INDIAN STATISTICAL INSTITUTE

Mid Semester Examination, 1st Semester, 2010-11

Statistical Methods I, B.Stat I

Total Points 100

Date: September 01, 2010 Time: 3 hours

1. Consider a 2×2 contingency table, and suppose that we are interested in determining whether a relation exists between the two variables.

   (a) With respect to a suitable example, explain the difference between a prospective study and a retrospective study.

   (b) With respect to your example, define the odds ratio and interpret it.

   (c) With respect to your example or otherwise, explain why the analysis based on odds ratios is still meaningful when the data are collected through retrospective sampling, while the analysis based on proportions is not. [10+10+10=30 points]

2. Consider two groups of numerical observations of sizes n_1 and n_2 respectively. Let $\bar{x}_1$ and s_1^2 represent the mean and variance for group 1, while $\bar{x}_2$ and s_2^2 represent the corresponding quantities for the second group.

   (a) Let s^2 represent the variance of the combined group of $n_1 + n_2$ observations. Show that

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} + \frac{n_1 n_2}{(n_1 + n_2)^2}(\bar{x}_1 - \bar{x}_2)^2.$$

   (b) In a batch of 10 children, the I.Q. of a dull boy is 36 below the average of the other nine children. Show that the standard deviation of the IQ of the entire group cannot be smaller than 10.8. If this standard deviation is actually 11.4, determine what the standard deviation is when the dull boy is left out. [10+10=20 points]

3. Assuming that the height distribution of a group of men is normal, find the mean and standard deviation, if 84.13% of the men have heights less than 65.2 inches and 68.26% have heights lying between 65.2 and 62.8 inches.
[15 points]

4. Write down the simple linear regression model for Y on X with appropriate assumptions.

   (a) Find the least square estimates of the regression parameters.

   (b) Suppose that the error ϵ in the simple linear regression model has a normal distribution. Find the maximum likelihood estimates of the set of parameters under the normality assumption based on a bivariate sample of size n. Show that the maximum likelihood estimates of the regression parameters are the same as their least squares estimates.
[10+15 points]

5. Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ represent a sample of n bivariate observations of the variables X and Y. Let s_x^2, s_y^2 represent the variance of the X and Y observations respectively.

   Define two new variables U and V as $u_i = x_i + y_i$ and $v_i = x_i - y_i$. If $s_x^2 = s_y^2$, show that $r_{uv} = 0$, where r_{uv} represents the correlation coefficient between U and V.
[10 points]

Indian Statistical Institute, Kolkata

Analysis I B. Stat.-I Mid-semester Examination

2010-11 First semester

Maximum Marks: 40 Maximum Time: $2\frac{1}{2}$ Hrs.

---

(1) Give brief answers. [5×3]

(a) Find the limsup and liminf of the sequence $(x_n)_{n=1}^{\infty}$, where $x_n = (-1)^n(1+\frac{1}{n}), n \in \mathbb{N}$.

(b) Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers and $l = \liminf a_n$. Prove that for every $x > l$ and every $N \geq 1$, there exists $n \geq N$ such that $a_n < x$.

(c) Find the sum of the series $\sum\limits_{n=1}^{\infty} \frac{1}{n(n+3)}$.

(d) Let $\sum\limits_{n=1}^{\infty} x_n$ be a convergent series of positive terms and $(y_n)_{n=1}^{\infty}$ be a bounded sequence of non-negative terms. Show that the series $\sum\limits_{n=1}^{\infty} x_n y_n$ converges.

(e) Let $\sum\limits_{n=1}^{\infty} x_n$ be a convergent series of positive terms. Prove that $\sum\limits_{n=1}^{\infty} \frac{\sqrt{x_n}}{n}$ is convergent.

(2) Let $(a_n)_{n=1}^{\infty}$ be a sequence such that $a_n > 0$ for all $n \in \mathbb{N}$. Show that $\lim\limits_{n\to\infty} a_n = \infty$ if and only if $\lim\limits_{n\to\infty} \frac{1}{a_n} = 0$. [6]

(3) Suppose that the sequence $(a_n)_{n=1}^{\infty}$ satisfies the condition [5]

$$0 < a_n < 1, \quad a_n(1-a_{n+1}) > \frac{1}{4}, \quad n \in \mathbb{N}.$$

Use the arithmetic-geoemtric mean inequality to show that the sequence $(a_n)_{n=1}^{\infty}$ is convergent and then find its limit.

(4) Test the following series for convergence. [6]

(a) $\sum\limits_{n=1}^{\infty} \frac{(n!)^2}{(2n)!}$, (b) $\sum\limits_{n=1}^{\infty} 3^n(\frac{n}{n+1})^{n^2}$.

(5) Let $a \in \mathbb{R}$. Study the convergence and absolute convergence of the series [5]

$$\sum_{n=1}^{\infty}(-1)^n \sin\frac{a}{n}.$$

(6) Let S be the sum of the convergent series [6]

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \cdots$$

Find a rearrangement of this series which is divergent. Provide only first few terms of the rearranged series.

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2010-2011
B.Stat. (Hons.) 1st Year. 1st Semester
Probability Theory I

Date: November 26, 2010 **Maximum Marks: 75** **Duration: 3 and 1/2 hours**

• This question paper carries 78 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.

• Whenever applicable, you should (1) write clearly, explaining all your notations, a suitable sample space Ω for the answers, and (2) state, with adequate justification, assignment of probability to the sample points.

---

1. Consider a random experiment that results in one of m outcomes, denoted by $O_1, \ldots, O_m$, each with probability $1/m$. Suppose this experiment is conducted independently and indefinitely until each O_j appears once. It can be proved that the experiment terminates with probability 1. Find the distribution of X, the number of trials required. [11]

2. Suppose X is a random variable which is independent of itself. Show that $\exists\, a \in \mathbb{R}$ such that X is degenerate at a. [9]

3. Suppose an infinite sequence of random variables $\{X_i : i \geq 1\}$ is defined on a probability space $(\Omega, \mathcal{A}, P)$ such that for every $n \geq 1$, $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) Bin$(1, p)$ random variables, where $0 < p < 1$. We have proved (in class) that infinitely many 1's occur with probability 1. Introduce relevant random variables to establish the following: *given that there are r successes during the first n trials, the trials at which these successes occur constitute a random sample of size r (without replacement) from the "population" of possible positions.* [11]

4. Let X_1, X_2 be non-negative integer-valued i.i.d. random variables such that $p_k := P(X = k) > 0$ for every $k \geq 0$. Suppose

$$P(X_1 = n | X_1 + X_2 = n) = P(X_1 = n-1 | X_1 + X_2 = n) = \frac{1}{n+1}, \; n \geq 1.$$

Show that X_1, X_2 are Geometric(p) random variables for some $0 < p < 1$. [9]

P.T.O.

5. Let a random variable X have uniform distribution over $\{0, 1, \ldots, r-1\}$. Let r be a composite number, say, $r = ab$, where $1 < a, b < r$. Show that the distribution of X can be recognized as that of the sum of two independent random variables. [9]

6. Assume that the number of insect colonies in a certain area follows the Poisson distribution with parameter λ, and that the number of insects in a colony has a logarithmic series distribution with parameter p. [*A logarithmic series distribution with parameter* p, $0 < p < 1$, *is defined by the probability mass function* $f(x) := -p^x/[x \log(1-p)]$, $x = 1, 2, \ldots$.]

(a) Describe the set-up above using language and notations of compound distributions.

(b) Use (a) to find the distribution of the total number of insects in the area.

[4+9 = 13]

7. A population consists of r classes whose sizes are in the proportion $p_1 : p_2 : \cdots : p_r$. A random sample of size n is taken with replacement. Find the expectation and variance of the number of classes not represented in the sample. [8+8 = 16]

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

INDIAN STATISTICAL INSTITUTE LIBRARY
★ Need membership card to check-out documents
★ Books are not to be loaned to anyone
★ Loans on behalf of another person are not permitted
★ Damaged or lost documents must be paid for

Indian Statistical Institute

First Semester Examination (2010-2011)

B. Stat (First year)

Computational Techniques and Programming-I

Date : 29.11.10 Time: 3 hours

(Answer all questions)

1. (a) Give the algorithm/flowchart for producing a 3×3 Magic Square where the entire row sums, column sums, and diagonal sums are equal. You have to give the dry run of your algorithm.

   (b) (I) Give the corresponding C program of the above algorithm/flowchart with full documentation.

   (II) Give also the C program to test in all possible ways whether a given square is really a magic square.

   (8+(6+6)

2. (a) Explain with the help of 'C' code the various operations associated with "STACK' data structure.

   (b) Describe an algorithm which is associated with a suitable application of "STACK". Also give the corresponding C program.

   (5+(5+10))

3. (a) Distinguish between

   (I) Recursive algorithm and Iterative algorithm.

   (II) Subroutine call and Interrupt scheme.

   (b) What is a concurrent program? Distinguish between a procedure call and a process creation?

   (10+10)

P.T.O.

41

4. (A) In the theory of 2's complement arithmetic, illustrate all the following cases of two decimal number additions (Convert the decimal numbers into binary for your illustration).

(a) When both the numbers are **positive**.

(i) Normal addition without any overflow.

(ii) Abnormal addition with an overflow.

(b) When both the numbers are **negative**.

(i) Normal addition without any underflow.

(ii) Abnormal addition with an underflow.

(c) When one number is **positive and another** is negative.

(B) In the event of overflow and underflow how would it be notified within the computer for the purpose of creating an interrupt?

(C) Z is expressed as Boolean expression as the following:

$a'b'c + ab'c' + ab'c + abc' + abc$ (+ denotes OR logic)

(i) Simplify the expression to the minimum extent possible.

(ii) How many logic gates would you require to implement the minimum expression if you are given their input lines as non-complemented a, non-complemented b and non-complemented c.

(9+5+6)

5. Illustrate the following programming techniques to swap two integers variables (with appropriate "C" code)

(i) Using only two variables.

(ii) Using three variables.

(iii) Using "**swap**" function call.

(iv) Using **Exclusive-OR** Boolean logic.

(4+4+6+6)

INDIAN STATISTICAL INSTITUTE

Mid-semester Examination: 2010-2011

B.Stat. (Hons.) 1st Year. 1st Semester

Probability Theory I

Date: September 03, 2010 **Maximum Marks: 50** **Duration: 2 and 1/2 hours**

- This question paper carries 55 points. Answer as much as you can. However, the maximum you can score is 50.

- 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.

- Whenever applicable, you should (1) write clearly, explaining all your notations, a suitable sample space Ω for the answers, and (2) state, with adequate justification, assignment of probability to the sample points.

---

1. Let $A := \{p/2^q : p > 0, q \geq 0 \text{ are integers}\}$. Show that A is countable. [9]

2. In bridge, prove that the probability p of West's receiving exactly k aces is the same as the probability that an arbitrary hand of 13 cards contains exactly k aces. [9]

3. Consider a random experiment of distributing r *indistinguishable* balls among n cells and assume that all distinguishable arrangements have equal probabilities. Show that the probability that a group of m prescribed cells contains a total of exactly j balls is given by $[{}^{m+j-1}C_{m-1}\ {}^{n-m+r-j-1}C_{r-j}]/{}^{n+r-1}C_r$. [9]

4. Consider random arrangements of r_1 alphas and r_2 betas and assume that all arrangements are equally probable. Show that the probability of having exactly k runs of alphas is given by $[{}^{r_1-1}C_{k-1}\ {}^{r_2+1}C_k]/{}^{r_1+r_2}C_{r_1}$. [9]

5. Consider events A_r, $1 \leq r \leq n$, such that at least one of them is certain to occur, but certainly no more than two occur. If $P(A_r) = p$ for $1 \leq r \leq n$ and $P(A_r \cap A_s) = q$ for $1 \leq r \neq s \leq n$, show that $p \geq 1/n$ and $q \leq 2/n$. [9]

6. Two similar decks of N distinct cards each are matched simultaneously against a similar target deck. Find the probability u_m of having exactly m double matches. [10]

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

INDIAN STATISTICAL INSTITUTE

SECOND SEMESTER BACKPAPER EXAMINATION (2009–10)

B. STAT. I YEAR

ANALYSIS II

Date : 16/9/10 Maximum Marks : 100 Time : 3 hours

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

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

$$f_+(x) = \begin{cases} f(x) & \text{if} \quad f(x) > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and } f_-(x) = \begin{cases} -f(x) & \text{if} \quad f(x) < 0 \\ 0 & \text{otherwise} \end{cases}$$

Show that $f \in \mathcal{R}[a, b]$ if and only if both $f_+ \in \mathcal{R}[a, b]$ and $f_- \in \mathcal{R}[a, b]$. Moreover,

$$\int_a^b f(x)dx = \int_a^b f_+(x)dx - \int_a^b f_-(x)dx$$ [10]

(b) Let $f : [a, b] \to \mathbb{R}$ be a function with a continuous derivative. Show that f is the sum of a continuous increasing function and a continuous decreasing function. [7]

2. Show that $\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{\sqrt{n^2 + k^2}} = \log(1 + \sqrt{2})$. [10]

3. For various values of s, test the following improper integral for convergence

$$\int_2^\infty \frac{dx}{x(\log x)^s}$$ [10]

4. Let $\{f_n\}$ be a sequence of continuous functions which converges uniformly to a function f on an interval I. Prove that

$$\lim_{n\to\infty} f_n(x_n) = f(x)$$

for every sequence $\{x_n\} \subseteq I$ such that $x_n \to x \in I$.

Show that the conclusion may fail if the convergence is not uniform. [7 + 3 = 10]

5. Let $\{f_n\}$ be a sequence of continuous functions on $[0,1]$ decreasing pointwise to the constant function 0. Is it true that

$$\int_0^1 f_n(x)dx \longrightarrow 0 \text{ as } n \to \infty?$$

Briefly justify your answer. [8]

6. Let $f_n(x) = n^\alpha x(1-x^2)^n$ for $x \in [0,1]$, $n \geq 1$. Show that $\{f_n\}$ converges pointwise on $[0,1]$ for any $\alpha \in \mathbb{R}$. Find all α such that the convergence is uniform on $[0,1]$ and all α such that

$$\lim_{n\to\infty}\left(\int_0^1 f_n(x)dx\right) = \int_0^1 \left(\lim_{n\to\infty} f_n(x)\right) dx.$$

[15]

7. Find all real x for which the power series $\sum_{n=1}^{\infty}\left[n + \frac{1-(-1)^n}{2}\right] x^n$ converges and show that for all such x, the series represents the function $\frac{2x}{(1-x)(1-x^2)}$ [4 + 6 = 10]

8. (a) Let $f(x) = (\pi - |x|)^2$, $x \in [-\pi,\pi]$. Compute the Fourier coefficients of f and show that

$$f(x) = \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4}{n^2}\cos nx$$

(b) Use (a) to prove that

(i) $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ (ii) $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$

[10 + (4 + 6) = 20]

Indian Statistical Institute

B.Stat Mid-Semester Examination 2010

Computational Techniques & Programming I

Date: 6/9/10

Full Marks: 30 | **Time: Two hours**

Answer all the questions

1. (a) Design the flowchart for **efficiently** finding all the prime numbers between 1 to200.

   (b) Following the above flowchart design a 'C' program and give the dry run for finding all prime numbers from 1 to 11.

[5+5]

2. (a) What is the value of sum after the following program is executed? Give the dry run.

```
main()
{
        int sum , index ;
        sum = 1 ;
        index = 9 ;
        do{
                index = index – 1 ;
                sum = 2 * sum ;
            } while(index > 9)
}
```

(b) What is the value of average? Give the dry run.

```
main()
{
        int average , sum , index ;
        index = 0 ;
        sum = 0 ;
        for( ; ; ){
                sum = sum + index ;
                ++index ;
                if(sum >= 100)
                        break ;
            }
        average = sum / index ;
}
```

[2+2]

3. Input one 5 X 5 square entity (with all rows and columns filled up) and test it in all possible ways whether it is a magic square or not.

[6]

4. (a) Design the steps for getting a compiler for a high level language (L) for machine A when the same for machine B is available.

   (b) When nothing is available, state the steps for getting a compiler for a high level language L for machine B.

[2+2]

5. What is the function of an Operating System?
What are the different varieties of Operating System developed over the years since the starting generation of Computers? Explain briefly.

[6]

INDIAN STATISTICAL INSTITUTE

Midsemester Examination, 2nd Semester, 2010-11

Statistical Methods II, B.Stat I

Total Points 100

Date: Time: 2 hours

1. Suppose that X follows p dimensional multivariate normal with mean vector μ and variance-covariance matrix Σ. Suppose that Σ is nonsingular.

   (a) Show that $(X-\mu)^T\Sigma^{-1}(X-\mu)$ has a chi-square distribution with degrees of freedom p. (Note that a nonsingular matrix A has a nonsingular symmetric square root $A^{1/2}$ so that $A^{1/2}A^{1/2} = A$.)

   (b) Consider the collection of all points where the probability density function of the above random variable X is equal to a fixed constant a. Show that the collection represents an ellipsoid in p dimensions.

   (c) Let $p = 3$, and the distribution of $X = (X_1, X_2, X_3)^T$ be multivariate normal with

   $$\mu = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, \text{ and } \Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

   Find the distribution of $3X_1 + 2X_2 + X_3$. [10+10+10=25]

2. Given five pairs of observations (X, Y) as (10.5, 23.1), (16.7, 32.8), (18.2, 31.8), (17.0, 32.0) and (16.3, 30.4), find the least squares regression line, and, under the assumption of normality of errors, test at level $\alpha = 0.05$ whether the slope of the regression line is different from zero (two sided critical value fot the t distribution with 3 degrees of freedom is 3.18) [15]

3. Consider the multiple linear regression model

   $$Y_{n\times 1} = X_{n\times p+1}\beta_{(p+1)} + \epsilon_{p+1}.$$

   Show that under the least squares estimate of β, the error sum of squares can be written as Y^TAY, where A is a suitable idempotent matrix. [15]

4. Consider testing goodness of fit under multinomial models. Let the number of cells be k, and let the proportion of observations in the i-th cell under a random sample of size n be d_i. Consider any goodness-of-fit test statistic of the form $\rho_C(d, p_0) = \sum_{i=1}^k C(\delta_i)p_{0i}$, where the null hypothesis of interest is $H_0 : p_i = p_{0i}, i = 1, \ldots, k$, p_i is the theoretical proportion of the observations in the i-th cell, and $\delta_i = p_i/p_{i0} - 1$.

Here C is a convex function on $[-1, \infty)$ with $C(0) = 0$, $C'(0) = 0$ and $C''(0) = 1$. Also the third derivative of $C(\cdot)$ is finite and continuous at zero. Show that under the null hypothesis the asymptotic distribution of the test statistic $2n\rho_C(d, p_0)$ is the same as that of the corresponding Pearson's chi-square statistic (You need only show that the general statistic and the Pearson's chi-square statistic are separated only by an $o_p(1)$ term). [15]

5. (a) Explain with suitable justification how you would generate an observation from the density

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

where using a random observation from $U(0,1)$, where λ and θ are known constants.

(b) Suppose that you are given a random observation X from the exponential distribution with density

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

where θ is a positive constant. Show that the transformation $Y = [X]$ would generate an observation from a geometric distribution with probability mass function

$$f(x) = (1-p)^x p, \quad x = 1, 2, \ldots.$$

Find the expression for p in terms of θ.

[12+13=25]

INDIAN STATISTICAL INSTITUTE

Vectors and Matrices I : B. Stat 1st year: Mid Semester Examination: 2010-11
September 10, 2010.

Maximum Marks 40 Maximum Time 2:30 hrs.

Answer all questions. Each question has 6 marks.

(1) Find the dimension of all 2×2 trace zero matrices with real entries as a vector space over $\mathbb{R}$.

(2) Let V be a finite dimensional vector space and $A = \{x_1, x_2, \cdots, x_n\} \subset V$. Suppose that A generates V (i.e. $\operatorname{Span} A = V$). Show that if there is a vector $v \in V$, such that v can be expressed as a unique linear combination of elements of A, then A is a basis.

(3) Let V and W be two finite dimensional vector spaces over $\mathbb{R}$ and $T : V \to W$ be a linear map. Let S_2 be a convex subset of W and $S_1 = \{v \in V \mid Tv \in S_2\}$. Show that S_1 is a convex subset of V. (A set S is called convex if for any two points $x, y \in S$ and for any $0 < t < 1$, $tx + (1-t)y \in S$.)

(4) Consider $\mathbb{R}^2$ as a vector space over $\mathbb{R}$. Let $\{(x_1, x_2), (y_1, y_2)\}$ be a set of linearly independent vectors in $\mathbb{R}^2$. Show that the set $\{(x_1, y_1), (x_2, y_2)\}$ is also a linearly independent set in $\mathbb{R}^2$.

(5) Let $W = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid x_1 = x_2, x_3 = x_4 = x_5\}$ be the subspace of $\mathbb{R}^5$. Find the dimension of W.

Prove that there does not exist a linear map T from $\mathbb{R}^5$ to $\mathbb{R}^2$, such that its null space $N(T) = W$.

(6) Let V be a finite dimensional vector space over $\mathbb{C}$ and $T : V \to V$ be a liner transformation. Suppose that there are two nonzero vectors v_1 and v_2 such that $Tv_1 = \lambda_1 v_1$ and $Tv_2 = \lambda_2 v_2$ where λ_1, λ_2 are two complex numbers and $\lambda_1 \neq \lambda_2$. Show that v_1, v_2 are linearly independent.

(7) Let V be a finite dimensional vector space and W be a subspace of V. Let $\dim V = n$ and $\dim W = m$ with $n > m$. Find the dimension of the quotient space V/W.

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2010-11

Course Name: B. Stat. I Yr.

Subject Name: Analysis I

Date: 22/11/2010 Maximum Marks: 60 Duration: 3 Hrs

---

(1) Which of the following infinite series converge? Give reasons for your answers.

(a) $\sum_{n=2}^{\infty} \frac{n}{(\log n)^{n/2}}$,

(b) $\sum_{n=1}^{\infty} a_n$, where $a_1 = 3$, $a_{n+1} = \frac{n}{n+1} a_n$. [10]

(2) (a) Show that $\lim_{x \to 0} \sin \frac{1}{x}$ does not exist.

(b) Let $f(x) = \frac{1}{e^{1/x}+1}$, $x \neq 0$. Find $\lim_{x \to 0+} f(x)$ and $\lim_{x \to 0-} f(x)$.

(c) Let $I = [a, b]$ and $f : I \to \mathbb{R}$ be a continuous function on I. Suppose that for each $x \in I$, there exists $y \in I$ such that $|f(y)| \leq \frac{1}{2}|f(x)|$. Prove that there exists a point $c \in I$ such that $f(c) = 0$.

(d) Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous at 0, $f(0) = 0$, and satisfies

$$f(x + y) \leq f(x) + f(y) \quad \text{for all } x, y \in \mathbb{R}.$$

Show that f is uniformly continuous on $\mathbb{R}$. [3+3+4+4]

(3) (a) Let f be continuous on $[0, 1]$ and differentiable on $(0, 1)$. Suppose that $f(0) = f(1) = 0$ and that there exists a point $x_0 \in (0, 1), x_0 \neq \frac{1}{2}$, such that $f(x_0) = 1$. Prove that there exists $c \in (0, 1)$ such that $|f'(c)| > 2$.

(b) Let g be continuous on $[a, b], a > 0$, and differentiable on (a, b). Show that there exists $x_0 \in (a, b)$ such that

$$\frac{bg(a) - ag(b)}{b - a} = g(x_0) - x_0 g'(x_0).$$

[4+4]

(4) Consider the function

$$f(x) = \begin{cases} \frac{\sin x}{x}, & x \in (0, \pi/2], \\ 1, & x = 0. \end{cases}$$

(a) Show that f is strictly decreasing on $[0, \pi/2]$.

(b) Using (a) prove that $\sin x \geq \frac{2}{\pi} x$ for $x \in [0, \pi/2]$. [4]

(5) (a) Find $\lim_{x \to 0} \frac{\sin^2 x}{\tan x^2}$.

(b) Show that the Taylor series of $f(x) = \log(1 + x)$ about $x = 0$ converges to $f(x)$ for all $x \in (-1, 1]$. [4+8]

(6) (a) The sum of two nonnegative numbers is 36. Find the numbers if the sum of their square roots is to be as large as possible.

(b) Apply Newton's method to the function $f(x) = x^{1/3}$ with the initial approximation $x_0 = 1$ and calculate x_1, x_2, x_3, and x_4. What happens to $|x_n|$ as $n \to \infty$? [4+3]

(7) Consider the function $f(x) = x\sqrt{8 - x^2}$.

(a) Find the local and absolute extrema of f.

(b) Find the intervals on which f is increasing and the intervals on which f is decreasing.

(c) Find where the graph of f is convex and where it is concave.

(d) Find the points of inflection of f.

(e) Sketch the general shape of the graph for f. [8]

INDIAN STATISTICAL INSTITUTE

Semestral Examination, 1^{st} Semester, 2010-11

Statistical Methods I, B.Stat I

Total Points 100

Date: 24.11.2010 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 numbers, 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. Suppose that X and Y have the joint probability density function given by

$$f(x,y) = 2, \quad 0 \leq x \leq y \leq 1.$$

   (a) Find $E(X)$, $E(Y)$, $Var(X)$, $Var(Y)$, $Cov(X,Y)$ and $Corr(X,Y)$.

   (b) Find $P(1/2 \leq X + Y \leq 1)$.

   (c) Find the conditional probability density function of X given $Y = y$.

   [12+5+5=22 points]

3. Suppose that Y, U_1 and U_2 are nonnegative random variables such that $Y = U_1 + U_2$, and U_1 and U_2 are independent.

P. T. O.

(a) Suppose $Y \sim \chi^2(p)$ and $U_1 \sim \chi^2(q)$, where $p > q$. Show that $U_2 \sim \chi^2(p-q)$.

(b) Let $X_1, X_2, \ldots, X_n$ represent an independently and identically distributed sample from a $N(\mu, \sigma^2)$ distribution. Let

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

be the sample variance based on the above sample, where $\bar{X}$ is the usual sample mean. It is known that the sample variance s^2 and sample mean $\bar{X}$ are independent. Show, using part (3a) or otherwise, that

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1),$$

[10+10=20 points]

4. 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]

5. Let $X_1, X_2, \ldots, X_n$ represent an independently and identically distributed sample from the $U(0,1)$ distribution. Let $X_{(r)}$ be the r-th order statistic of the sample.

   (a) Find the probability density function of $X_{(r)}$.

   (b) Using part (5a) or otherwise, find the expectation $X_{(r)}$.

[10+10=20 points]

INDIAN STATISTICAL INSTITUTE

Vectors and Matrices I : B. Stat 1st year: Final Examination: 2010-11
December 2, 2010.

Maximum Marks 60 | Maximum Time 2:30 hrs.

Answer all questions.

(1) Let V be a real vector space of dimension n and let $T : V \longrightarrow V$ be a linear transformation.

(a) Show that if for some natural number m, Image T^m = Image T^{m+1} then Image T^{m+k} = Image T^{m+k+1} for any positive integer k.

(b) Show that Null T^n = Null T^{n+k} for k as above. 5+5

(2) Let A be a $n \times n$ real matrix which is not invertible. Suppose that $E_1, E_2, \ldots, E_k$ be a finite sequence of elementary matrices such that $E_k E_{k-1} \ldots E_2 E_1 A = H$ and H is in Hermite Canonical form (HCF).

(a) Show that $AE_k E_{k-1} \ldots E_2 E_1 A = A$.

(b) General solution of the consistent system $Ax = b$ can be written as

$$E_k E_{k-1} \ldots E_2 E_1 b + (I - H)z, \ z \in \mathbb{R}^n.$$

(Write all steps. Do not assume any formula for the general solution.) 5+10

(3) Let A be a $n \times n$ real nonzero matrix which is not invertible. Show that by changing at most $n - 1$ elements of A, it can be made invertible. 10

(4) Let A be a $m \times n$ real matrix with Rank $A = r > 0$. Show that there exists invertible matrices P, Q such that

$$A = P \begin{bmatrix} I_r & O_{r \times (n-r)} \\ O_{(m-r) \times r} & O_{(m-r) \times (n-r)} \end{bmatrix} Q.$$

Can you define a generalized inverse G of A of rank $s > r$ using the matrices P, Q and which is of the form above? 8 +5

(5) Let $Ax = b$ be a system of linear equations where A is an $m \times n$ real matrix.

(a) Prove that Rank A = Rank $[A : b] < n$ if and only if $Ax = b$ is consistent and solution of $Ax = b$ is not unique.

(b) If the system is genuinely non-homogenous (i.e. if $b \neq 0$) show that the set of all solutions is given by $\{Gb \mid G \text{ is a generalized inverse of } A\}$. 10+5

INDIAN STATISTICAL INSTITUTE
First-Semester Examination: 2010-11
Course Name: B. Stat. I Yr.
Subject Name: Analysis I (Back paper)

Date: 12.1.11 Maximum Marks: 100 Duration: 3 Hrs

---

(1) Give brief answers to the following questions.

(a) Find the supremum and infimum of the set $\left\{\frac{m}{m+n} : m, n \in \mathbb{N}\right\}$.

(b) Let (a_n) be a bounded sequence which satisfies the condition

$$a_{n+1} \geq a_n - \frac{1}{2^n}, \quad n \in \mathbb{N}.$$

Show that the sequence (a_n) is convergent.

(c) Let $a_n = \frac{1}{\sqrt{n}}\left(\frac{1}{\sqrt{1}+\sqrt{3}} + \frac{1}{\sqrt{3}+\sqrt{5}} + \cdots + \frac{1}{\sqrt{2n-1}+\sqrt{2n+1}}\right), n \in \mathbb{N}$. Find $\lim_{n\to\infty} a_n$.

(d) Give an example of a bounded function on $[0,1]$ which achieves neither an infimum nor a supremum.

[3×4]

(2) (a) If $\lim_{n\to\infty} a_n = +\infty$ or $\lim_{n\to\infty} a_n = -\infty$, then show that

$$\lim_{n\to\infty}\left(1+\frac{1}{a_n}\right)^{a_n} = e.$$

(b) Let $a \in \mathbb{R}$ and let (a_n) be defined as follows:

$$a_1 \in \mathbb{R} \text{ and } a_{n+1} = a_n^2 + (1-2a)a_n + a^2 \text{ for } n \in \mathbb{N}.$$

Determine all a_1 such that the sequence (a_n) converges and in such a case find its limit.

[6+6]

(3) Determine whether the following series are absolutely convergent, conditionally convergent, or divergent.

(a) $\sum_{n=2}^{\infty}(-1)^n \frac{1}{\ln(n^3)}$;

(b) $\sum_{n=1}^{\infty}(-1)^n \frac{\ln n}{n-\ln n}$.

[10]

(4) Let $f(x) = [x] + (x-[x])^{[x]}$ for $x \geq \frac{1}{2}$. Show that f is continuous at every $x \geq \frac{1}{2}$ and that it is strictly increasing on $[1,\infty)$.

($[x]$ denotes the greatest integer less than or equal to x.)

[6]

(5) Find the derivatives (if they exist) of the following functions:

(a) $f(x) = [x]\sin^2(\pi x)$, $x \in \mathbb{R}$,

(b) $g(x) = \begin{cases} x^2 e^{-x^2}, & |x| \leq 1, \\ 1/e, & |x| > 1. \end{cases}$

[4+4]

P.T.O.

(6) (a) Show that $f(x) = \ln(1 + x^2)$ is uniformly continuous on $[0, \infty)$.

(b) Let $a_1, a_2, \ldots, a_n$ be real numbers such that

$$|a_1 \sin x + a_2 \sin 2x + \cdots + a_n \sin nx| \leq |\sin x| \text{ for all } x \in \mathbb{R}.$$

Prove that $|a_1 + 2a_2 + \cdots + na_n| \leq 1$. [5+5]

(7) (a) Let $f : [0, 2] \to \mathbb{R}$ be continuous and $f(0) = f(2)$. Prove that there exist x_1 and x_2 in $[0, 2]$ such that

$$x_2 - x_1 = 1 \text{ and } f(x_1) = f(x_2).$$

(b) Assume that f is twice differentiable on (a, b), and that there exists $M \geq 0$ such that $|f''(x)| \leq M$ for all $x \in (a, b)$. Prove that f is uniformly continuous on (a, b). [4+6]

(8) (a) For any $k \in \mathbb{N}$, and for all $x > 0$, prove that

$$x - \tfrac{1}{2}x^2 + \cdots - \tfrac{1}{2k}x^{2k} < \ln(1 + x) < x - \tfrac{1}{2}x^2 + \cdots + \tfrac{1}{2k+1}x^{2k+1}.$$

(b) Let f be twice continuously differentiable on $\mathbb{R}$ such that $f(0) = 1, f'(0) = 0$ and $f''(0) = -1$. Let $a \in \mathbb{R}$. Show that

$$\lim_{x \to \infty} \left(f\left(\tfrac{a}{\sqrt{x}}\right)\right)^x = e^{-a^2/2}.$$

(c) Let P be a polynomial of degree $n \geq 2$. If all roots of P are real, then prove that all roots of P' are also real. [5+6+4]

(9) (a) Show that the equation $x^3 + 3x + 1 = 0$ has exactly one real solution. Use Newton's method to estimate the solution. Start with $x_0 = 0$ and find x_1 and x_2.

(b) An isosceles triangle has its vertex at the origin and its base parallel to the x-axis with the vertices above the x-axis on the curve $y = 27 - x^2$. Find the largest area the triangle can have. [4+5]

(10) Consider the function $f(x) = x^5 - 5x^4$.

(a) Find the local and absolute extrema of f.

(b) Find the intervals on which f is increasing and the intervals on which f is decreasing.

(c) Find where the graph of f is convex and where it is concave.

(d) Find the points of inflection of f.

(e) Sketch the general shape of the graph for f. [8]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2010-11 (Second semester)

B. Stat. I Yr. Analysis II

Date: 21 /02/2011 Maximum Marks: 40 Duration: 3 Hrs

---

(1) (a) A function $\varphi : [a,b] \to \mathbb{R}$ is called a *step function* on $[a,b]$ if there exists a partition $\{a = x_0, x_1, x_2, \ldots, x_n = b\}$ of $[a,b]$ and real numbers $c_1, c_2, \ldots, c_n$ such that $\varphi(x) = c_k$ for all $x \in (x_{k-1}, x_k), k = 1, 2, \ldots, n$.

Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Show that for any $\epsilon > 0$, there exists a step function φ on $[a,b]$ such that

$$\int_a^b |f(x) - \varphi(x)|\, dx < \epsilon.$$

(b) Let $f : [a,b] \to \mathbb{R}$ be continuous. Show that for every $x \in [a,b]$

$$\int_a^x \left(\int_a^u f(t)\, dt\right) du = \int_a^x (x-u) f(u)\, du.$$

(c) Let f and g be continuous on $[a,b]$ and $g(x) \geq 0$ for all $x \in [a,b]$. Prove that there exists $c \in [a,b]$ such that

$$\int_a^b f(x) g(x)\, dx = f(c) \int_a^b g(x)\, dx.$$

(d) Let f be continuous on $[0,1]$. For positive a and b, find the limit

$$\lim_{\epsilon \to 0+} \int_{a\epsilon}^{b\epsilon} \frac{f(x)}{x}\, dx.$$

[3+3+3+3]

(2) Let f be a function which is continuous on $[0,1]$, differentiable on $(0,1)$, $f(0) = 0$, and $0 < f'(x) \leq 1$ on $(0,1)$.

(a) Define $\Phi(t) = \left(\int_0^t f(x)\, dx\right)^2 - \int_0^t (f(x))^3\, dx$, $t \in [0,1]$. Show that Φ is a monotone function. Also, prove that $\left(\int_0^1 f(x)\, dx\right)^2 \geq \int_0^1 (f(x))^3\, dx$.

(b) Find all such functions f so that equality holds in (a). [4+4]

(3) Study the convergence of the following improper integrals.

(a) $\displaystyle\int_1^2 \frac{x^3 + 27}{(1-x^2)\sqrt{2 - x - x^2}}\, dx$, (b) $\displaystyle\int_1^\infty \frac{1}{x\left(\sqrt{\ln x} + (\ln x)^2\right)}\, dx$. [3+7]

(4) (a) Let $f_n : [0,1] \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{nx^2}{1 + nx}, \quad n \in \mathbb{N}.$$

Does the sequence (f_n) converge uniformly on $[0,1]$?

(b) Let $f_n : \mathbb{R} \to \mathbb{R}$ be uniformly continuous for each $n \in \mathbb{N}$. Suppose that $f_n \to f$ uniformly on $\mathbb{R}$. Show that f is uniformly continuous.

(c) Study the uniform convergence of the following series of functions on the given set E:

$$\sum_{n=1}^\infty \frac{n^2}{\sqrt{n!}} (x^n + x^{-n}), \quad E = \{x \in \mathbb{R} : \tfrac{1}{2} \leq |x| \leq 2\}.$$

[4+4+4]

INDIAN STATISTICAL INSTITUTE

Mid-semester Examination: 2010-2011

B.Stat. (Hons.) 1st Year. 2nd Semester

Probability Theory II

Date: February 22, 2011 **Maximum Marks: 50** **Duration: 2 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. Consider a symmetric random walk $\{S_n : n \geq 0\}$ as we have introduced in class. Fix an integer $n \geq 1$. Show that $P(S_{2n} = 0, \max(S_1, \ldots, S_{2n-1}) = k) = P(S_{2n} = 2k) - P(S_{2n} = 2k + 2)$. [10]

2. Let $X \sim \mathrm{N}(0,1)$. Show that for every $a > 0$, $\lim_{x \to +\infty} P(X > x + a/x | X > x) = \exp(-a)$. [8]

3. Suppose $X \sim \mathrm{U}(0,1)$. Find the distribution of $Y := -2 \log X$, in terms of a standard distribution, to be identified by you. [8]

4. For $\theta > 0$, define

$$f(x,\theta) = \begin{cases} \left(1 + \frac{\theta}{1-x}\right) \exp\left(-\frac{\theta x}{1-x}\right), & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Show the following.

(a) For every $\theta > 0$, $f(x, \theta)$ is a probability density function.

(b) The function

$$g(\theta) := \int_0^1 x^4 f(x,\theta)\, dx, \ \ \theta > 0,$$

is strictly decreasing on $(0, \infty)$. [8+8=16]

5. Suppose X is a random variable such that $\lim_{x \to +\infty} x^p P(| X | \geq x) = 0$ for some $p > 0$. Show that $\mathrm{E}(| X |^q) < \infty$ for every $q \in (0, p)$. [8]

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

Computational Techniques and Programming II

Indian Statistical Institute, Kolkata
B.Stat (hons.) I
(2010-11)
Semester 2
Midsemstral examination

Date: Feb 25, 2011 Duration: 2hrs.

Attempt all problems. Each problem carries 7 marks. The maximum you can score is 30. This is a closed note, closed book examination. You may use your own calculator. Laptops are not allowed. If you think that there is a mistake in some problem you must justify your point to get credit for that problem.

1. Recall the algorithm for in-place inversion using Gauss-Jordan method. At each step we choose the pivot `a[p][p]`, and then we modify the matrix. Discuss if the following C code achieves this modification.

```
for(i=0;i<n;i++) {
  for(j=0;j<n;j++) {
    if(i==p) {
      if(j==p) {
        a[i][j] = 1/a[i][j]; /*Invert pivot*/
      }
      else {
        a[i][j] /= a[p][p]; /*Divide pivotal row by pivot*/
      }
    }
    else {
      if(j==p) {
        a[i][j] /= -a[p][p]; /*Divide pivotal col by -pivot*/
      }
      else {
        a[i][j] -= a[i][p]*a[p][j]/a[p][p];
      }
    }
  }
}
```

2. Write a C program that will take two permutations π, ξ of $\{0, ..., n-1\}$ as input (in the form of two integer arrays of length n) and output the permutation $\mu = \pi \circ \xi \circ \pi^{-1}$. Your program must not compute π^{-1} explicitly.

3. Write down a flowchart for computing the Cholesky decomposition of a given positive definite matrix, A. By Cholesky decomposition we mean factorising the matrix as

$$A = LL'$$

where L is a lower triangular matrix. You should not use separate memory for storing L.

4. Discuss how memory leakage occurs in the following program. Also suggest how you should remedy the problem.

```
float **x;
int *y;

x = (float **) calloc(10,sizeof(float *));
for(i=0;i<10;i++)
  x[i] = (float *) calloc(10,sizeof(float));

/*Work with x*/

free(x);

y = (int *) calloc(20,sizeof(int));

/*Work with y*/
```

5. Let $L_i(x)$ be the i-th Lagrange polynomial interpolating the $n+1$ points $(x_0, y_0), ..., (x_n, y_n)$, where $n \geq 2$, and the x_i's are all distinct. Find the value of

$$\sum_{i=0}^{n}(3x_i^2 - 2x_i + 1)L_i(3).$$

Justify your answer.

Good Luck!!

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: Second Semester(2010-2011)

B. STAT. I year

Vectors and Matrices II

Date: 28 Feb. 2011. Maximum Marks: 30 Duration: 3 Hrs.

Notes: *(i) All the matrices and vectors considered are over real field unless otherwise stated. (ii) Class room notation is used. (iii) State clearly the results used.*

1 *Prove or disprove the following:*

(a) $r(ABC) = r(AC)$ *if* B *is nonsingular.*

(b) A *is n.n.d. matrix implies* $A = B^2$ *for some symmetric matrix* B.

(c) A *is p.d. and* B *is n.n.d. implies* $A + B$ *is p.d.*

(d) *Let* A *and* B *be real symmetric matrices and* A *is p.d. Then each eigenvalue of* $A^{-1}B$ *is* 1 *implies* $A = B$.

(e) $x'Ax = 0$ *for all real vectors* x *implies* $Ay = \phi$ *for some nonnull vector* y.

(f) A *is a generalized inverse of itself implies* A *is idempotent.*

(g) A *and* B *are nonnegative definite matrices implies all the eigenvalues of* AB *are nonnegative.*

(h) λ *is a nonzero eigenvalue of* A *implies* $1/\lambda$ *is an eigenvalue of every reflexive g-inverse of* A.

(i) G *is a reflexive g-inverse of* A *if and only if* $G = C^{-R}B^{-L}$ *for some left inverse* B^{-L} *of* B *and some right inverse* C^{-R} *of* C *where* $A = BC$ *is any rank factorization of* A.

(j) $x'Jx \leq n\ x'x$, *where* J *is a square matrix of order* n *with each element equal to* 1.

[$10 \times 2 = 20$]

2 *Let*

$$A = \begin{pmatrix} B & C \\ C' & E \end{pmatrix},$$

be a p.d. matrix where B *is nonsingular. Then show that the matrix* $(E - C'B^{-1}C)$ *is p.d.*

[5]

3 *Consider the system of linear equations* $Ax = b$ *where* A *is a nonsingular matrix of order* n. *Then show that* x_i, *the* i^{th} *element of the solution vector* x, *is given by* $|A_i|/|A|$, *where* A_i *is the matrix obtained from* A *by replacing the* i^{th} *column of* A *by the vector* b.

[5]

4 *State and prove Cayley-Hamilton theorem.*

[5]

$\cdots xXx \cdots$

INDIAN STATISTICAL INSTITUTE

Semester Examination: Second Semester(2010-2011)

B. STAT. I year

Vectors and Matrices II

Date: 2 May 2011. Maximum Marks: 70 Duration: 3 Hrs.

Notes: *(i) All the matrices and vectors considered are over real field unless otherwise stated. (ii) Class room notation is used. (iii) State clearly the results used.(iv) 5 marks are alloted for neat presentation of the answers.*

1 Prove or disprove the following:

(a) Let A be a matrix of order $m \times n$ such that $C(A) \subseteq C(B)$ and $R(A) \subseteq R(C)$ then $A = BXC$ for some matrix X.

(b) Let A and B be idempotent matrices satisfying $C(A) \subseteq C(B)$ and $R(B) \subseteq R(A)$ then $A = B$.

(c) $C(A) = C(A^*) \Leftrightarrow r(A) = r(A^2)$.

(d) λ is an eigenvalue of A implies $|\lambda| \leq \delta$, where δ is the maximum singular value of A.

(e) For a block triangular matrix $A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$, $r(A) = r(B) + r(D)$.

(f) Number of nonzero eigenvalues of A is same as the number of its singular values implies $r(A) = r(A^2)$.

(g) $ABA = 0$ implies B can be decomposed as $B = C + D$ where $AC = 0$ and $DA = 0$.

(h) Every square matrix can be decomposed as $A = B + C$ where $r(B) = r(B^2)$, C is nilpotent, satisfying $BC = 0 = CB$.

(i) $|A| \neq 0$ implies A can be expressed as $A = BC$, where B is positive definite and C is orthogonal.

(j) A and B are matrices of same order and $C(A) = C(B) \Rightarrow B = AZ$ for some nonsingular matrix Z.

[$10 \times 2 = 20$]

2 Let N be a positive definite matrix and $(x, y) = y'Nx$ be the inner product defined on R^n. Let $\|x\|$ be the norm induced by this inner product.

(a) Let $A_{n \times m}$ and $B_{n \times k}$ be two matrices. Then show that $\|Ax\| \leq \|Ax + By\|$ for all x and y implies that the columns of A are orthogonal (w.r.t. above inner product) to the columns of B.

(b) Show that a set of necessary and sufficient conditions for the matrix $G_{n \times m}$ to be a minimum norm generalized inverse (w.r.t. above norm) of $A_{m \times n}$ is $AGA = A$ and NGA is symmetric.

[$2 \times 5 = 10$]

[P.T.O.]

3 Let A be a positive definite matrix given by

$$A = \begin{pmatrix} B & C \\ C' & E \end{pmatrix} \text{ and } A^{-1} = \begin{pmatrix} P & Q \\ Q' & R \end{pmatrix},$$

where B is nonsingular.

(a) Show that the matrix $(E - C'B^{-1}C)$ is positive definite.
(b) Show that the matrix $(P - B^{-1})$ is nonnegative definite.

[$2 \times 5 = 10$]

4 Prove the following:

(a) Let $A = B + C$ and $r(A) = r(B) + r(C)$. Then A is nonnegative definite and B is symmetric imply B and C are nonnegative definite.
(b) H is idempotent matrix implies that there exists a positive definite matrix M such that HM is hermitian.

[$2 \times 5 = 10$]

5 (a) Derive a set of necessary and sufficient conditions for the matrix P_A to be the orthogonal projector of R^n onto $C(A_{n \times n})$.
(b) Give an expression for P_A in terms of A and justify your answer.
(c) Obtain the matrix P_A for the given matrix A and justify your answer.

$$A = \begin{pmatrix} 1 & 0 & -2 & 1 & 2 \\ 0 & 2 & 1 & 3 & -1 \\ 0 & 0 & 6 & 7 & 5 \\ 0 & 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix},$$

[$6 + 4 + 2 = 12$]

6 (a) For the given nonnegative definite matrix A obtain the lower triangular matrix L satisfying $A = LL'$.

$$A = \begin{pmatrix} 4 & 2 & -2 & 0 \\ 2 & 2 & 1 & 3 \\ -2 & 1 & 6 & 7 \\ 0 & 3 & 7 & 14 \end{pmatrix},$$

(b) For the above given matrix A, derive a nonsingular matrix of transformation (of the variables) which reduces the quadratic form $x'Ax$ to diagonal form.

[$6 + 2 = 8$]

$\cdots xXx \cdots$

INDIAN STATISTICAL INSTITUTE

Semestral Examination, 2^{nd} Semester, 2010-11

Statistical Methods II, B.Stat I

Total Points 100

Date: Time: 3 hours

1. Define intra-class correlation and distinguish it from inter-class correlation. Derive the formula for the intra-class correlation when the variable x is measured over p classes, with the i-th class having k_i members.

[15 points]

2. (a) Define the multiple correlation coefficient $r_{1.23...p}$ between x_1 and the set of variables $x_2, x_3, \ldots, x_p$. Show that

$$r^2_{1.23...p} = \frac{\text{var}(X_{1.23...p})}{\text{var}(x_1)}$$

where $X_{1.23...p}$ is the value of x_1 predicted by the multiple regression equation of x_1 on $x_2, \ldots, x_p$.

(b) For a set of variables $x_1, x_2, \ldots, x_p$, define the partial correlation coefficient $r_{12.34...p}$. For three variables x_1, x_2, x_3, show that

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}},$$

where r_{ij} is the simple correlation coefficient between x_i and x_j.

[10+10=20 points]

3. (a) Suppose that you have a random observation U from the Uniform (0, 1) distribution at your disposal. Explain how you can draw a random observation from the Pareto distribution given by the density

$$f(x) = \frac{ca^c}{x^{c+1}}, \quad a, c \geq 0, \quad x \geq a.$$

What is the median of this distribution?

(b) Suppose you have a random number generator such that you can generate random numbers from any binomial distribution. Explain how you can drawn random numbers X and Y such that X is distributed as binomial (8, 2/3) and Y is distributed as binomial (18, 2/3), and the correlation between X and Y is 0.5.

[15+10=25 points]

4. (a) Write down the simple linear regression model of y on x, and state the standard assumptions. Derive the least squares estimates of the regression parameters.

(b) For 20 army personnel, the regression of weight of kidneys (y) on weight of heart (x), both measured in oz., is

$$y = 0.399x + 6.934,$$

and the regression of weight of heart on weight of kidneys is

$$x = 1.212y - 2.461.$$

Find the correlation between the two variables, and their means.

(c) A simple linear regression performed with six paired observations gives the following fitted regression line for Y on X:

$$y = 189.4389 - 0.1901x.$$

The values of the dependent variable Y, and the fitted values $\hat{Y}$ for the six cases are as given in the table below.

Observations Number	Y	$\hat{Y}$
1	156	158.6516
2	157	159.0317
3	159	158.6516
4	160	159.9819
5	161	159.0317
6	161	158.6516

Using the above numbers, perform a one sided test for the null hypothesis H_0 : $\beta_1 = 0$ versus $H_1 : \beta_1 < 0$. (t tables attached).

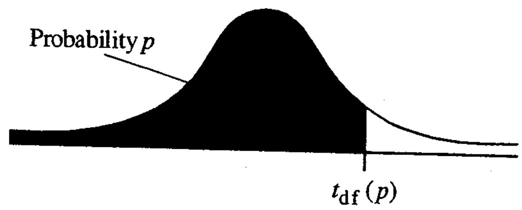
[7+6+7=20 points]

5. The following table gives the frequencies of the number of weed seeds in 196 half kg. packets of a variety of pulses. Fit a Poisson model to this data, calculate the expected frequencies, and perform a goodness of fit test to determine whether the Poisson model is appropriate for this data.

Number of weed seeds	Observed Frequency
0	7
1	33
2	54
3	37
4	34
5	16
6	8
7	5
8	1
9	1
10 or more	0

[20 points]

Table A.2 Selected Percentiles of t-Distributions



Tabled values are $t_{df}(p)$

d.f.	.75	.80	.85	.90	.95	.975	.98	.99	.995	.9975	.999	.9995
	Probability p											
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.67	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.32	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	1.215	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
70	0.678	0.847	1.044	1.294	1.667	1.994	2.093	2.381	2.648	2.899	3.211	3.435
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
90	0.677	0.846	1.042	1.291	1.662	1.987	2.084	2.368	2.632	2.878	3.183	3.402
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
500	0.675	0.842	1.038	1.283	1.648	1.965	2.059	2.334	2.586	2.820	3.107	3.310
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090	3.291

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2010-2011

B.Stat. (Hons.) 1st Year. 2nd Semester

Probability Theory II

Date: May 06, 2011 **Maximum Marks: 85** **Duration: 3 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. Let $0 < a < 1$. Define $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$\begin{aligned} F(x,y) &= 1 - e^{-x} - e^{-y} + e^{-x-y-axy} \quad \text{if } x, y > 0, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

(1) Show that F is a bivariate cumulative distribution function (cdf).

(2) Let (X,Y) have cdf F. Show that $\text{Cov}(X,Y) = \int_0^\infty [e^{-x}/(1+ax)]\, dx - 1$.

[8+6 = 14]

2. It is known that the arithmetic mean of a finite set of positive numbers is greater than or equal to the harmonic mean of the same. State and prove an analogous result for a positive random variable, stating clearly the assumptions you need. [10]

3. Let $X_i \sim \text{Beta}(\gamma_i, \delta_i)$, $i = 1, \ldots, n$, be independently distributed random variables. Suppose that $\gamma_i = \gamma_{i-1} + \delta_{i-1}$, $i = 2, \ldots, n$. Show that $\prod_{i=1}^n X_i \sim \text{Beta}(\gamma_1, \delta_1 + \cdots + \delta_n)$.

[12]

4. Let D be the distance between two points picked independently at random from a uniform distribution inside a disc of radius r. Show that $\text{E}(D^2) = r^2$. [11]

5. Let $X_1, \ldots, X_n$ be independent and identically distributed (i.i.d.) random variables with $X_1 \sim \text{N}(\mu, \sigma^2)$. Let $\bar{X} := \sum_{i=1}^n X_i/n$, $R := \max_{1 \le i \le n} X_i - \min_{1 \le i \le n} X_i$. Show that $\bar{X}$ and R are independently distributed. [12]

P.T.O.

6. For $i \geq 1$, let $X_i \sim$ exponential(λ) be i.i.d. random variables. Let $S_n := X_1 + \cdots + X_n$ for $n \geq 1$, and $:= 0$ for $n = 0$. Fix $t > 0$. Define N_t by $N_t = \max\{n \geq 1 : S_n \leq t\}$ if $\{n \geq 1 : S_n \leq t\} \neq \emptyset$ and $= 0$, otherwise. Assume that $P(N_t < \infty) = 1$. Find the distribution function and mean of $t - S_{N_t}$. [10+4 = 14]

7. Suppose $X_1, \ldots, X_n$ $(n > 2)$ are i.i.d. random variables with a common distribution function F. Let F be continuous. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the corresponding order statistics. Show that $[F(X_{(n)}) - F(X_{(2)})]/[F(X_{(n)}) - F(X_{(1)})] \sim \text{Beta}(n-2, 1)$. [12]

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

INDIAN STATISTICAL INSTITUTE

Second Semester Examination: 2010-11

B. Stat. I Yr. Analysis II

Date: 09/05/2011 Maximum Marks: 60 Duration: 3 Hours

---

(1) (a) Find the value of

$$\lim_{n\to\infty}\left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots\frac{1}{3n}\right)$$

by using the Riemann integral of a suitable function.

(b) Let f be a nonnegative and continuous function on $[0,\infty)$ such that the improper integral $\int_0^\infty f(x)\,dx$ is convergent. Prove that

$$\lim_{n\to\infty}\frac{1}{n}\int_0^n xf(x)\,dx=0.$$

(c) Let $\alpha>0$. Study the convergence of the improper integral $\displaystyle\int_1^\infty \frac{e^{\sin x}\sin 2x}{x^\alpha}\,dx.$

[Observe that $e^{\sin x}\sin 2x$ is the derivative of $2e^{\sin x}(\sin x-1)$.] **[4+5+5]**

(2) (a) Let $f:\mathbb{R}\to\mathbb{R}$ be a differentiable function such that f' is uniformly continuous on $\mathbb{R}$. Define $g_n:\mathbb{R}\to\mathbb{R}$ by $g_n(x)=n[f(x+1/n)-f(x)]$. Show that (g_n) converges uniformly to f' on $\mathbb{R}$.

(b) Let $f:[0,1]\to\mathbb{R}$ be an infinitely differentiable function such that $f\not\equiv 0$, $f^{(n)}(0)=0$ for $n=0,1,2,\ldots$, and let there exist a sequence (a_n) of real numbers such that $\sum_{n=1}^{\infty}a_nf^{(n)}(x)$ converges uniformly on $[0,1]$. Prove that $\lim_{n\to\infty}n!a_n=0$. **[6+6]**

(3) (a) Let R,R_1, and R_2 be the radii of convergence of the power series $\sum_{n=0}^{\infty}a_nx^n$, $\sum_{n=0}^{\infty}b_nx^n$, and $\sum_{n=0}^{\infty}\frac{a_n}{b_n}x^n$, respectively. Assume that $R_1,R_2\in(0,\infty)$ and $b_n\neq 0$ for all $n\geq 0$. Show that $R\leq\frac{R_1}{R_2}$. Give an example to show that the inequality can be strict.

(b) Let $\alpha>0$ and $L>0$ such that $\lim_{n\to\infty}|a_n\alpha^n|=L$. Find the radius of convergence of the power series $\sum_{n=0}^{\infty}a_nx^n$.

(c) Show that the function $f(x)=\frac{1}{1-x}$ is real-analytic at $x=3$. **[6+6+3]**

(4) (a) Show that $\sum_{k=1}^{n}\sin kx=\frac{\cos\frac{x}{2}-\cos(n+1/2)x}{2\sin\frac{x}{2}}$, $x\neq 2l\pi, l\in\mathbb{Z}$.

(b) Let (c_n) be a decreasing sequence of nonnegative real numbers such that $\lim_{n\to\infty}c_n=0$. Show that the trigonometric series $\sum_{n=1}^{\infty}c_n\sin nx$ converges for all $x\in\mathbb{R}$.

(c) Does there exist a 2π-periodic Riemann integrable function f such that $\sum_{n=1}^{\infty}\frac{\sin nx}{\sqrt{n}}$ is the Fourier series of f? Justify your answer. **[3+4+3]**

(5) Consider the 2π-periodic function f whose values on $[-\pi,\pi]$ are given by $f(x)=(\pi-|x|)^2$.

(a) Find the Fourier series of f.

(b) For which values of x does the series converge to $f(x)$?

(c) Using the Fourier series of f prove that $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$ and $\sum_{n=1}^{\infty}\frac{1}{n^4}=\frac{\pi^4}{90}$. **[4+2+6]**

Computational Techniques and Programming II

Indian Statistical Institute, Kolkata
B.Stat (hons.) I
(2010-11)
Semester 2
Semstral examination

Date: May 11, 2011 Duration: 3hrs.

Attempt all problems. The maximum you can score is 50. This is a closed note, closed book examination. You may use your own calculator. Laptops are not allowed. If you think that there is a mistake in some problem you must justify your point to get credit for that problem.

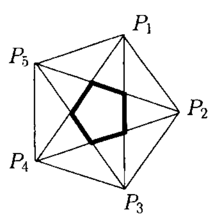
1. Consider mn cells laid out in an $m \times n$ rectangular grid. Associated with each cell (i,j) there is an unknown constant x_{ij} and a known constant c_{ij}. The values of the constants are related as follows. For each cell

$$x_{ij} = c_{ij} + \frac{\text{sum of neighbouring } x\text{-values}}{\text{number of neighbours} + 1}.$$

By neighbours we mean cells sharing a common *side*. Show that this uniquely specifies the values of x_{ij} for all i,j.

Suggest an efficient numerical method to find these unique values using given values of c_{ij}'s. Justify your suggestion in terms of efficiency for large m,n. Also prove that your method is correct. You may use theorems stated in class, provided you state them clearly. [10]

2. Given the coordinates of the vertices $P_1, ..., P_5$ of a regular pentagon, you are to produce the following diagram using R. (No need to produce the labels). Your answer should be in the form of a function `draw(P)` where `P` is a 5×2 matrix with the i-th row storing the coordinates of P_i.



[Hint: The parameter `lwd` controls the thickness of the line in any line drawing command in R. For example, `lwd=3` will produce thick lines.][10]

3. Consider the following C program to compute the sum of two fractions. Here each fraction is stored as an `int` array of length 2, the 0-th entry is the numerator, the 1-th entry is the denominator.

```
int *sum(int *a, int *b) {
  int res[2];

  res[1]  = a[1]*b[1];
  res[0]  = a[0]*b[1]  + a[1]*b[0];

  return res;
}
```

Show why use of this function gives rise to a dangling pointer error. Also suggest how you can correct the flaw. [The denominators are given to be nonzero, and the result need not be in reduced form.] [5+5]

4. Let the simple Newton-Cotes quadrature rule of order $2n$ be

$$\int_a^b f(x)dx \approx \sum_{i=0}^{2n} a_i f(x_i),$$

where $x_i = a + i(\frac{b-a}{2n})$, and a_i's are as in the definition of the simple Newton-Cotes rule. Then prove that

$$\int_a^b x^{2n+1}dx = \sum_{i=0}^{2n} a_i x_i^{2n+1}.$$

[Hint: Consider the function $x^{2n+1} - p(x)$, where $p(x)$ is the unique polynomial of degree $\leq 2n$ interpolating x^{2n+1} at the points x_i's.] [10]

5. Derive the 3-rd order Taylor's method for solving $dy/dx = e^{-xy}$ with $y(0) = 2$. Compute $y(0.5), y(1)$ approximately with step size 0.5 using this method. [5+5]

6. Suppose that you have a computer that can only do addition, subtraction and multiplication, but not division! Use Newton-Raphson method to construct an iterative scheme to compute $1/a$ for any given $a \neq 0$ using this computer. Your iteration must not use division anywhere. [5]

Good Luck!!

End Semester Examinations (2011-12)

B Stat – 1 year

Remedial English

100 marks

1 1/2 hours

Date: 12.05.2011

1. Write an essay on any one of the following topics. Five paragraphs are expected.
a) Life at the ISI
b) Cyberspace
c) Rabindra Nath Tagore: 150 years

(60 marks)

2. Fill in the blanks with appropriate prepositions – Write the whole sentence.
a) Can we meet _______ an hour and lunch ______ the Grand?
b) Please switch _______ the radio. I want ______listen _____ music.
c) I will take the train ______Mumbai ______Delhi.
d) I want ______ vote _____ my favourite candidate .
e) Please give me a cup ______ tea. I prefer tea ______ coffee.
f) He is still capable _____ being humbled ______beauty.
g) I will write _____ him ____ this matter.
h) He spoke _____ _____ _____ him.
i) He fell in _____ _____ the car.

(20 marks)

3. Fill in the blanks with appropriate words.
Some ________ later, _______ afternoon ______ ______ visited by Bari and Natesh. It ______ ______ holiday and Swami ______ ______ home. He ______ fussy and _____ the available furniture from ______ and there, dashed ______ door ______ borrowed ______ folding ______ , and ______ seats for ______ . Veena threw a brief _____ at ______ visitors and ______ past them.

(20 marks)

INDIAN STATISTICAL INSTITUTE

Semester Examination:BACK PAPER Second Semester(2010-2011)

B. STAT. I year

Vectors and Matrices II

Date:27 June 2011. Maximum Marks: 100 Duration: 3 Hrs.

Notes: *(i) All the matrices and vectors considered are over real field and the inner products are Euclidean unless otherwise stated. (ii) Class room notation is used. (iii) State clearly the results used. (iv) 5 marks are alloted for neat presentation of the answers.*

1 Prove or disprove the following:

(a) $|A| \neq 0$ implies A can be expressed as $A = BC$, where B is positive definite and C is orthogonal.

(b) A and B are matrices of same order and $C(A) = C(B) \Rightarrow B = AZ$ for some nonsingular matrix Z.

(c) $x^3 - 3x^2$ is the minimal polynomial of the matrix A implies A is singular.

(d) $A^2 = A$ and $r(A) = r$ imply that A can be expressed as sum of r non null idempotent matrices.

(e) $r(A) = r$ and A has an r^{th} order nonzero principal minor imply $r(A) = r(A^2)$.

[$5 \times 2 = 10$]

2 Prove the following:

(a) Given a vector $u_{m\times 1}$ and a nonnull vector $v_{n\times 1}$, there exists a matrix W such that $Wv = u$.

(b) x_0 is a solution of nonhomogenius equations $Ax = b$ if and only if $x_0 = Gb$ for some g-inverse G of A.

(c) If a solution x_0 of $Ax = b$ is in the row space of A then it is a minimum norm solution of $Ax = b$.

(d) Minimum norm solution is unique.

(e) x_0 is the minimum norm solution of $Ax = b$ implies $x_0 = Gb$ for some minimum norm g-inverse G of A.

[$5 \times 2 = 10$]

3 Prove the following:

(a) A real quadratic form $x'Ax$ can be written as the product of two linearly independent linear forms in x if and only if A has rank 2 and signature 0.

(b) If $x + iy$, where x and y are real vectors, is an eigenvector corresponding to a non null eigenvalue of a skew symmetric matrix, then x and y are orthogonal and of the same norm.

(c) Every singular value of an idempotent matrix A is 1 implies A is symmetric.

(d) For any real square matrix B of order n such that $|b_{ij}| \leq 1$, show that $|B|^2 \leq n^n$ and equality occurs if and only if $b_{ij} = \pm 1$ for all i and j and the rows of B are pairwise orthogonal.

[$4 \times 5 = 20$]

[P.T.O.]

4 Derive Jordan Canonical Form of a real square matrix of A order n.

[20]

5 Let M be a positive definite matrix and $(x, y) = y'Mx$ be the inner product defined on R^m. Let $\|x\|$ be the norm induced by this inner product.

(a) Show that $\|Px\| \leq \|Px + Qy\|$ for all x and y implies that the columns of P are orthogonal (w.r.t. above inner product) to the columns of Q.

(b) Show that a set of necessary and sufficient conditions for the matrix G to be a least squares generalized inverse (w.r.t. above norm) of A is $AGA = A$ and MAG is symmetric.

[$2 \times 10 = 20$]

6 Let $S = M\{(1 \;\; 0 \;\; 0)', (1 \;\; 1 \;\; 0)'\}$. Define *orthogonal projection of* R^3 *onto* S. Obtain a matrix that represents the above orthogonal projection operator. Justify your answer.

[$2 + 8 = 10$]

7 Find all the eigenvalues of the matrix $A = BC$, where B' and C are given below

$$B' = \begin{pmatrix} 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & 1 & -4 & -3 \\ 1 & -1 & 1 & -1 & 0 \end{pmatrix} \text{ and } C = \begin{pmatrix} 4 & 2 & -2 & 0 & 1 \\ 2 & 2 & 1 & 1 & 0 \\ -4 & 2 & 0 & -5 & 6 \end{pmatrix},$$

[5]

$\cdots xXx \cdots$

INDIAN STATISTICAL INSTITUTE

Backpaper Examination: 2010-2011
B.Stat. (Hons.) 1st Year. 2nd Semester
Probability Theory II

Date: 28-06-11 **Maximum Marks: 100** **Duration: 4 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. For $i \geq 1$, let $X_i \sim$ Binomial$(1, 1/2)$ be independent and identically distributed (i.i.d.) random variables. Let $S_n := X_1 + \cdots + X_n$ for $n \geq 1$, and $:= 0$ for $n = 0$. Fix an integer $n \geq 1$. Show that $P(S_1 \neq 0, \ldots, S_{2n} \neq 0) = P(S_{2n} = 0)$. [12]

2. Let $|\alpha| \leq 1, \sigma, \theta > 0$. Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$\begin{aligned} f(x,y) &= \sigma\theta e^{-\sigma x - \theta y}[1 + \alpha\{2e^{-\sigma x} - 1\}\{2e^{-\theta y} - 1\}] && \text{if } x, y > 0, \\ &= 0 && \text{otherwise.} \end{aligned}$$

(1) Show that f is a bivariate probability density function (pdf).

(2) Let (X, Y) have pdf f. Show that $\mathrm{Cov}(X, Y) = \alpha/(4\sigma\theta)$. [8+8 = 16]

3. Suppose X and Y are independent random variables. Suppose, moreover, that the distribution function of X, denoted by F, is continuous. Show that $P(X = Y) = 0$. [11]

4. Let D be the distance between two points picked independently at random from a uniform distribution inside a square of side a. Show that $\mathrm{E}(D^2) = a^2/3$. [12]

5. Let $X_i \sim$ Gamma$(\alpha_i, 1/\beta)$, $i = 1, \ldots, n$, be independently distributed random variables. For $i = 1, \ldots, n-1$, let $Y_i := (X_1 + \cdots + X_i)/(X_1 + \cdots + X_{i+1})$. Also, let $Y_n = X_1 + \cdots + X_n$.

(1) Show that $Y_1, \ldots, Y_n$ are independent.

(2) For $i = 1, \ldots, n$, find the distributions of Y_i, in terms of standard distributions, to be identified by you. [9+4 = 13]

P.T.O.

6. Let $X_1, \ldots, X_n$ be i.i.d. random variables with $X_1 \sim \mathrm{N}(\mu, \sigma^2)$. Let $\bar{X} := \sum_{i=1}^n X_i/n$, $S := \left[\sum_{i=1}^n (X_i - \bar{X})^2/(n-1)\right]^{1/2}$. Show that $\bar{X}$ and S are independently distributed. [12]

7. For $i \geq 1$, let $X_i \sim$ exponential(λ) be i.i.d. random variables. Let $S_n := X_1 + \cdots + X_n$ for $n \geq 1$, and $:= 0$ for $n = 0$. Fix $t > 0$. Define N_t by $N_t = \max\{n \geq 1 : S_n \leq t\}$ if $\{n \geq 1 : S_n \leq t\} \neq \emptyset$ and $= 0$, otherwise. Assume that $P(N_t < \infty) = 1$. Show that $S_{N_t+1} - t \sim$ exponential(λ). [12]

8. Suppose $X_1, \ldots, X_n$ $(n > 2)$ are i.i.d. exponential$(0, 1)$ variables. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the corresponding order statistics. Show that the joint distribution of $(X_{(2)} - X_{(1)}, \ldots, X_{(n)} - X_{(1)})$ is same as that of $(Y_{(1)}, \ldots, Y_{(n-1)})$, where $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n-1)}$ denote the order statistics corresponding to $n-1$ i.i.d. exponential$(0, 1)$ variables $Y_1, \ldots, Y_{n-1}$. [12]

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

INDIAN STATISTICAL INSTITUTE
Second Semester Examination: 2010-11
B. Stat. I Yr. Analysis II (Backpaper)
Date: 30/06/2011 Maximum Marks: 100 Duration: 3 Hours

---

(1) Let $f : [0,1] \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ \frac{1}{q}, & \text{if } x \in \mathbb{Q} \cap [0,1] \text{ and } x = \frac{p}{q}, \text{ where } p, q \in \mathbb{N} \text{ with } \gcd(p,q) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Let $\epsilon > 0$ be given. Show that the set $\{x \in [0,1] : f(x) \geq \epsilon\}$ is a finite set.

(b) Show that f is Riemann integrable over $[0,1]$ and find the value of $\int_0^1 f(x)\, dx$. **[3+7]**

(2) (a) Find the value of

$$\lim_{n\to\infty} \left(\frac{1}{\sqrt{n^2+n}} + \frac{1}{\sqrt{n^2+2n}} + \cdots + \frac{1}{\sqrt{n^2+n^2}} \right)$$

by using the Riemann integral of a suitable function.

(b) Let $F(x) = x^2 \cos\frac{1}{x} - 2\int_0^x t\cos\frac{1}{t}\, dt$ for $x \neq 0$ and $F(0) = 0$. Show that F is an antiderivative of the function f defined by $f(x) = \sin\frac{1}{x}$ for $x \neq 0$ and $f(0) = 0$.

Now, let $c \in [-1,1]$ and let $g(x) = \sin\frac{1}{x}$ for $x \neq 0$ and $g(0) = c$. For which values of c does there exist an antiderivative of g? Justify your answer. **[7+8]**

(3) (a) For a function f defined on $[a,b]$, let $f_n(x) = \frac{[nf(x)]}{n}, x \in [a,b], n \in \mathbb{N}$, where $[x]$ is the greatest integer less than or equal to x. Show that $f_n \to f$ pointwise on $[a,b]$. Is this convergence uniform?

(b) Let $f_n(x) = x/n$, if n is even, and $f_n(x) = 1/n$ if n is odd. Show that the sequence of functions (f_n) is pointwise convergent but not uniformly convergent on $\mathbb{R}$. Find a uniformly convergent subsequence.

(c) Show that $f(x) = \sum\limits_{n=1}^{\infty} \frac{1}{n^2+x^2}$ is differentiable on $\mathbb{R}$. **[6+6+6]**

(4) (a) Find all values of $x \in \mathbb{R}$ for which the series $\sum\limits_{n=1}^{\infty} \frac{n4^n}{3^n} x^n(1-x)^n$ converges.

(b) Let R be the radius of convergence of the power series $\sum\limits_{n=1}^{\infty} a_n x^n$, where $0 < R < \infty$. Find the radius of convergence of $\sum\limits_{n=1}^{\infty} n^n a_n x^n$. **[6+6]**

(5) Let (S_n) be the sequence of partial sums of $\sum\limits_{n=0}^{\infty} a_n$ and let $T_n = \frac{1}{n+1}(S_0 + S_1 + \cdots + S_n)$. Prove that if (T_n) is a bounded sequence, then the power series $\sum\limits_{n=0}^{\infty} a_n x^n$, $\sum\limits_{n=0}^{\infty} S_n x^n$ and $\sum\limits_{n=0}^{\infty} (n+1) T_n x^n$ converge for $|x| < 1$. **[12]**

(6) Let $f(x) = \frac{\pi - x}{2}$ for $x \in (0, 2\pi)$, $f(0) = 0$, and extend this function 2π-periodically to $\mathbb{R}$. Use the Fourier series of f to prove that $\frac{\pi - x}{2} = \sum\limits_{n=1}^{\infty} \frac{\sin nx}{n}$. Also, for $x \in (0, 2\pi)$, find the sum of the series $\sum\limits_{n=1}^{\infty} \frac{\cos nx}{n^2}$. **[12]**

(7) (a) Find the Fourier series of the 2π-periodic function f whose values on the interva $[-\pi, \pi)$ are given by

$$f(x) = \begin{cases} -1, & -\pi \leq x \leq 0, \\ 1, & 0 < x < \pi. \end{cases}$$

Also show that $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$.

(b) Let $d_n(x) = \sum_{k=1}^{n} \sin kx$. Let $f : \mathbb{R} \to \mathbb{R}$ be a 2π-periodic function which is Riemam integrable over $[-\pi, \pi]$. Show that

$$\sigma_n f(x) = -\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) d_n(t - x)\, dt.$$

(c) Let $f : \mathbb{R} \to \mathbb{R}$ be a 2π-periodic function which is Riemann integrable over $[-\pi, \pi]$. I $f(x+)$ and $f(x-)$ are both finite and

$$\lim_{t \to 0+} \frac{f(x+t) - f(x+)}{t} \quad \text{and} \quad \lim_{t \to 0+} \frac{f(x-t) - f(x-)}{t}$$

both exist and are finite, then show that the Fourier series of f at x converges t $\frac{1}{2}[f(x+) + f(x-)]$. **[7+7+7**

INDIAN STATISTICAL INSTITUTE

Backpaper Examination: 2010-2011
B.Stat. (Hons.) 1st Year. 1st Semester
Probability Theory I

Date: 20.1.2011 **Maximum Marks: 100** **Duration: 4 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.
- Whenever applicable, you should (1) write clearly, explaining all your notations, a suitable sample space Ω for the answers, and (2) state, with adequate justification, assignment of probability to the sample points.

---

1. If n distinguishable balls are placed at random into n cells, find the probability that exactly one cell remains empty. [9]

2. Suppose that Ω is a finite sample space, and that $A_1, \ldots, A_n$ are mutually independent events such that $0 < P(A_i) < 1$ for all i. Show that $\# \Omega \geq 2^n$, the equality being attainable. [9+3 = 12]

3. Consider the experiment of permutating randomly a decks of N cards. Suppose that this experiment is conducted independently thrice. Denote by p_m, the probability of having exactly m positions at each of which the same card appears in all three experiments. Show that $p_m = \sum_{j=0}^{N-m} (-1)^j \ (N-m-j)!/[m!j!N!]$. [12]

4. Suppose F is a distribution function. Let $D(F) := \{t \in \mathbb{R} : F \text{ is discontinous at } t\}$. Show that $D(F)$ is at most a countably infinite set. [10]

5. Suppose that $X_i \sim \text{Geometric}(p_i)$, $i = 1, \ldots, k$, are independent random variables, $0 < p_i < 1 \; \forall \, i$. Show that $P(\min(X_1, \ldots, X_k) = X_1) = p_1/[1 - (1-p_1) \cdots (1-p_k)]$. [10]

6. Suppose that $X_1, \ldots, X_k$ are independent random variables such that X_i $(i = 1, \ldots, k)$ has negative binomial distribution with parameters α_i and p, $\alpha_i > 0 \; \forall \, i$, $0 < p < 1$. Use tools of probability generating function to find the distribution of $\sum_{i=1}^{k} X_i$. [10]

P.T.O.

7. Suppose that $X_i \sim \text{Poisson}(\lambda_i)$, $i = 1, \ldots, k$, are independent random variables, $\lambda_i > 0 \; \forall \, i$. Let $1 \leq j < k - 1$. Find the conditional distribution (pmf) of $(X_{j+1}, \ldots, X_k)$ given $X_1 = x_1, \ldots, X_j = x_j, X_1 + \cdots + X_k = n$. [10]

8. Consider the problem of random distribution of r distinguishable balls in n cells, and assume that each arrangement has probability n^{-r}. For $i = 1, \ldots, n$, let $X_i := 1$ or 0 according as the i-th cell is occupied or empty. Find the probability mass function (pmf) of $(X_1, \ldots, X_n)$. [11]

9. Consider the random experiment of distributing randomly n letters in n directed envelopes, one letter in one envelope. Find the expectation and variance of the number of wrongly distributed letters. [8+8 = 16]

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

INDIAN STATISTICAL INSTITUTE
FIRST SEMESTER BACK PAPER EXAMINATION
(2010-2011)
B. STAT (First Year)
Computing Techniques and Programming I
Full Marks-100, Duration-Three hours

Date : 12.1.11

Note: Answer all the Questions.

1. (a) Give the algorithm/flowchart of producing a 3×3 Magic Square where the entire row sums, column sums, and diagonal sums are equal. You have to give the dry run of your algorithm.
(b) (I) Give the corresponding C program of the above algorithm/flowchart with full documentation.
(II) Give also the C program to test in all possible ways whether a given square is really a magic square. [8+ (6+6)]

2. (a) With the help of C program you have to prepare two files (input file and output file) which have to be program controlled. In the input file go on adding arbitrary real numbers in the form of two-dimensional array, say $m\times n$. In the output file print all the elements of the said two-dimensional array ($m\times n$) and also the average of all the numbers.
(b) (I) Write a C program to count the number of characters including the blank spaces in a given text.
(II) Write a C program to count the number of words in the above text file [10+ (4+6)]

3. (a) Given any real number x, you have to calculate:
$p_0(x) = a_0 + a_1x + a_2x^2 + a_3x^3$
$= a_0 + x.p_1(x)$, where $p_1(x)$ is recursively defined as $p_1(x) = a_1 + x.p_2(x)$
and in general $p_n(x) = a_n + x.p_{n+1}(x)$.
For the purpose of the calculation of $p_0(x)$, you have to use one '**for**' loop and one given value of x.

(b) (i) What do you understand by big O in the context of the complexity of algorithms?

P.T.O.

Most of the classical sort algorithms take the time ranging from $O(n.\log n)$ to $O(n^2)$ where n denotes the number of elements to be sorted. Consider two different values for n. The first value is n_0 whereas the second one is $100.n_0$. What would be the effect of the two values on $O(n.\log n)$ to $O(n^2)$?

(ii) For a cyclic redundancy code, an n-bit encoded message (where k is the number of message bits) has to be formed. What should be the length of check bits?

(iii) For four message bits 1100 and the generator polynomial x^3+x+1, find the encoded message and hence explain how you would proceed to locate all the single bit errors.

[6 + (6 +2 +6)]

4. (a) Distinguish between

(i) Recursive algorithm and Iterative algorithm

(ii) Subroutine call and Interrupt scheme

(b) What is a concurrent program? Distinguish between a procedure call and a process creation. [10 + 10]

5. (a) In the theory of 2's complement arithmetic describe all the cases of addition of two numbers (which may be positive or negative).

(b) We find that in two occasions we encounter overflow or underflow. In the event of any such occasion how would it be notified within the computer for the purpose of creating an interrupt?

[10+10]

INDIAN STATISTICAL INSTITUTE

Vectors and Matrices I : B. Stat 1st year: Back paper Examination: 2010-11

Date..10..1..11

Maximum Marks 45 | Maximum Time 3 hrs.

Answer all questions.

(1) Let $\mathcal{P}_n$ be the vector space of all polynomials of degree $< n$, in one variable and with real coefficients. Let $a_1, a_2, \ldots, a_n$ be distinct real numbers. Suppose that

$$P(x) = (x - a_1)(x - a_2)\ldots(x - a_n) \text{ and } Q_i(x) = \frac{P(x)}{(x - a_i)}.$$

Show that $Q_1, Q_2, \ldots, Q_n$ form a basis of $\mathcal{P}_n$. 10

(2) Let $W = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid x_1 = x_2, x_3 = x_4 = x_5\}$ be the subspace of $\mathbb{R}^5$. Prove that there does not exist a linear map T from $\mathbb{R}^5$ to $\mathbb{R}^2$, such that its null space $N(T) = W$. 7

(3) Let V be a finite dimensional vector space and W be a subspace of V. Let $\dim V = n$ and $\dim W = m$ with $n > m$. Find the dimension of the quotient space V/W. 8

(4) Let $Ax = b$ be a system of linear equations where A is an $m \times n$ real matrix. Show that it has a unique solution if and only if Rank A = Rank $[A : b] = n$. 10

(5) Let A be an $m \times n$ real matrix.

(a) Suppose that rank of A is n. Show that G is a generalized inverse of A if and only if $GA = I$.

Suppose that $m = n$ for questions (b), (c) and (d).

(b) If $A^T = A$, then show that A has a generalized inverse G such that $G^T = G$.

(c) If H_1 and H_2 are two matrices in HCF such that A is row-equivalent to both of them, then show that $H_1 = H_2$.

(d) If the rank of A is equal to the rank of A^2 and if $A = PQ$ is a rank factorization of A, then prove that QP is invertible. 5+5+5+5

(6) Let A, B, C be three matrices of the same order. Suppose that the row-space of A is a subspace of the row space of B and the column space of C is a subspace of the column space of B. Let B_1 and B_2 be two generalized inverses of B.

Show that $AB_1C = AB_2C$. 10

P. T. O.

(7) Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map and dimension of the image of T is $r > 0$. Show that there are ordered bases X and Y of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively such that the matrix of T with respect to X in the domain and Y in the range is

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix},$$

where by O we denote the zero matrices of the appropriate order. 10

(8) Let A be 3×5 matrix which is reduced to

$$H = \begin{bmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

by the following row operations (in that order) $R_{1,2}, R_1(1/2), R_{2,1}(-1), R_{3,1}(-1/2), R_{2,3}, R_2(3/4), R_{3,2}(1/2)$.

(a) Get back A.

(b) Obtain a rank factorization of A (using H).

(c) Find a generalized inverse of H and that of A. 5+5+5

(9) Let $T : V \longrightarrow V$ be a linear transformation where V is a n dimensional complex vector space. Fix an ordered basis $\{x_1, x_2, \ldots, x_n\}$. Suppose that T is upper triangular with respect to this basis. Show that if T is not injective then for some i $(1 \leq i \leq n)$ $Tx_i \in \text{Span}\{x_1, \ldots, x_{i-1}\}$. 10

Computational Techniques and Programming II

Indian Statistical Institute, Kolkata
B.Stat (hons.) I
(2010-11)
Semester 2
Back paper examination

Date: 01-07-11 Duration: 3hrs.

This paper carries 100 marks. Attempt all problems. This is a closed note, closed book examination. You may use your own calculator. Laptops are not allowed. If you think that there is a mistake in some problem you must justify your point to get credit for that problem.

1. Prove that the Gauss-Jacobi iteration for solving

$$A_{2\times 2}\mathbf{x} = \mathbf{b}$$

converges for any initial value, if A is a positive definite matrix. Also construct an $A_{2\times 2}$ where the Gauss-Seidel method fails to converge.[10+5]

2. A user has the following 4 files in the same folder. Explain the error that she will encounter when she tries to compile and run `main.c`.

a.h
```
int i;
```

b.h
```
#include "a.h"
i = 7;
```

c.h
```
#include "a.h"
i = 8;
```

main.c
```
#include "b.h"
#include "c.h"

main() {
   printf("i = %d\n",i);
}
```

Correct the error by modifying *only* the file `a.h`. Justify your procedure. Explain the output of the corrected program. [3+8+4]

3. Let $x_1, ..., x_n$ be the roots of the n-th orthogonal polynomial under the inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx.$$

Show that there exist numbers $a_1, ..., a_n$ such that

$$\int_0^1 x^k dx = \sum_{i=1}^n a_i x_i^k$$

for $k = 0, ..., 2n-1$. [10]

(16)

4. Numerically find the *global* maximum value of $5\log p + 3\log(1-p) + 4\log(1-\frac{p}{2})$ for $p \in (0,1)$. It is enough to check convergence up to 4 decimal places. Show the iterations. Justify your answer. [15]

5. Let $x_i = \frac{i}{2}$ and $y_i = \Phi(x_i)$ for $i = 0, 1, 2$, where $\Phi(x)$ denotes the $N(0,1)$ c.d.f. Let $p(x)$ be the polynomial interpolating (x_i, y_i)'s. Obtain an upper bound (possibly involving x) on $|p(x) - \Phi(x)|$ based on Newton's difference method. Clearly mention any theorem you are using. [15]

6. Write an R function called `bisect(f,a,b,maxIter,eps)` that can be used to solve $f(x) = 0$ for $x \in [a,b]$ by bisection method. Your function should check appropriate conditions on $f(a), f(b)$. Here `maxIter` is the maximum number of iterations allowed, and `eps` is the error limit. [15]

7. Suppose that you have a software to compute QR decomposition of any full column rank matrix A. Consider the (possibly inconsistent) system

$$A\mathbf{x} = \mathbf{b},$$

where A is full column rank. Let $\hat{\mathbf{x}}$ be the least squares solution. Show how you can use QR decomposition of A to compute $\|\mathbf{b} - A\hat{\mathbf{x}}\|^2$ without any need to compute $\hat{\mathbf{x}}$ explicitly. [15]

Good Luck!!