly independent. 5

2. If f_1 and f_2 are two linear functionals on a vector space $\mathcal{V}$ over a field F satisfying $f_1(x) = 0$ whenever $f_2(x) = 0$ for $x \in \mathcal{V}$, show that $f_1 = \alpha f_2$ for some α in F. 6

3. $\mathcal{T}$ is a transformation from V to W. Show that for existence of a right inverse of $\mathcal{T}$, it is necessary that $\mathcal{T}$ is surjective. If $\mathcal{T}$ is surjective, show that any right inverse of $\mathcal{T}$ is injective. 3+3

4. Let $\mathcal{T}$ be a linear map from V to W. Show that there always exists a linear map S from W to V, such that $\mathcal{T}S\mathcal{T} = \mathcal{T}$. 5

5. Let $\mathcal{T}$ be a linear map from V to V. Show that V is the direct sum of $R(\mathcal{T}^k)$, the Range Space and $N(\mathcal{T}^k)$, the Null Space, for some positive integer k. 6

6. Show that a subset S of the vector space $\mathcal{F}^n$ is a subspace of $\mathcal{F}^n$ if and only if S is the null space of a matrix, where , $\mathcal{F}$ is the field of real numbers or complex numbers. 6

7. Let the reduced row echelon form of A be

$$\begin{bmatrix} 1 & 0 & 2 & 0 & -2 \\ 0 & 1 & -5 & 0 & -3 \\ 0 & 0 & 0 & 1 & 6 \end{bmatrix}.$$

Determine A if the first, second and fourth columns of A are

$$\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}.$$

9

8. Find a non-singular matrix P such that $P^t AP$ is the normal form of A under congruence, where

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

9

9. Let V and W be finite dimensional vector spaces over a field F with $dimV = n$ and $dimW = m$ and A be a $m \times n$ matrix over F. Relative to two different pairs of ordered bases of V and W, let A represent two linear maps $\mathcal{T}_1$ and $\mathcal{T}_2$ from V to W. Show that there exists non-singular linear maps $S \in L(V, V)$ and $R \in L(W, W)$ such that $\mathcal{T}_1 = R^{-1}\mathcal{T}_2 S$.

9

10. Suppose that the augmented matrix of a system $Ax = b$ is transformed into a matrix$(A'|b')$ in reduced row echelon form by a finite sequence of elementary row operations. Show that $Ax = b$ is consistent if and only if $(A'|b')$ contains no row in which the only nonzero entry lies in the last column. 5

11. Let A and B be two matrices of orders $m \times n$ and $n \times s$ respectively. Then show that $Rank\ (AB) = Rank(B)$ if and only if $N(A) \cap R(B) = \{0\}$, where $N(A)$ and $R(B)$ denote the null space of A and the range of B. 6

Indian Statistical Institute
Mid-Semestral Examination: 2013-14
B. Stat (Hons.) First Year
Vectors & Matrices –I

Date: **02-09-2013** Maximum Marks: **40** Duration: **2Hrs.**

1. Find a linear map $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ such that the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ of $\mathbb{R}^3$ is mapped to itself, T not being the identity map. (4)

2. Let $f(x)$ be a polynomial of degree n with real coefficients. Prove that the $n + 1$ polynomials $f(x), f^{(1)}(x), \dots, f^{(n)}(x)$ are a basis of $\mathcal{P}_n(\mathbb{R})$, where, $f^{(n)}(x)$ is the n-th derivative of $f(x)$ and $\mathcal{P}_n(\mathbb{R})$ is the vector space of polynomials in $\mathbb{R}$ of degree less than or equal to n. (4)

3. Let $S: V \longrightarrow W$ and $T: W \longrightarrow V$ be linear maps between finite-dimensional vector spaces over F. Suppose that TS is the identity map on V. Prove that T is surjective (onto), S is injective (1-1) and dim $V \leq$ dim W. (4)

4. Let $V = \mathbb{R}^4$ and W be a subspace of V generated by the vectors (1,0,0,0) and (1,1,0,0). Find a basis of the quotient space V/W. (4)

5. Let $V = \mathbb{R}^2$ and define f, g $\in$ V', the dual space, as follows:
f(x, y) = x + y and g(x, y) = x - 2y.
Find a basis $\{v_1, v_2\}$ for V such that its dual basis is {f, g}. (4)

6. Let $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be given by $T(x, y) = (3x + 2y, x)$ with respect to the standard basis. Let $f \in (\mathbb{R}^2)'$ be given by $f(x, y) = 2x + y$. Compute $T^t(f)$, where T^t is the transpose. Find also the matrix of T^t with respect to the dual bases. (6)

7. Let $v_1 = (1,0,-1,2)$ and $v_2 = (2,3,1,1)$ and W be the subspace of $\mathbb{R}^4$ spanned by v_1 and v_2. Describe the subspace W^0, where W^0 is the annihilator of W. (4)

8. If W is a subspace of V and $x \notin W$, prove that there exists f $\in W^0$ such that f(x)$\neq$0. (6)

9. Let V be vector space over F such that it contains at least three elements. Let $w \in V$. Then show that V is the linear span of S, where, $S = \{v \in V | v \neq w\}$. (7)

10. Let V and W be finite dimensional vector spaces with dim V = n and dim W = m. Let $T: V \longrightarrow W$ be a linear transformation of rank r. Then show that there exists a basis $\{v_1, \dots, v_n\}$ of V and a basis $\{w_1, \dots, w_m\}$ for W such that $T(v_i) = w_i$ for $i = 1, 2, \dots, r$ and $T(v_i) = \theta'$, the zero element of W, for $i = r + 1, \dots, n$. Hence, find the matrix of the transformation with respect to these bases. (7)

INDIAN STATISTICAL INSTITUTE

Midsemester Examination : (2013-2014)

B. Stat 1st Year

Statistical Methods I

Date: 30. 08. 2013 Maximum marks: 100 Duration: 3 hours.

1. Explain the meaning of the following terms, clearly pointing out the differences between them: (i) randomized study, (ii) observational study. Why can causal inference not be drawn from observational studies? [10]

2. Given n random realizations $x_1, \ldots, x_n$ of a random variable X, show that the difference between the sample mean and the sample median never exceeds the sample standard deviation of the variable (with divisor n). [10]

3. In class we have talked about the raw moment and the central moment. In practice, there is no conceptual difficulty in determining the moment against any value of interest (other than the mean). Thus, given the sample $X_1, \ldots, X_n$, the rth moment about the value a is:

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - a)^r.$$

Now suppose that the first four moments of a distribution about the value 4 (i.e. $\frac{1}{n}\sum_{i=1}^{n}(X_i - 4)^r, r = 1, \ldots, 4$) are -1.5, 17, -30 and 108, respectively. Find the first four raw and central moments. [10 + 10 = 20]

4. Suppose that the probability distribution of the number of customers entering a store by time t has a Poisson distribution with parameter λt. (For definiteness, we consider the time when the store opens to be $t = 0$). Show that the time of arrival of the first customer has an exponential distribution with mean $1/\lambda$. [15]

[The probability mass function of the Poisson random variable with parameter θ is $f(x) = \frac{e^{-\theta}\theta^x}{x!}$, $x = 0, 1, \ldots$, and the probability density function of an exponentially distributed random variable is $f(x) = \frac{1}{\mu}e^{-x/\mu}$, $x \geq 0$.]

P.T.O

5. Suppose that thirty six heat lamps are connected in a greenhouse so that when one lamp fails the next one takes over immediately (only one lamp is turned on at one time). The lamps operate independently, and each has a mean lifetimie of 50 hours when turned on, and a standard deviation of 4 hours. If the greenhouse is not checked for 1850 hours after the lamp system is turned on, what is the probability that the lamp will be burning when it is inspected at the end of the 1850 hour period? [15]

6. Let $(x_1, y_1), \ldots, (x_n, y_n)$ represent n bivariate observations. We are interested in the regression of y on x.

   (a) Derive the least squares estimates of α and β, the intercept and the slope parameter of the regression line of y on x.

   (b) Show that $Var(Y) = r^2 s_y^2$, where r is the sample correlation coefficient between x and y, s_y^2 is the sample variance of y, and Y is the predicted value of y based on the fitted regression line in (a).

   (c) Show that $Cov(Y, e) = 0$, where $e = y - Y$ is the residual (error) in prediction.

   (d) Using part (c) or otherwise, show that $Var(e) = (1 - r^2)s_y^2$.

   (e) Show that $Corr(y, Y) = |r|$. [6 + 4 + 3 + 3 + 4 = 20]

INDIAN STATISTICAL INSTITUTE
Mid Semestral Examination: 2013-14

Course Name: B. STAT. I YEAR

Subject Name: Introduction to Programming and Data Structure

Date: 28. 08. 2013

Maximum Marks: 60 Duration: 2 hours 30 Min.

Answer as much as you can.

1. a) Let the function **f** be defined as follows:

```
int f ( int n )
{
int t;
t = 100000 * (n % 10) + (n / 10);
return t / n;
}
```

What is the value returned by the call **f(142857)**? Briefly explain your answer. (4)

b) The following function expects a non-negative integer argument.

```
unsigned int doit ( unsigned int n )
{
unsigned int i, s, t;
s = 0; t = 1;
for (i=0; i<n; ++i) {
s = s + i + n;
t = t * 2;
}
return s * t;
}
```

What value does **doit(4)** return? Describe as a mathematical function of **n** the value returned by the function **doit(n)**.

2+4 = 6

c) How many bytes are allocated to the pointer **p** after the following call?

```
#define MAXSIZE 100
p = (long int *)malloc(MAXSIZE * sizeof(long int));
```

(2)

d) What value will be stored in **count** after the execution of the following code and why?

```
int count, n = 100;
count = printf("\nn:%d\n",n);
```

(3)

2. a) Complete the following program that reads an integer $n > 2$ and prints the smallest integer $d > 2$ such that n is an integral multiple of d^2. If no such d exists, then print -1. For example, for n = 49, n = 50, and n = 51, your program should respectively print d = 7, d = 5, and d = −1. Do not make any function calls (including math library calls). Use built-in arithmetic and conditional operators only. Do not use any variables other than n, d, t. You may use t as a temporary variable. (10)

b) Suppose that an array A of size $n > 1$ is passed to the following function?

```
void someFunc ( int A[] , int n )
{
int i,j,k;
for (i=0; i<n; ++i) {
for (j=1; j<=i; ++j) A[j] -= A[j-1];
for (k=j; k<n/2; ++k) A[k] += A[k+1];
}
}
```

Which of the following is a proper estimate of the running time of the above function? Explain your choice.

(A) $O(n)$ (B) $O(n.\log n)$ (C) $O(n^2)$ (D) None of the above (5)

3. a) Divisibility of a number by 9 is defined recursively as follows: 0 and 9 are divisible by 9, any other number is divisible by 9 if and only if the sum of its digits is divisible by 9. You are required to fill up the parts of the code that are left blank so that the overall code tests whether the given number is divisible by 9.

```
#include <stdio.h>

int main ( )
{
  int num, digitSum;
  scanf ("%d", &num);                    // read num, assume num ≥0
                                         // reduce as per recursive definition, if necessary
  while (------------------------------------)
  {
                                         // find the sum of the digits of num
   ---------------------------- ;        // initialise
while (----------------)
 {                                       // digits remain

------------------------------------ ;   // add digit
-------------------------------------- ; // drop digit
 }                                       // end-while
                                         // prepare for next round of reduction
---------------------------- ;
 }                                       // end-while, reduction complete
                                         // now test the base cases
 if ( ------------------------------------ )
    printf ("given number is divisible by 9\n");
 else
    printf ("given number is not divisible by 9\n");
 return 0;
```

(8)

b) Write an iterative C function `void append (char *s, char *t)` which appends the string `t` immediately after the end of the string `s`. For example, if `s` and `t` are passed respectively as `"mumbo"` and `"humbo jumbo"`, then your function should change the string `s` to `"mumbohumbo jumbo"`.

Do not use any string library functions. (7)

4. a) Express $f(n)$ as a function of n, where f is defined as follows. Show your calculations. (5)

```
unsigned int f ( unsigned int n )
{
  unsigned int s = 0, i, j;
  for (i=1; i<=n; ++i)
  for (j=i; j<=n; ++j)
  s += i + j;
  return s;
}
```

b) Write a suitable **typedef** for representing ordered pairs of integers of the form $<x, y>$ as a structure. (2)

c) Suppose we would like to represent a real number $x \times 10^y$ as an ordered pair $<x, y>$, with the first member satisfying $1000 \le |x| \le 9999$. For example, 6.023×10^{23} is represented as $<6023, 20>$ whereas -1.6×10^{-19} is represented as $<-1600, -22>$. Write a C function which takes two real numbers represented as two ordered pairs as given above and returns an ordered pair representing the product of the pairs passed as parameters.

Do not use any math library functions. (8)

5. a) Given a polynomial of degree n,

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x^1 + a_0,\ a_n \neq 0,$$

it can be rewritten as:

$$p(x) = (((a_n x + a_{n-1})x + \ldots + a_1)x + a_0.$$

This is the Horner's scheme for evaluating $p(x)$. The advantage of this method of evaluation is that explicit exponentiation is avoided. Let the coefficients for the various powers of x of $p(x)$ be stored in an array (say) **P[ ]**, as float P[ ] = $\{a_n, a_{n-1}, \ldots, a_1, a_0\}$.

Write an iterative C function based on the above scheme to evaluate a polynomial. The function should be declared as:

hornerpoly(float P[ ], int n, float x) (6)

b) Write a recursive C function `char *lastOccur (char *s, char c)` which takes a string `s` and a character `c` as input; it returns a pointer to the last occurrence of `c` in `s`, or `NULL` if the character `c` does not occur in `s`. *Do not use any string library functions.* (6)

c) Briefly distinguish between structures and unions on the basis of their memory usage. (3)

6. a) Consider the following definition:

int **B;

Based on this definition, dynamically allocate memory for a two dimensional array that will have n rows and m columns where n and m are variables to be taken from the users. (4)

```
#include <stdio.h>
int t = 10;
main()
{
int x = 0;
void funct1();
funct1();
printf("After first call \n");
funct1();
printf("After second call \n");
funct1();
printf("After third call \n");
}
void funct1()
{
   static int y = 0;
   int z = 10;
   printf("value of y %d z %d",y,z);
   y=y+10;
}
```

(4)

c) A two-dimensional character array is used to store a list of names. Each name is stored in a row. An empty string indicates the end of the list. Complete the ***recursive*** function `printNames()` to print the names stored in the two-dimensional array supplied as **p**. (7)

```
void printNames ( char (*p)[100] )
{

--------------------------------------
--------------------------------------
--------------------------------------

}

void main()
{

char names[20][100] = { "Anurag", "Nilashis", "Anamitra", "Sarthak",
"Arnab", "Dheeraj", "Sagnik", ""};

printNames(names);
}
```

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2013-14 (First Semester)
Bachelor of Statistics (B. Stat.) I Year
Probability Theory I

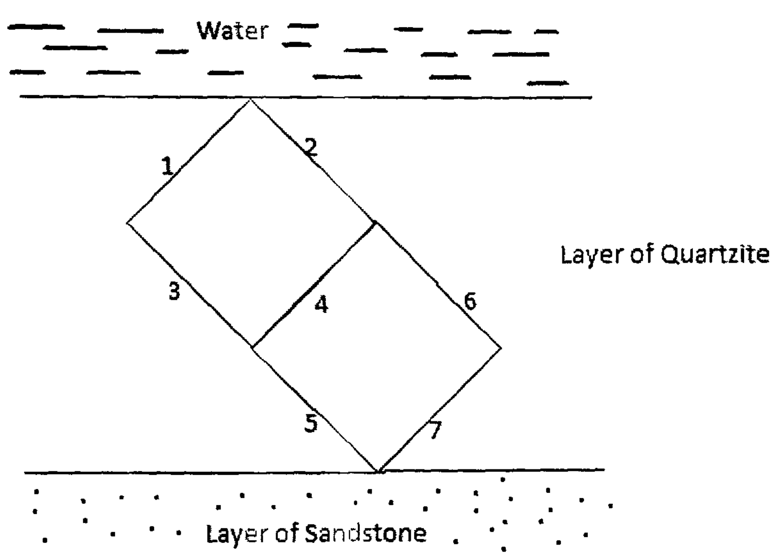
Teacher: Parthanil Roy

Date: 26/08/2013 **Maximum Marks: 30** **Duration: 10:30 - 12:30 pm**

Note:

- **Please write your roll number on top of your answer paper.**
- **There are three problems each carrying 10 marks with a total of 30 points. Solve as many as you can. Show all your works and write explanations when needed. Maximum you can score is 30 marks.**
- **You are NOT allowed to use class notes, books, homework solutions, list of theorems, formulas etc. If you are caught using any, you will get a zero grade in the course.**

1. Consider the following schematic diagram of a drainage network model, where each of Paths 1 - 7 (as shown in the figure below) behave independently of each other. Suppose that each of the odd-numbered paths is open with probability $p_o \in (0,1)$ and each of the even-numbered paths is open with probability $p_e \in (0,1)$.



Recall that water will be able to pass through a particular path if and only if it is open. If it is given that water has passed through the layer of quartzite to the layer of sandstone, calculate the probability that either Path 6 or Path 7 is open. [10]

2. A closet contains 13 different pairs of gloves. These twenty six gloves are randomly arranged into 13 pairs.

   (a) Find the probability that each pair contains a left glove and a right glove. [5]

   (b) Find the probability that all the gloves are paired correctly. [5]

3. There are two books, say A and B, and you would like to choose one of them randomly with probabilities $1/3$ and $2/3$, respectively. Suppose you have a coin for which the probability of obtaining a head in one toss is unknown but constant. Assuming that various tosses produce independent results, design an algorithm for the random selection mentioned above. Justify that your algorithm works. $[2+8]$

Wish you all the best